

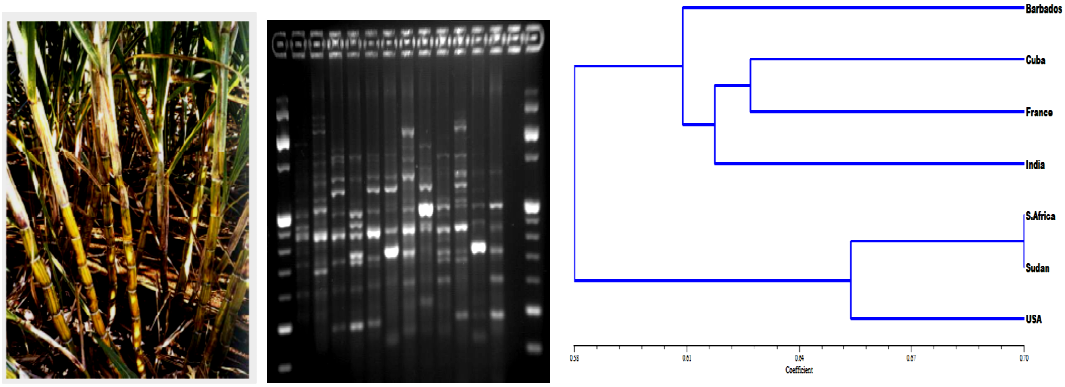
ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

INSTITUTE OF BIOTECHNOLOGY

Molecular Characterization of Introduced Sugarcane (*Saccharum officinarum* L.) Genotypes in Ethiopia Using Inter Simple Sequence Repeat (ISSR) Markers

BY: ADEBABAY TAZEB



A Thesis Submitted to the School of Graduate Studies, Addis Ababa University, in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biotechnology

ADDIS ABABA, ETHIOPIA

MARCH, 2015

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

INSTITUTE OF BIOTECHNOLOGY

This is to certify that the thesis prepared by Adebabay Tazeb, entitled: **Molecular Characterization of Introduced Sugarcane (*Saccharum officinarum* L.) Genotypes in Ethiopia Using Inter Simple Sequence Repeat (ISSR) Markers** and submitted in fulfillments of requirements for the degree of Master of Science in Biotechnology complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by Examining Committee:

Examiner: Dr. Kifle Dagne Signature_____ Date _____

Examiner: Dr. Tigist Wondimu Signature_____ Date_____

Advisor: Dr. Teklehaimanot Haileselassie Signature_____ Date_____

Advisor: Dr. Kassahun Tesfaye Signature _____ Date _____

Title: Molecular Characterization of Introduced Sugarcane (*Saccharum officinarum* L.) Genotypes in Ethiopia Using Inter Simple Sequence Repeat (ISSR) Markers

By: Adebabay Tazeb

Year: 2015, Addis Ababa University

ABSTRACT

The major thrust of sugarcane (*Saccharum officinarum* L.) variety improvement programs is to increase sugar yield, while due to its genetic complexity, sugarcane has received very little research interest, despite its economic importance, and molecular techniques are being developed only in recent times. Furthermore, until the present date no molecular studies were carried out to evaluate sugarcane germplasm from Ethiopian fields. Therefore, the objective of this study was to evaluate genetic diversity and establish relationships among populations of sugarcane in Ethiopia using ISSR markers. Genomic DNA was extracted from leave samples of 82 sugarcane genotypes according to modified CTAB method. A total of 149 scorable and reproducible bands were generated using 12 ISSR primers among which 124 were polymorphic and attributed to PPL = 83.22 %, h = 0.31, and I = 0.45 at species level. Intra-population diversity based on polymorphic loci ranged from 28.86 % to 47.65 % with mean of 38.35 %, Nei's gene diversity of 0.097 to 0.171 with mean of 0.137, Shannon's information index of 0.147 to 0.255 with mean of 0.205, and within population AMOVA of 63.56 %. All diversity parameters confirm that, the highest diversity was obtained from those that were obtained from France and Cuba whilst the lowest was from those of Barbados and South Africa. AMOVA revealed 36.4 % among population variation which is significantly lower than that of within population variation. From Jaccard's pairwise similarity coefficient, those from Sudan and South Africa were most related populations exhibiting 0.70 similarity and USA and Barbados were the most distantly related populations with similarity of 0.53. With all clustering analysis, most of the genotypes clustered to their respective origins. Thus, ISSR markers detected a range of diversity from sugarcane varieties with their unique identity that deserve conservation attention and improvement programs.

Key words: Commercial crop, genetic diversity, ISSR markers, *Saccharum officinarum*

ACKNOWLEDGEMENTS

I thank the almighty whose blessings have enabled me to accomplish my thesis work successfully. Next it is my proud privilege to express my sincere thanks and deep sense of gratitude to my advisors Dr. Teklehaimanot Haileselassie and Dr. Kassahun Tesfaye for their unreserved guidance, encouragement, invaluable advices, and relentless corrections of this thesis from preparation of the proposal and the conceptualization of its structure up to the final accomplishment of the thesis. Their constant encouragement and confidence imbibing attitude has always been a moral support for me and also their politeness, friendly approach, motivation, and professional support throughout my study were indeed special. I am also very grateful to Dr. Kassahun Tesfaye for offering his laboratory facilities and follow up in each and every Laboratory procedures was so special and unforgettable.

I would like to acknowledge Ethiopian Sugar Corporation, specially Research and Training Directorate, for their financial support and allowed to provide plant material for molecular analysis. Next I would like to extend my heartfelt gratitude to Mr. Feyissa Tadesse, director of varietal development at Wonji Sugar Factory, for his unlimited invaluable advices and totally for the success of this thesis work.

I want to express my deep gratitude to Addis Ababa University for accepting me to the graduate program of the Institute of Biotechnology and the staff members of the University had been overall helpful to me in many ways.

I am pleased to acknowledge my home University (Madawalabu University) for giving me study leave, at Addis Ababa University.

I also extend my cheerful thanks to all my classmates and friends who helped me in various aspects, especially Haftom Birhane, Aleshenafi Alemu, Edossa Fikiru, Hawi Negussie, Kedra Mohammed, and Semegn should deserve thanks. Last, but not the least, I appreciate and thanks my beloved parents and family members for their encouragement and continuous support they gave me during my entire stay in the University.

Table of Contents **page**

LIST OF FIGURES ix
LIST OF TABLES x
LIST OF ABBREVIATIONS xi
1. INTRODUCTION 1
2. LITERATURE REVIEW 6
2.1. Origin and distribution of sugarcane 6
2. 2. Domestication of sugarcane 9
2.3. Taxonomy and botanical descriptions of sugarcane 11
2.3.1. Taxonomy 11
2.3.2. Botanical descriptions 13
2.4. Polyploidy and genetic base of sugarcane 14
2.5. Ecology of sugarcane 15
2.6. Mode of propagation 16
2.7. Economic importance 17
2.8. Importance of genetic diversity analysis in sugarcane germplasm 18
2.9. Genetic distance and genetic similarity in sugarcane varieties 20
2.10. Genetic markers and their applications for genetic diversity analysis in plants 22
2.10.1. Morphological markers 23
2.10.2. Biochemical markers 24
2.10.3. Molecular markers 26
2.10.3.1. Non PCR-based molecular marker (hybridization-based marker) 27
2.10.3.1.1. Restriction Fragment Length Polymorphism (RFLP) 27
2.10.3.2. PCR-based molecular markers 29
2.10.3.2.1. Random Amplified Polymorphic DNA (RAPD) 29

2.10.3.2.2. Amplified Fragment Length Polymorphism (AFLP)..... 31
2.10.3.2.3. Microsatellite or Simple Sequence Repeat (SSR)..... 32
2.10.3.2.4. Inter Simple Sequence Repeat (ISSR) 34
3. OBJECTIVE OF THE STUDY 39
3.1. General objective..... 39
3.2. Specific objectives..... 39
4. MATERIALS AND METHODS..... 40
4.1. Plant material and sampling strategy 40
4.2. Genomic DNA extraction..... 41
4.3. Concentration and quality checking of DNA..... 43
4.4. Screening and optimization of ISSR primers..... 43
4.5. PCR amplification and gel electrophoresis 46
4.6. Band scoring and data analysis 48
5. RESULTS 50
5.1. ISSR primer variation and level of polymorphism 50
5.2. Level of genetic diversity in sugarcane populations..... 56
5.3. Analysis of Molecular Variance (AMOVA)..... 62
5.4. Genetic similarity among the sugarcane populations..... 64
5.5. Clustering analysis 67
5.6. Principal Coordinate Analysis (PCoA) 71
6. DISCUSSION..... 74
6.1. ISSR polymorphism and application for genetic diversity study in sugarcane..... 74
6.2. Diversity among sugarcane genotypes in Ethiopia 77
6.3. Genetic distance among sugarcane genotypes and its implications for varietal improvement..... 80

6.4. Genetic differentiation and structure among sugarcane genotypes 82
6.5. Genetic relationships in sugarcane genotypes 86
7. CONCLUSION 89
8. RECOMMENDATIONS 91
9. REFFERENCES 92
10. APPENDICES 109

LIST OF FIGURES

Figure 1. Map of origin and distribution of sugarcane……8

Figure 2. Hypothetical pathway for sugarcane evolution and domestication……10

Figure 3. ISSR profiles of sugarcane individuals generated by di-nucleotide primers (810 (A)) and (811 (B)) using agarose gels…….53

Figure 4. ISSR profiles of sugarcane individuals generated by tri-nucleotide primer (866 (D)) and tetra-nucleotide primer (873 (E)) using agarose gels……54

Figure 5. ISSR profiles of sugarcane individuals generated by penta-nucleotide primers (880 (F) and 881 (G)) using agarose gels.……55

Figure 6. UPGMA based dendrogram for the seven populations of sugarcane.……..66

Figure 7. Dendrogram obtained with UPGMA based on Jaccard's similarity coefficients for 82 genotypes of sugarcane using 12 ISSR primers……..68

Figure 8. NJ based analysis of 82 genotypes of sugarcane using 12 ISSR primers after pairwise comparisons using Jaccard's similarity coefficient.……70

Figure 9. Two dimensional (2D) representation and relationships of 82 genotypes of seven sugarcane populations on Principal Coordinate analysis based on Jaccard's similarity coefficients……72

Figure 10. Three dimensional (3D) representation of Principal Coordinate analysis of genetic relationships among 82 genotypes of sugarcane as revealed by ISSR marker generated by 12 primers.……73

LIST OF TABLES

Table 1. Number of genotypes of sugarcane used for molecular diversity analysis……41

Table 2. Total ISSR primers used for initial screening and PCR optimization……45

Table 3. Selected ISSR primers used for total amplification of all DNA samples.……46

Table 4. Lists of all primers used in the analysis, primer sequence, estimated band size range, number of amplified loci (NAL), number of polymorphic loci (NPL), percent of polymorphic loci (PPL), level of gene diversity (h), and Shannon's information index (I).……52

Table 5. Percent of polymorphic loci (PPL), number of polymorphic loci (NPL) population differentiation (G_{ST}), and gene flow (Nm) estimate in sugarcane populations. The analysis was carried out at country level and grouping was made based on geographic origin.……59

Table 6. Nei's gene diversity (h) in sugarcane populations. The analysis was carried out at country level and grouping was made based on geographic origin.……60

Table 7. Shannon's information index (I) in sugarcane populations. The analysis was carried out at country level and grouping was made based on geographic origin.……61

Table 8. Partitioning of the genetic variation into among groups, among populations within groups, and within populations by analysis of molecular variance (AMOVA) based on 149 ISSR bands generated from sugarcane.……63

Table 9. Partitioning of genetic variations into within and among populations by analysis of molecular variance (AMOVA) based on 149 ISSR bands generated from sugarcane.……63

Table 10. Similarity matrix among the sugarcane populations obtained by Jaccard's similarity coefficient……65

LIST OF ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
bp	base pair
CTAB	Cethyl Trimethyl Ammonium Bromide
dNTP	deoxy Nucleotide Tri-Phosphate
EDTA	Ethylene Diamine Tetra Acetic acid
h	Nei's gene diversity
I	Shannon's Information index
ISSR	Inter Simple Sequence Repeats
m.a.s.l	meter above sea level
NJ	Neighbor Joining
NPL	Number of Polymorphic Loci
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
PPL	Percent of Polymorphic Loci
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SSR	Simple Sequence Repeats
TBE	Tri-Borate-EDTA buffer
TE	Tris-EDTA buffer
UPGMA	Unweighted Pair Group Method with Arithmetic Mean

1. INTRODUCTION

Sugarcane (*Saccharum officinarum* L., 2n = 80) belongs to the genus *Saccharum* of the family Poaceae within the tribe Andropogoneae (Phillips, 1995). It is an economically important tropical grass that accounts for 70 % of the world's sugar production (Cunff *et al*., 2008). It recently has gained increased attention because of the potential of its by-products; ethanol, molasses, and bagasse, as important renewable biofuel sources. This tropical plant has a C_4 carbohydrate metabolism which, allied with its perennial nature, makes it one of the most productive cultivated plants (Cunff *et al*., 2008), with an average yield of cane stalk is 60-70 tons per hectare per year, but this figure can vary between 30 and 180 tons per hectare depending on knowledge and crop management approach used in sugarcane cultivation. However, sugarcane probably the most complex of all crop genomes is studied to date, mainly due to its very high degree of polyploidy and interspecific origin (Cunff *et al*., 2008). It thus represents a major challenge for genetic studies.

Sugarcane is the world's largest commercial crop by production quantity grown in over 127 countries and is one of the staple sources of sweetening agents in the world; as such it is a very important crop from an agro-economic perspective (Cordeiro, 2001). Brazil is the largest producer of sugarcane in the world and the next five major producers, in decreasing amounts of production are India, China, Thailand, Pakistan, and Mexico. The world demand for sugar is the primary driver of sugarcane agriculture.

Sugarcane production in Ethiopia started in 1954/55 when the Wonji Sugar Factory was commissioned and produced 15,843 tons of sugar in the first campaign. When sugarcane development began in 1951, the company was owned by Dutch Company, HVA (Handlers Vereeniging Amsterdam). The presence of large areas of suitable low lands, adequate water resources, favorable climate, and suitable soil types are all proven to be highly conducive for sugarcane development and productivity in Ethiopia. Experiences of existing sugar factories show that because of the suitable conditions in the country, an average sugarcane production per hectare per month of the land under irrigation is very high as compared to other countries i.e. 9-11 tons in Ethiopia against 6-8 tons elsewhere (EIA, 2008). This would make Ethiopia a very attractive location for private investors to invest in production and processing of sugarcane. Currently, Ethiopia produces 300,000 tons of sugar every year out of the 500,000 tons demand, which only covers 60 % of the annual demand for domestic consumption (ESDA, 2014). The difference is being bridged through importation from abroad.

In Ethiopia, sugar industry has been increasing very promisingly because of favorable government policy for sugar industry. This can be explained by the expansion projects of the previous sugar estates and/or factories, the establishment of other new sugar factories and participation of investors in the industry. Furthermore, the industry has become one of the major pillars of the government's five year economic Growth and Transformation Plan (GTP). In the coming few years, it is expected that the sugar estates will be expanded to over 300,000 hectares and 10 new sugar factories will be built where some of them will start production very soon. With this,

the annual sugar production of Ethiopia is expected to reach 2.25 million tons at the end of 2015 which will suffice local demand and will be exported (ESDA, 2013). Due to this, Ethiopian Sugar Corporation is working vigorously to raise the current production of sugar in the country and is engaged in building new sugar factories at various regional states as well as conducting expansion work at the existing ones with the aim to significantly contribute to the overall economic development of the country through satisfying the local demand for sugar supply and seizing a remarkable market share at the international level.

Despite the potential of the country, the per capita annual consumption of sugar in Ethiopia is one of the lowest in the world with about 5.1 kg as estimated by International Sugar Organization (ISO). This is considered very low as compared to 16.3 kg per capita of African average (ISO, 2009). This is quite far below the existing potential being exploited by other countries (ESDA, 2009). This low consumption is associated with low production of sugar in the country, and then low production is because of (i) unavailability of novel, high yielding and disease/pest resistant varieties, (ii) lack of implementation of advanced technologies in sugarcane breeding and production systems, and (iii) genetic improvement of sugarcane has been based mainly on the production of hybrids through conventional breeding methods. However, the genetic complexity, narrow genetic base, susceptibility to various diseases or pests, and the long breeding/selection cycle make the improvement of this crop through conventional breeding laborious (Lakshmanan *et al*., 2005).

In spite of its commercial importance, molecular characterization studies of sugarcane in Ethiopia have not yet been conducted for conservation and sustainable utilization. However, we anticipate that the rapid advancements in molecular biology and emerging modern biotechnology innovations would play a significant role in the future sugarcane crop improvement programs and offer many new opportunities to develop it as a new generation of industrial crop (Lakshmanan *et al*., 2005). Therefore, this study is the first of its kind to characterize sugarcane genotypes using molecular marker technique for future breeding and sustainable utilization of its genetic resources.

Molecular characterization of sugarcane germplasm therefore, provides essential information on the extent of genetic diversity among parental lines and the difference in genetic background among germplasm for breeders to utilize it in sugarcane improvement programs (Singh *et al*., 2010). The knowledge on phylogenetic relationship among genotypes of sugarcane or germplasm collection will guide parental selection for the development of new varieties (Ming *et al*., 2006). Hence, genetic diversity analysis of sugarcane genotypes using molecular technologies is very crucial for the success of sugarcane breeding. However, the assessment of genetic diversity with morphological markers is often difficult and unreliable, particularly for closely related varieties, because of (i) the influence of environmental or management practices on the character and (ii) the limited diversity among varieties with a highly similar pedigree (Morell *et al*., 1995). The problems associated with phenotypic characterization have fostered the need for an accurate, fast, reliable, and cost effective genetic method to distinguish between sugarcane varieties, verify

pedigrees, and estimate genetic diversity. Moreover, such a technique would provide impetus for the development of new and superior varieties.

In the present study, Inter Simple Sequence Repeat (ISSR) markers were used for the investigation in an attempt to assess their potential as molecular marker system to evaluate the genetic diversity among and within sugarcane genotypes. ISSR is found to be quick and cost-effective technique based on PCR amplification of inter-microsatellite sequences to target multiple loci in the genome (Kassahun Tesfaye *et al*., 2014) and in this technique, no DNA sequence information for primer construction is needed (Zietkiewicz *et al*., 1994). Since the analytical procedures include PCR, only low quantities of template DNA are required (5-50 ng per reaction). Furthermore, ISSRs are randomly distributed throughout the genome and is a multi-locus marker system so that it can randomly sample the whole genome and gives multiple fragments. ISSR analysis can be applied in studies involving genetic identity, parentage, clone and strain identification, and taxonomic studies of closely related species (Zietkiewicz *et al*., 1994; Esayas Aga *et al.*, 2005). Moreover, ISSR is very important marker system in discriminating within and among population genetic variation without prior DNA information especially for crops that are very difficult to classify morphologically. Therefore, this study aims at using Inter Simple Sequence Repeat (ISSR) molecular marker based technique to study molecular diversity of sugarcane genotypes representing different sampling sites and provide baseline information for future breeding and improvement of the crop.

2. LITERATURE REVIEW

2.1. Origin and distribution of sugarcane

The center of origin of sugarcane was in dispute until Artschwager and Brandes (1958) exhaustively surveyed the evidence and concluded that it originated in New Guinea, where from in ancient times, various forms of this thick, tall, and tropical cane has been grown as a native domestic garden crop for chewing. Prior to this, Barber (1931) had discussed about the origin of sugarcane from historical, linguistic, and botanical perspectives. The derivation of all names, for sugar and sugarcane is from the Arabic word Sanskrit "Sharkara" meaning 'white sugar' but this terminology has Indian imitation on the basis of its meaning 'a new crop from the east'. This is a linguistic evidence of the Indian origin of the species according to the author (Barber, 1931). Barber also concludes that botanical studies indicate two separate classes of cane with differing origins: the thin Indian indigenous canes from North India and tropical forms from space islands of Oceania, with New Guinea as a possible nucleus.

Finally, after much examination and discussion of available evidence by Artschwager and Brandes, their firm conclusion was that New Guinea is the undoubted center of origin of the species and it represents the generally accepted view today on existence of wild relative as main criteria. However, the genus *Saccharum* probably originated before the continents assumed their current shapes and locations. The genus consists of 35-40 species and has two centers of diversity: the Old World (Asia and Africa)

and the New World (North, Central, and South America) (Cheavegatti-Gianotto *et al*., 2011).

The dispersal of sugarcane forms from New Guinea and its eastern outlier islands began in prehistoric times and still continues. Artschwager and Brandes (1958) distinguish three main movements of sugarcane and first of these brought about the introduction of sugarcane to the Solomon islands, New Hebrides, and New Caledonia from 8000 B.C. onwards. The second dates from 6000 B.C. and took a westerly direction to Indonesia, Philippines, and ultimately to Northern India. The third is considered to have occurred about 600 to 1100 A.D., reaching various island groups east of the New Hebrides, including Fiji, Tonga, Samoa, the Cook Islands, and Hawaii, as well as other parts of Oceania. In general, it has been shown that the general direction of movements of the species in early times was westward from the Pacific (Artschwager and Brandes, 1958).

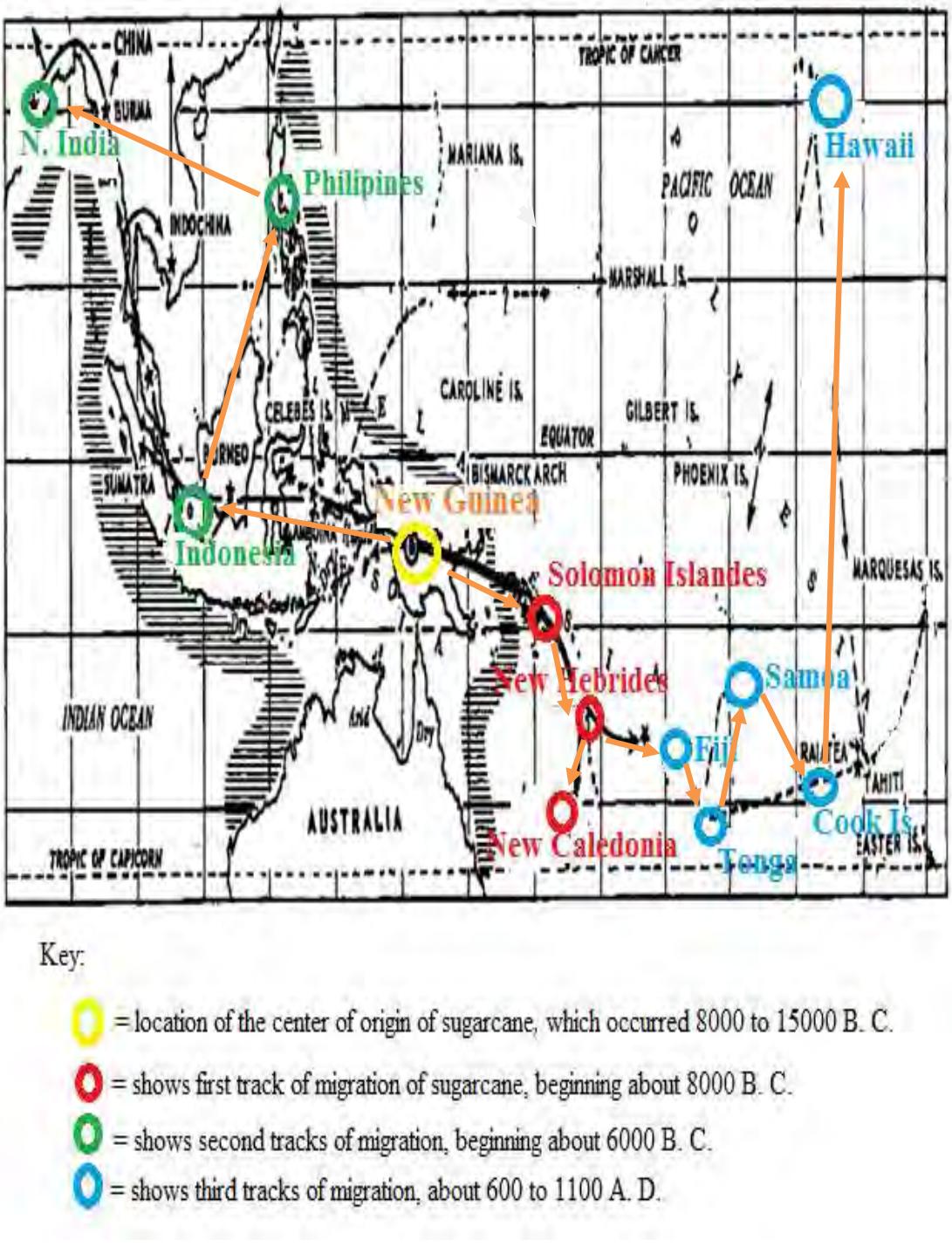


Figure 1. Map of origin and distribution of sugarcane, Source: (Artschwager and Brandes, 1958).

2. 2. Domestication of sugarcane

Domestication described as the process by which genetic changes in wild plants are brought about through a selection process imposed by humans. It is an evolutionary process in which selection (both natural and artificial) operates to change plants genetically (Acquaah, 2007). For the last century, exhaustive investigation was carried out for the domestication and the early evolution of sugarcane since centers of crop domestication are of interest to researchers from various disciplines. Particularly, plant breeders are interested in centers of crop domestication as regions of genetic diversity and variability are critical to the success of crop improvement (Acquaah, 2007). Therefore, the most popular scenario has been developed by Artschwager and Brandes; *S. officinarum* was domesticated from the wild species of *S. robustum* in New Guinea and was then dispersed in the Pacific and mainland Asia during human migrations. In mainland Asia, they hybridized with local *S. spontaneum* giving rise to North Indian *(S. barberi)* and Chinese *(S. sinense)* cultivars. The direct emergence of sugarcane from *S. robustum* is now generally accepted and it has been hypothesized that *S. robustum* was the result of complex introgressions between *S. spontaneum* and other related genera, particularly *Erianthus* and *Miscanthus* (Daniels and Roach, 1987). The emergence of sugarcane from a wild species other than *S. robustum* was considered a minor hypothesis (Grivet *et al*., 2004).

Genus *Saccharum* is a well-defined lineage that diverged over a long evolutionary period from the lineages leading to the *Erianthus* and *Miscanthus* genera (Grivet *et al*., 2004). The implication of this is that cultivated sugarcane species most likely

emerged from wild *Saccharum* species (Grivet *et al*., 2004). The members of the *Saccharum* clade therefore, contributed directly to sugarcane cultivars through allopatric speciation of *S. spontaneum* and *S. robustum* (Moore *et al*., 2014). *Miscanthus* and *Erianthus*, previously proposed as contributing to sugarcane, are recognized as sharing common ancestors. Following *Saccharum* speciation, *S. robustum* contributed to cultivars of *S. officinarum* for sugar, *S. edule* for vegetables, and *S. robustum* used for fencing and construction purpose (Moore *et al*., 2014).

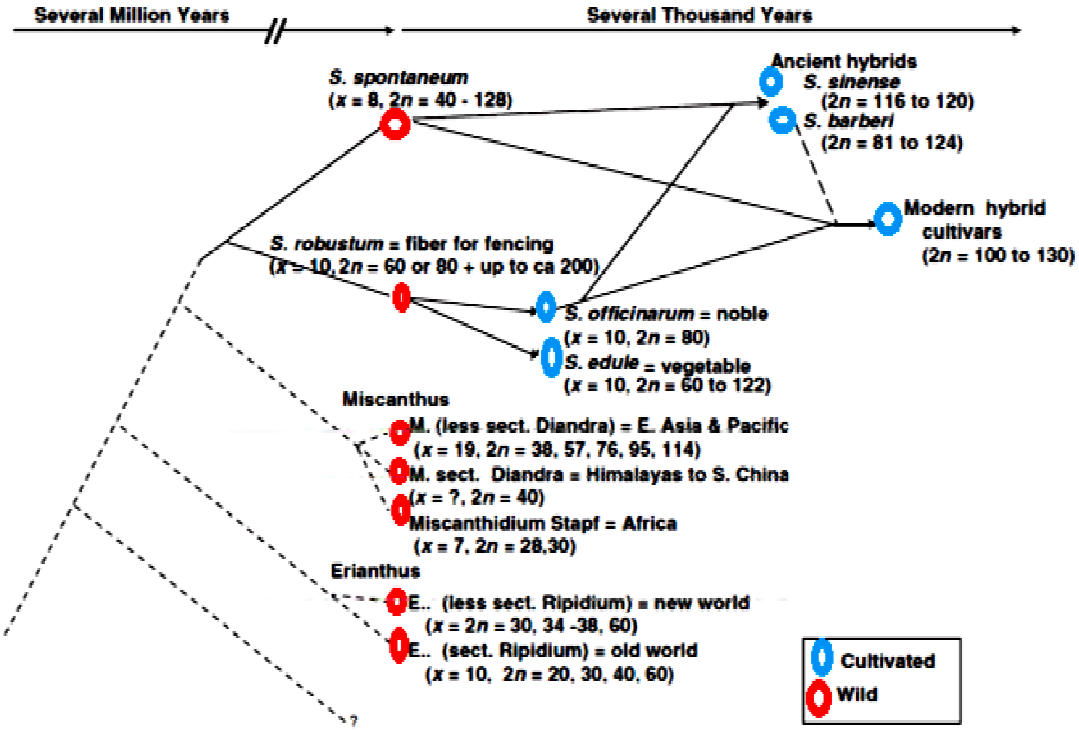


Figure 2. Hypothetical pathway for sugarcane evolution and domestication: strongest evidence for the evolution and domestication are indicated with solid lines; weaker evidence is indicated with dashed lines, Source: (Moore *et al*., 2014).

2.3. Taxonomy and botanical descriptions of sugarcane

2.3.1. Taxonomy

The taxonomy of sugarcane, based on morphology, chromosome numbers, and geographical distribution has been controversial since the original classification of *S. officinarum* by Linnaeus in 1753 (Ming *et al*., 2006). Thus, taxonomic classification of sugarcane can be put as:

Kingdom: Plantae (plants)

Division: Magnoliophyta (flowering plants)

Sub division: Angiospermae

Class: Liliopsida (Monocotyledoneae)

Order: Cyperales/ Glumiflorae (grasses, sedges)

Family: Gramineae (Poaceae) (grasses and bamboos)

Sub family: Panicoideae

Tribe: Andropogoneae

Sub tribe: Saccharinae

Genus: *Saccharum*

Species: *Saccharum officinarum* (sugarcane)

According to its taxonomic classification, sugarcane belongs to the family Gramineae or grass family, which is one of the largest and most important families of flowering plants that have approximately 10,000 species classified into 600 to 700 genera (Ha *et al*., 1999; Moore *et al*., 2014). This makes it the fifth largest family in number of

species in the plant kingdom, but in terms of use by humans it is by far the most important.

The genus *Saccharum* is characterized by complex polyploid-aneuploid genome (Jannoo *et al*., 1999), and formally comprises two wild species and four groups of former cultivated species (D'Hont, *et al*., 1998; Jannoo *et al*., 1999) for being diverse in genome content and organization. These are; (i) *S. officinarum (x* = 10, 2*n* = 80, sweet chewing cane found in native gardens in New Guinea and other South Pacific islands), (ii) *S. spontaneum* (*x* = 8, 2*n* = 40-128, wild cane found throughout Asia), (iii) *S. robustum* (*x* = 10, 2*n* = 60, 80, putative ancestor of *S. officinarum* found most commonly on river banks in Asia and New Guinea), (iv) *S. edule* (2*n* = 60-122, in New Guinea and other South Pacific islands), (v) *S. barberi* (2*n* = 81-124, semi-sweet Indian cane), and (vi) *S. sinense* (2n = 116-120, semi-sweet Chinese cane).

Of the cultivated groups of genus *Saccharum*, *S. officinarum* (*2n* = 80) was named first and is the primary group for production of sugar (Moore *et al*., 2014). *Saccharum officinarum* is also known as "noble cane" due to the sweetness of its stalk juice and the primary source of genes for sucrose accumulation (Lu *et al*., 1994b). A general description of the species is that it often possesses colorful large diameter stalks, broad leaves, short internodes, high sugar content, low fiber content, and is relatively intolerant to the more subtropical environments where sugarcane is commercially grown, especially those where freezes can occur (Moore *et al*., 2014). However, among the wild species, *S. spontaneum* (*2n* = 40-128, with chromosome numbers frequently as multiples of eight), is highly variable in morphology, cytology,

and geographic distribution broadly throughout tropical and subtropical regions. It is a perennial grass, from short grassy appearing narrow-leafed types with no stalks, and very low sucrose content and has contributed toward the development of modern sugarcane cultivars by conferring resistance to most major diseases, providing vigour and hardiness for increased abiotic stress tolerance (such as cold and drought), increasing tillering, and improving ratoon ability (Moore *et al*., 2014).The species *S. sinense*, *S. barberi*, and *S. robustum* also provided minor contributions toward the development of some modern sugarcane varieties.

2.3.2. Botanical descriptions

Sugarcane botany is similar to other grass species, with the major exception that the harvested economic product is sucrose which is stored in the stalk rather than grain or fodder. Sugarcane is a tall perennial tropical grass, which tillers at the base to produce unbranched stems of 2-8 m tall, and of around 5 cm in diameter. It could be called as giant grass that is cultivated for these thick stalks or canes, from which the sugar is, extracted (Miller *et al*., 2009).

The main parts of the sugarcane plant are the stalk, leaf, and root system. The stalk consists of segments called joints. Each joint is made up of a node and an internode. The leaf of the sugarcane plant is divided into two parts: sheath and blade, separated by a blade joint. The sheath, as its name implies, completely sheaths the stalk, extending over at least one complete internode. Two kinds of roots develop from a planted seed piece. The set roots, which arise from the root band, are thin and highly branched; the shoot roots, originating from the lower root bands of the shoots, are

thick, fleshy, and less branched (Miller *et al*., 2009; Verheye, 2010). The inflorescence, or tassel, of sugarcane is an open-branched panicle. The inflorescence, which is known as arrow, emerges above the mass of foliage, presumably an adaptation to wind pollination. The flower is bisexual (i.e. hermaphrodite or perfect flower) (Verheye, 2010) since sugarcane is a cross-pollinating species (Berding *et al*., 2004); natural pollination is assisted by wind and after pollination, it takes 21-25 days for the seed to fill and mature.

2.4. Polyploidy and genetic base of sugarcane

Sugarcane commercial varieties, which is an allopolyploid with genome contributions from *S. officinarum* and *S. spontaneum*, is having high chromosome number of 2n = 100 to 130 in different varieties. *Saccharum officinarum* is an octoploid, having 80 chromosomes with a base chromosome number of x = 10, while *S. spontaneum* forms a polyploid series from 5 to 16 ploid number, with a base chromosome number of x = 8 (D'Hont *et al*., 1998). However, polyploidy is regarded as an important mechanism of speciation and adaptation (Premachandran *et al*., 2011) and there are documented advantages in sugarcane such as (i) heterosis that provides increased vigour, (ii) gene redundancy as a result of gene duplication, and (iii) asexual reproduction that enables the polyploids to reproduce efficiently which are present in allopolyploid sugarcane. Moreover, polyploids are more vigorous due to heterosis and gene redundancy the ability to diversify gene function by altering redundant copies of important genes.

However, in the case of genetic base, most commercial varieties exhibited narrow genetic base and this prompted sugarcane researchers and breeders to be more

conscious of the role of sugarcane germplasm in increasing the variability since the commercial varieties of sugarcane are becoming increasingly similar to each other due to commonness of one or more parents in their ancestry. This has led to the narrowing down of the genetic base of these varieties which eventually creates genetic vulnerability; which refers to the susceptibility of most of the cultivated varieties of sugarcane to disease, insect or some other stress due to similarity in their genotype (Roach, 1989). In general, the sugarcane genetic base is said to be already exhausted or the proper characterization and evaluation is limited. The lack of universal, standard, and systematic characterization and evaluation procedures of sugarcane germplasm is a big reason for the limited use of readily available germplasm. Furthermore, the lack of it is a setback in the assessment of level of the genetic diversity present in the sugarcane germplasm.

2.5. Ecology of sugarcane

Sugarcane grows most successfully in those regions where the climate is more or less tropical but it can grow in subtropical regions as well (Rao, 2007). An average mean temperature of 26-32 ^{0}C is best suited for the growth of sugarcane. Temperatures above 50 ^{0}C arrest its growth; those below 20 ^{0}C slow it down and then generally reduce tillering. The crop does best in the tropical regions receiving a rainfall of 750-1200 mm per annum. Sugarcane requires a long growing season, from 10-12 months, because certain number of heat units is required to bring the plant to maturity (Rao, 2007). In addition to temperature and rainfall, sunlight also affects growth and development of cane. Under bright sunlight conditions the stems are thicker but shorter, and leaves are broader and greener, while under low sunshine, the stems are

slender and long with narrower and yellowish leaves (Rao, 2007). Sugarcane thrives best on fertile, deep, and well aerated loamy soils (Rao, 2007; Verheye, 2010). It can also be raised successfully on lighter soils provided that there are adequate irrigation facilities and there is proper drainage and addition of organic matter on heavy clay soils. However, saline, alkaline, and acidic soils are not at all suitable for this crop (Rao, 2007).

2.6. Mode of propagation

Sugarcane can be propagated from seeds and seedlings selected for the purpose of developing new varieties since sugarcane freely flowers and produces viable seed in tropical climates. However, commercial sugarcane cultivars are highly heterozygous polyploid crops and for all commercial purposes, it is propagated vegetatively using stalk cuttings called 'sett' (seed cane). Due to its high heterozygosity, the resultant F_1 obtained after crossing of the desired parents provides enough variability for an early selection.

In case of sugarcane planting, which consists of two operationally different but sequentially connected labour intensive operations; (i) sett preparation (seed cane cutting, transportation of seed cane, seed cane preparation of 2-3 eye-budded setts, and seed cane treatment using chemicals) and (ii) sett placement (distribution of setts, opening of furrows, placement of setts in furrows by following appropriate geometry, application of fertilizers, covering of setts, and pressing of soil). Sugarcane takes about 40-45 days for germination after planting (Verheye, 2010; Duttamajumder *et al.*, 2011).

The vegetative mode of propagation aids in hassle free maintenance of hybrid vigour of F_1 over clonal generations since the offsprings are genetically identical with their parents and therefore, beneficial genetic traits can be preserved. Another advantage of sugarcane is its capability in giving rise to a stubble/ratoon crop after harvest of the plant crop. This reduces the cost of cane cultivation in terms of land preparation, cost of seed and costs incurred in planting. However, the clonal propagation has inherent weakness of accumulation of cryptic pathogens and pests over time, which severely impairs the productivity potential of a genotype because large number of sugarcane varieties were gradually phased out due to various biotic and abiotic stresses, and the varieties gradually lose their vigour as there is no genetic variation (Duttamajumder *et al.*, 2011). Therefore, sustaining sugar productivity at the existing level necessitated quick replacement of the cane varieties. Even if suitable replacement is available, considerable time elapses before it covers sizeable area, primarily on account of limited availability of healthy seed, due to slow rate of seed multiplication in sugarcane.

2.7. Economic importance

Sugarcane is an important commercial crop which besides serving as food also fulfills requirements of energy and feedstock for industry (Martin *et al.*, 2002; Solomon, 2011). Sugarcane contributes significantly to the sugar production as most of the world's sugar is prepared by using raw material derived from sugarcane (Cunff *et al.*, 2008). Among the by-products of sugarcane, bagasse is mostly used in production of steam and electricity to fulfill the requirements for the sugar industry. Excessive

availability due to industrial wastes has made it an ideal substrate for ethanol production (Solomon, 2011). Production of bagasse along with sugar is more than 200 million tons per year in all sugarcane producing countries, majority of which (about 95 %) is consumed as fuel in the mills thus saving about 40 million tons of oil. The remaining 5 % (that is not used as fuel) is utilized in manufacturing other products like pulp and paper, boards, and animal feeds. The most extensive use of bagasse around the world is for the production of pulp and paper (Solomon, 2011). The molasses, one of the by-products of sugarcane, is widely used in alcohol production, fertilizer, and livestock feed. Molasses is also rich in potash and nitrogen and is widely used for fertilizer synthesis (Solomon, 2011).

2.8. Importance of genetic diversity analysis in sugarcane germplasm

Genetic diversity is the total of heritable variation present within and among populations of organisms that serves an important role in evolution by allowing a species to adapt to a new environment (Kremer *et al*., 1998). The ultimate source of genetic diversity is gene mutation which is a permanent change in the DNA sequence, molded and shaped by selection, recombination, gene flow, genetic drift, and migration in heterogeneous environments in space and time (Hartl and Clark, 1997). Natural selection chooses the best fit among and within a population, thus there can be no adaptive evolution without genetic variation. Hence, genetic diversity is an essential raw material for evolution, which enables populations of the crop species to survive, adapt to new circumstances, and evolve to produce new genetic variants, where some of them may become the fit variants that meet long-term changes in the environment (Hedrick, 2011).

Analysis of the extent and distribution of genetic variation in a crop species is also essential in understanding the evolutionary relationships between cultivars and to sample genetic resources in a more systematic fashion for breeding and conservation purposes (Ejeta *et al*., 1999). Accurate assessment of the levels and patterns of genetic diversity can be invaluable in plant breeding for diverse applications including (i) analysis of genetic variability in cultivars, (ii) identifying diverse parental combinations to create segregating progenies with maximum genetic variability for further selection, and (iii) introgressing desirable genes from diverse germplasm into the available genetic base (Mohammadi and Prasanna, 2003). To achieve the above, genetically diverse germplasm is selected in breeding programs to enhance the productivity and diversity of cultivars. Therefore, a major focus of research in genetics has been to determine the amount of genetic variation in both natural and domestic populations and describing the possible mechanisms of maintaining such variability in meeting new environmental challenges (Weir, 1996).

According to the heterosis theory, parental lines with wider genetic bases must be selected for crossing during variety development on the premise that the main attributes of the parental lines are complementary (Ming *et al*., 2006). However, in terms of genetic diversity within the *Saccharum* genus, numerous studies have shown that sugarcane appears to be the least variable species than other *Saccharum* species (Jannoo *et al*., 1999; Selvi *et al*., 2003; Alwala *et al*., 2006). Selection pressure during domestication for high sucrose types (i.e. severely erodes the genetic variability created from recombination among selected parental clones) and the vegetative mode

of propagation of the crop has been suggested as the major reason for the low genetic diversity within the *S. officinarum* species (Selvi *et al*., 2003; Berding *et al*., 2004).

However, there is molecular evidence that due to its polyploid nature, interspecific origin, and vegetative mode of propagation (i.e., maintenance of hybrid vigour), high levels of heterozygosity exist among modern sugarcane cultivars using RFLP markers (Lu *et al*., 1994b; Janno *et al*., 1999). The major part of diversity was contributed to the 15-25 % chromosome complement that was inherited from *S. spontaneum* by random assortment of half its chromosomes which has the greatest intraspecific diversity (D'Hont, *et al*., 1996). Similar patterns of molecular diversity were also detected using AFLP (Lima *et al*., 2002) and SSR markers (Pandey *et al*., 2011).

In general, genetic diversity information that can be derived from molecular analyses of sugarcane germplasm will help to identify outstanding morphological and agronomic characters present in the germplasm, determine the degree of relationships of the different germplasm within and among breeding populations, and design collection and conservation strategies. Also, this will provide benchmark data for the effective utilization of the germplasm in sugarcane breeding programs and pre-breeding activities.

2.9. Genetic distance and genetic similarity in sugarcane varieties

Genetic distance is a measure of the average genetic divergence between two species or between populations within a species or taxa (Nei, 1987; Souza and Sorrells, 1991) and/or it is the extent of gene differences between cultivars, as measured by allele frequencies at a sample of loci. On the other hand, genetic similarity is the extent of

gene similarities among cultivars (Nei, 1987). However, the distance in gene frequency between the parent genotypes is important because the higher the difference in gene frequency, the higher the amount of heterosis, which indicated that a more distant genetic relationship (Carrera *et al*., 1996). Information concerning genetic relatedness is crucial, for it indicates the rate of adaptive evolution and the extent of response in crop improvement. Furthermore, it is essential as a guideline in the choice of parents for breeding programs and implementing an effective genetic conservation program (Muench *et al*., 1991).

In most cases, modern sugarcane varieties are interspecific hybrids mostly derived from crosses involving *Saccharum officinarum* and *S. spontaneum* (D'Hont *et al*., 1996) but the genetic base of the modern varieties appears to be narrow and this is reflected in the slow progress in sugarcane breeding at present. This could be associated with the limited number of parental clones involved in the primary crosses or their parental breeding lines, used to develop these varieties, are the same or are very close to each other, or use repeated backcrosses with *S. officinarum,* and clonal way of propagation (Nair *et al*., 1999; Nair *et al*., 2002; Rodriguez *et al*., 2005). Therefore, conscious efforts are to be made to diversify the parental genetic base to ensure high genetic variability among the cultivated varieties. Moreover, new sources from the interspecific/intergeneric hybrid gene pool need to be used along with proven parents to generate the variability that will be both commercially viable and genetically diverse. Thus, information on the genetic divergence/similarity of available germplasm is essential for the identification of potential germplasm groups

and for optimizing hybridization and selection procedures (Nair *et al*., 1999; Nair *et al*., 2002).

2.10. Genetic markers and their applications for genetic diversity analysis in plants

A genetic marker is a gene or DNA sequence with a known location on a chromosome that can detect variation in either a protein or DNA sequence. A difference, whether phenotypic or genotypic, may act as a genetic marker if it identifies characteristics of an individual's genotype and/or phenotype, and if its inheritance can be followed through different generations (De Vicente and Fulton, 2003). The analysis of genetic diversity within and among populations routinely involves the use of different genetic markers. Nowadays, genetic markers are used in both basic plant research and plant breeding programs to characterize plant germplasm, for gene isolation, marker assisted introgression of favorable alleles, production of improved varieties (Henry, 2001), and to obtain information about the genetic variation of populations.

Generally, there are two classes of traits to measure genetic variations. These are (i) traits controlled by many gene loci (polygenic traits/minor genes) and (ii) traits controlled at the single gene loci (monogenic traits/major genes). Therefore, to fully characterize the level and pattern of genetic diversity in a species and its evolutionary causes, studies involving both classes of traits are necessary. Basically, there are three classes of genetic markers for genetic diversity analysis. These are morphological,

biochemical, and molecular markers (De Vicente and Fulton, 2003; Semagn *et al.*, 2006; Mondini *et al.,* 2009; Xu, 2010).

In sugarcane, different authors used different genetic markers to characterize genotypes. The most commonly employed ones are: phenotypic characters (Akhtar *et al*., 2006), Allozymes (Glaszmann *et al*., 1989; Eksomtramage *et al*., 1992). Restriction Fragment Length Polymorphism (RFLP) (Lu *et al*., 1994b; Grivet *et al*., 1996; Jannoo *et al*., 1999), Random Amplified Polymorphic DNA (RAPD) (Nair *et al*., 1999), Amplified Fragment Length Polymorphism (AFLP) (Cordeiro, 2001; Lima *et al*., 2002), and Microsatellite or Simple Sequence Repeat (SSR) (Cordeiro, 2001; Chen *et al*., 2009; Glynn *et al*., 2009; Pandey *et al*., 2011; Smiullah *et al*., 2013).

2.10.1. Morphological markers

Morphological traits controlled by a single locus can be used as genetic markers if their expression is reproducible over a range of environments and are often visually described phenotypic characters such as flower color, seed shape, growth habits or pigmentation (Winter and Kahl, 1995). These markers are phenotypic markers with distinguishable traits that are evident to human eyes (Bagali *et al.,* 2010).

Traditionally, genetic diversity within and among populations was determined by assessing differences in morphology. These measures have the advantage of being readily available, do not require sophisticated equipment and are the most direct measure of phenotype, thus they are available for immediate use, which is an important attribute. However, morphological determinations need to be taken by an expert in the species, they are subject to changes due to environmental factors and

may vary at different developmental stages and their number is limited (De Vicente and Fulton, 2003; Mondini *et al.,* 2009). In sugarcane, used morphological markers to study, assessment of adaptation, performance of various cane varieties in different ecologies, and evaluation of agronomic characters before a variety is introduced for commercial cultivation since cane yield is a complex trait which is associated with number of stalks, stalk height, and weight etc (Akhtar *et al*., 2006).

In general, morphological markers cannot be reliable measures of genetic differences because of the influence of the environment on gene expression; therefore methods to assess and detect genetic diversity have been extended to include biochemical and molecular analysis (Singh *et al.,* 2010).

2.10.2. Biochemical markers

To address the limitations of morphological markers, other markers have been developed at both the protein level and the DNA level (De Vicente and Fulton, 2003). Protein markers are usually named 'biochemical markers' but, more than often, they are mistakenly considered as a common class under the so-called 'molecular markers'. Biochemical analysis is based on the separation of proteins into specific banding patterns. It is a fast method which requires only small amounts of biological material. However, only a limited number of enzymes are available and thus, the resolution of diversity is limited (Mondini *et al.,* 2009).

Allozymes, being allelic variants of enzymes, provide an estimate of gene and genotypic frequencies within and between populations (Mondini *et al.,* 2009). This

information can be used to measure genetic diversity, gene flow, genetic structure of species, population structure, and population divergence (Karp *et al.,* 1996). Major advantages of these types of markers include: codominant inheritance, absence of epistatic and pleiotropic effects, ease of use, and low costs of the assay, while limitations include: (i) there are only few isozyme systems per species (no more than 30) with correspondingly few markers, (ii) the number of polymorphic enzymatic systems available is limited and the enzymatic loci represent only a small and not random part of the genome (the expressed part) therefore, the observed variability may not be representative of the entire genome, (iii) although these markers allow large numbers of samples to be analyzed; comparisons of samples from different species, loci, and laboratories are problematic, since they are affected by extraction methodology, plant tissue and developmental stage of the plant (Mondini *et al.*, 2009).

Allozymes analyses in sugarcane have been used to differentiate between wild and noble canes and show progeny-parent relationships (Glaszmann *et al*., 1989). Allozymes research in sugarcane has also revealed that most of the diversity within sugarcane genotypes is related to the presence or absence of S. *spontaneum* genes (Eksomtramage *et al*., 1992). However, Glaszmann *et al*. (1989) found that the use of allozymes/isozymes in sugarcane is often beset with practical difficulties due to the high number of bands that may migrate at similar distances and the occurrence of multiple bands of unequal intensities, both of which arising due to the high ploidy of sugarcane.

2.10.3. Molecular markers

Molecular markers are based on naturally occurring polymorphisms in DNA sequences due to base pair deletions, insertions, and substitutions (Gupta *et al*., 1999). Molecular markers are superior to both morphological and biochemical markers because they are highly polymorphic and heritable (their expression is not affected by environmental variability), relatively simple to detect, abundant throughout the genome even in a highly inbred cultivars, completely independent of environmental conditions, and can be detected at virtually any stage of plant development. However, major disadvantage is the need for technically more complex equipment.

The rapid development of molecular techniques, over the last few decades, now offers a palette of technical approaches for plant genotyping or genome analysis. Which technique is most appropriate depends upon (i) the extent of genetic polymorphism required; (ii) the analytical or statistical approaches available for the technique's application, and (iii) the pragmatics of time and costs of materials (Parker *et al*., 1998).The discovery of the polymerase chain reaction (PCR) was a landmark in molecular marker evolution and has proved to be a unique process for the development and utilization of a battery of new very sensitive and quick approaches, such as AFLP or microsatellites (SSR) (Paglia and Morgante, 1998).

Molecular markers belong to the so-called anonymous DNA marker type and generally measure apparently neutral DNA variation. Suitable DNA markers should represent genetic polymorphism at the DNA level and should be expressed consistently across tissues, organs, developmental stages, and environments.

However, it is extremely difficult to find a molecular marker which would meet all the criteria (properties) required in a study. Depending on the type of study to be undertaken a marker system can be identified that would fulfill at least a few of the characteristics (Weising *et al*., 2005; Xu, 2010). Various types of molecular markers are utilized to evaluate DNA polymorphism generally classified as hybridization-based markers and PCR-based markers.

2.10.3.1. Non PCR-based molecular marker (hybridization-based marker)

In hybridization-based markers, DNA profiles are visualized by hybridizing the restriction enzyme digested DNA, to a labeled probe, which is a DNA fragment of known origin or sequence.

2.10.3.1.1. Restriction Fragment Length Polymorphism (RFLP)

Restriction Fragment Length Polymorphism (RFLP) was developed first and was initially used for human genome mapping (Botstein *et al*., 1980). Later, RFLP markers are one of the most important tools for plant genome mapping (Jiang, 2013), and they are classified as hybridization-based markers. RFLP involves the extraction of genomic DNA followed by its digestion with specific restriction enzymes, which cut the DNA into fragments. RFLP results when there is variation in restriction enzyme cleavage sites, arising due to base substitutions, insertions, deletions or translocations in the genomic DNA (Gupta *et al*., 2002).

The major strength of RFLP markers are their high reproducibility, codominant inheritance, and good transferability between laboratories which provides locus-specific markers that allow synteny (conserved order of genes between related

organisms) studies (Kesawat and Das, 2009; Jiang, 2013). For this, no sequence information is required and they are relatively easy to score due to large size differences between fragments. Still, there are several limitations for RFLP analysis: it requires the presence of high quantity and quality of template DNA. The requirement of radioactive isotope makes the analysis relatively expensive as well as the assay is time consuming and labor intensive (Kesawat and Das, 2009). RFLPs can be applied in diversity and phylogenetic studies ranging from individuals within populations or species, to closely related species. It is widely used in gene mapping studies because of their high genomic abundance due to the ample availability of different restriction enzymes and random distribution throughout the genome (Kesawat and Das, 2009).

RFLP markers have been shown a strong molecular differentiation between *S. officinarum* and *S. spontaneum* (Jannoo *et al*., 1999), and that the major part of the diversity among sugarcane cultivars arises solely from the S. *spontaneum* chromosomes. However, only a limited variability within *S. officinarum* was identified. Lu *et al*. (1994b) shown RFLP markers to be efficient for research on sugarcane genetic diversity and taxonomy and were able to identify the relationships between the different sugarcane cultivars and use the marker as a powerful tool for variety identification. Moreover, RFLP maps have been or are being constructed for numerous crop plants, including sugarcane (Grivet *et al*., 1996), to assess genetic variability (characterization of germplasm stocks), determine correlations between RFLP markers and qualitative or quantitative traits and in some cases to maximize the

benefits of marker-assisted selection or elucidate phylogenetic relationships (Besse *et al*., 1997).

2.10.3.2. PCR-based molecular markers

Polymerase chain reaction (PCR) is an in vitro technique that allows the amplification of a specific DNA region that lies between two regions of known DNA sequence (Saiki *et al*., 1985). This technique allows for the amplification of any DNA sequence of interest to high copy numbers, thereby circumventing the need for molecular cloning (Weising *et al*., 1995). Some of the advantages of PCR-based marker systems are: (i) PCR requires only small amounts of DNA, and often crude miniprep procedures yield DNA of sufficient quantity and quality, (ii) PCR is relatively quick to perform and technically straightforward, once PCR conditions have been established, and (iii) the range of primer sequences possible gives PCR-based techniques great diagnostic power (Morell *et al*., 1995). Various molecular techniques have developed utilizing PCR include: Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and microsatellites or Simple Sequence Repeats (SSR), and Inter Simple Sequence Repeats (ISSR).

2.10.3.2.1. Random Amplified Polymorphic DNA (RAPD)

Random Amplified Polymorphic DNA (RAPD) is based on the amplification of genomic DNA with single primers of arbitrary nucleotide sequence, usually 10 bp long (Williams *et al*., 1990). These primers detect polymorphisms in the absence of specific nucleotide sequence information and the polymorphisms function as genetic markers and can be used to construct genetic maps. RAPD polymorphisms arise when

genomic regions vary for the presence or absence of complementary primer annealing sites due to insertion or deletion between two priming sites, which results in different lengths of the amplification products (Bardakci, 2001). Since most of the RAPD markers are dominant, it is not possible to distinguish whether the amplified DNA fragment is heterozygous (two different copies) or homozygous (two identical copies) at a particular locus (Arif *et al*., 2010).

RAPD analysis has many potential applications, and may be used to assess kinship relationships or genetic diversity (Hadrys *et al*., 1992), construct genetic maps (Bardakci, 2001). The main advantages of the RAPD technology include: (i) suitability for work on anonymous genomes, (ii) applicability to problems where only limited quantities of DNA are available and (iii) low expense (Hadrys *et al*., 1992).

However, there are various limitations and considerations in RAPD analysis, which include specificity of the marker in genome scanning (Hadrys *et al*., 1992), reproducibility of amplification products (Williams *et al*., 1990), and unclear and non-reproducible fragments (Pan *et al*., 1997). Consequently, the analytical power of RAPD markers is not competitive with analysis using sequence information or single locus probe fingerprinting technologies and as such is not suitable for applications such as extensive fingerprinting analysis (Siles *et al*., 2000) or definitive parentage determination. Yet, in sugarcane, RAPDs have been used to evaluate genetic diversity of commercial sugarcane hybrids as well as between members of the *Saccharum* complex, to resolve taxonomical groups in sugarcane cluster analyses (Nair *et al*., 1999).

2.10.3.2.2. Amplified Fragment Length Polymorphism (AFLP)

Amplified Fragment Length Polymorphism (AFLP) method was originally developed as a universal DNA fingerprinting analysis (Vos *et al*., 1995) and is robust and relatively insensitive to PCR reaction conditions. AFLP is a PCR-based marker for the rapid screening of genetic diversity and intraspecific variation. It is a potent fingerprinting technique for genomic DNAs of any origin or complexity and rapidly generates a number of highly replicable markers that allow high resolution genotyping. The strength of AFLPs includes its high genomic abundance, high reproducibility, highly polymorphic, generation of many informative bands per reaction; small amount of template DNA is needed, and the fact that no sequence information for primer construction is required (Saal and Wricke, 2002). Possible reasons for AFLP polymorphisms are: (i) sequence variations in a restriction site, (ii) insertions or deletions within an amplified fragment, and (iii) differences in the nucleotide sequence immediately adjoining the restriction site.

AFLPs have been used for the analysis of genetic diversity (Russell *et al*, 1997), DNA fingerprinting (Powel *et al*., 1996), the construction of linkage maps (Cho *et al*., 1998), and to locate traits of interest (Hartl *et al*., 1999). For genome analysis in sugarcane, AFLPs have proved highly informative as a means of elucidating genetic diversity (Lima *et al*., 2002). With respect to sugarcane, AFLPs have the resolving power to analyze relationships within and between *Saccharum* species, with the additional benefit of being a multi-locus marker system (Cordeiro, 2001). Although AFLP is a powerful molecular marker, some reproducibility issues have been raised in sugarcane, and it is believed that these stem from: (i) partial digestion of the

template DNA as a result of insufficient enzymatic conditions or due to unexpected or inconsistent methylation of template DNA, (ii) poor amplification of PCR fragments, or (iii) DNA contamination (Cordeiro, 2001).

2.10.3.2.3. Microsatellite or Simple Sequence Repeat (SSR)

Microsatellite DNA markers, also known as Simple Sequence Repeats (SSRs), are one of the most powerful molecular genetic marker classes. Simple Sequence Repeats (SSRs), are very short motifs (about 1-6 bp) usually characterized by a high degree of repetition and occur at many thousand loci in the nuclear genome (Singh *et al*., 2010). They are ubiquitous and highly polymorphic, owing to the mutation affecting the number of repeat units. The hypervariability of SSRs among related organisms makes them an informative and excellent choice of markers for a wide range of applications in population and evolutionary biology (Varshney *et al*., 2007; Chen *et al*., 2009), which include estimate genetic diversity (Hokanson *et al*., 2001), identify and test the paternity of cultivars (Rosa *et al*., 2004), study population structure and gene flow (Schueler *et al*., 2006), and develop gene mapping (Hayden *et al*., 2006).

SSR polymorphism reflects differences in simple repetitive sequences of defined regions of the genome. With the advent of polymerase chain reaction (PCR) technology, this property was converted into a highly versatile molecular marker (Weber and May, 1989) and became the basis for SSR-based DNA fingerprinting. Products of different length can be amplified with primers flanking the variable microsatellite region and single locus are typically amplified, resulting in one or two

bands, depending on the homozygous or heterozygous state in diploid organisms. Therefore, SSRs are considered locus-specific and codominant markers.

The popularity of microsatellites stems from a unique combination of several important advantages: the relatively abundance with uniform genome coverage, the enormous extent of allelic diversity, the hypervariability, the codominant inheritance, the ease of detection by PCR using pair of flanking primers, and requirement for only a small amount of template DNA (Selvi *et al*., 2003; Chen *et al*., 2009; Singh *et al*., 2010; Singh *et al*., 2011).

Numerous molecular marker strategies have been developed, but the most common employs sequence information of repeat-flanking regions to design locus-specific PCR primer pairs. The necessity of sequence information for primer design is the more serious obstacle of this technique, in addition to the possible presence of undetected null alleles, which can interfere with the interpretation of molecular data. Null alleles can be due to mutations in one or both primer binding sites (Weber and May, 1989) and these mutations can prevent PCR amplification. Homozygous individuals for a null allele do not show any band at all, whereas heterozygotes have only one band and therefore mimic a homozygote on the electrophoresis gel. Erroneous interpretations due to null alleles may be solved by redesigning primer pairs for the locus, avoiding the mutated primer binding site and by examining multiple microsatellite loci, reducing the influence of null alleles.

SSRs are the most frequent DNA marker used for fingerprinting and have been developed for use in sugarcane. Screening and evaluating the available genetic

diversity in sugarcane in SSR markers could both optimize and facilitate the breeding process (Cordeiro, 2001). In sugarcane, various studies were carried out for characterizing genotypes in different countries with SSR markers. Some of these studies are: diversity among mainland USA sugarcane genotypes (Glynn *et al*., 2009), analysis of genetic relatedness among sugarcane genotypes (*Saccharum* spp. hybrids) from breeding programs in China and other countries (Chen *et al*., 2009), assessment of genetic diversity among sugarcane genotypes from India (Pandey *et al*., 2011), and diversity analysis of sugarcane genotypes from Pakistan (Smiullah *et al*., 2013).

2.10.3.2.4. Inter Simple Sequence Repeat (ISSR)

Inter Simple Sequence Repeat (ISSR) is a PCR based technique, which involves amplification of DNA segment present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (Joshi *et al.*, 2000; Reddy *et al*., 2002). ISSR markers were introduced in 1994 (Gupta *et al*., 1994; Zietkiewicz *et al*., 1994) for assessing genetic variation below the species level, mainly in studying population structure and differentiation of cultivated plants. The technique uses microsatellites, usually 16-25 bp long, as primers in a single primer PCR reaction targeting multiple genomic loci to amplify mainly the inter-SSR sequences of different sizes. The microsatellite repeats used as primers can be di-nucleotide, tri-nucleotide, tetra-nucleotide, or penta-nucleotide (Fang and Roose, 1997; Wolfe *et al.*, 1998; Reddy *et al*., 2002).

The extent of polymorphism varies with the nature (unanchored, 3'-anchored, or 5'-anchored) and sequence of the repeat motifs (di-, tri-, tetra-, or penta-repeats) in the primer employed. When unanchored i.e. only the SSRs are used as primers, the primer tends to slip within the repeat units during amplification leading to smears instead of clear bands, because unanchored primer can anneal anywhere in the repeat region on the template DNA, while extending the primer (anchoring) with 1-4 degenerate nucleotides at the 3'end or 5'end assures annealing only to the ends of a microsatellite in template DNA, thus obviating internal priming and smear formation (Reddy *et al*., 2002) since anchored primer at the 3'end or 5'end anneals at specific regions on the template DNA and produces clear bands. The primers anchored at 3' end give clear banding pattern as compared to those anchored at 5' end (Blair *et al*., 1999) since primers at the 5' end amplifies part of the repeat region and also leading to larger bands.

Usually di-nucleotide repeats, anchored either at 3' or 5' end reveal high polymorphism (Blair *et al*., 1999; Joshi *et al*., 2000). In rice and wheat, di-nucleotide simple sequence repeats used as primers gave the maximum number of bands and are therefore more common than any SSRs with larger units (Blair *et al*., 1999). Poly (GA) based 3'anchored primers produced 5 times as many bands as those with poly (GT) motif indicating low frequency or lack of clustering of (GT) motif (Blair *et al*., 1999). Poly (AT) repeats are the most abundant di-nucleotides in plants but the primers based on (AT) motif would self-anneal due to sequence complementarity and not amplify the target regions (Reddy *et al*., 2002). The (AG) and (GA) based primers have been shown to amplify clear bands in rice (Blair *et al*., 1999; Joshi *et al*., 2000;

Reddy *et al*., 2002). Similarly, (AG), (CA), (GA), (TC), and (TG) based primers have been shown to amplify clear bands and detect high polymorphism in sugarcane (Da Costa *et al*., 2011). Using ISSRs it has been shown that tetra-nucleotide repeats were abundant across eukaryotic genomes and that tetramers of tetra-nucleotides AGAC and GACA are scattered within the genome of grasses (Gupta *et al*., 1994).

ISSR markers are inherited in a dominant or codominant Mendelian fashion (Gupta *et al.*, 1994). However, they are interpreted as dominant markers and scored di-allelically with presence (1) or absence (0) of bands similar to RAPD markers (Wolfe *et al.*, 1998). The absence of a band is interpreted as sequence divergence at the site of primer binding. This occurs by loss of a locus through mutation i.e. the deletion of the site or chromosomal rearrangement which could prevent amplification of a fragment and thus give a presence or absence polymorphism or the amplified region would result in the absence of a product (Wolfe and Liston, 1998).

ISSRs have high reproducibility possibly due to the use of longer primers (16-25 bp) as compared to RAPD primers (10 bp) which permits the subsequent use of high annealing temperature (45-60 ^{0}C) leading to higher stringency. Various studies on reproducibility of ISSR markers were carried out. Among these, Fang and Roose (1997) reported a reproducibility level of more than 99 % after performing repeatability tests for ISSR markers by using DNA samples of the same cultivar grown in different locations, DNA extracted from different aged leaves of the same individual, and by performing separate PCR runs. In other cases, the reproducibility of ISSRs amplification products ranged from 86 to 94 % (Moreno *et al*., 1998).

ISSRs exhibit the specificity of microsatellite markers, but do not need sequence information for primer designing enjoying the advantage of random markers (Joshi *et al*., 2000). Furthermore, the potential advantages of this technique are: (i) simple and fast, (ii) low quantity of template DNA required, (iii) randomly distributed throughout the genome, (iv) small reaction volumes and amounts of enzyme are needed for PCR, (v) the hypervariability of banding patterns, (vi) large quantities of plant material for DNA extraction are not required, (vii) no specialized apparatus needed (e.g. automated sequencer, autorad development, etc.), (viii) banding patterns are easily scorable, and (ix) the higher annealing temperatures used for ISSR reactions may reduce the amount of template-primer mismatch artifacts than may be encountered with RAPD markers, which generally rely on lower annealing temperatures. However, Limitations of the technique are: (i) pure template DNA and similar concentrations among samples are required for standardization of reactions, (ii) optimization of initial reactions is needed, (iii) bands are scored as dominant markers and genetic diversity estimates are based on diallelic characters (Wolfe *et al.,* 1998).

ISSR markers have been used for genetic characterization of various plant species. Some of these are: *Hagenia abyssinica* (Tiliye Feyissa *et al*., 2007), coffee (*Coffea arabica* L.) (Esayas Aga *et al.*, 2005; Kassahun Tesfaye *et al*., 2014), rice (*Oryza sativa)* (Gezahegn Girma *et al.*, 2010), anchote (*Coccinia abyssinica* (Lam.) Cogn.) (Abreham Bekele *et al*., 2014), and lentil (*Lens culinaris*) (Edossa Fikiru *et al*., 2007). It is also used for different applications such as genetic diversity analysis, fingerprinting and genome mapping, parentage, phylogenetic studies, gene tagging,

clone and strain identification, and taxonomic studies of closely related species (Zietkiewicz *et al*., 1994). ISSRs are widely used in genetic diversity analyses because of no genome sequence information is required for primer designing, and the primers are not proprietary and can be synthesized by anyone, their development costs are low, highly discriminative, reproducible, informative, reliable, and the laboratory procedures can easily be applied to any plant species (Zietkiewicz *et al*., 1994; Esayas Aga *et al.*, 2005). Therefore, for this study, ISSR markers were selected because the technique is simple, fast, cost effective, and easily accessible or available on hand.

3. OBJECTIVE OF THE STUDY

3.1. General objective

- The general objective of this study was to evaluate the level and patterns of genetic diversity and establish relationships among introduced genotypes and populations of sugarcane (*Saccharum officinarum* L.) in Ethiopia using Inter Simple Sequence Repeat (ISSR) markers.

3.2. Specific objectives

- To assess the level of genetic diversity in introduced sugarcane genotypes and populations;
- To reveal genetic structure within and among introduced genotypes and populations of sugarcane;
- To determine genetic identity among introduced genotypes and populations of sugarcane;
- To identify divergent genotypes and populations that could be used as potential parents' in the future sugarcane breeding programs in Ethiopia.

4. MATERIALS AND METHODS

4.1. Plant material and sampling strategy

Sugarcane plant used in this study was collected from Wonji Sugarcane Plantation and Factory. This site is located about 107 km southeast of Addis Ababa at 8°24'25" Latitude and 39°16'18" Longitude and at an altitude of 1540 m.a.s.l. It receives an annual average rainfall of 831 mm and maximum and minimum average temperatures of 27 ^{0}C and 15 ^{0}C, respectively. The most common soil types in Wonji Sugarcane Plantation are light (coarse textured) and heavy soil types (black cotton) which are more common in Wonji Sugarcane Plantation.

A total of 82 genotypes of sugarcane were used for this study. These genotypes were previously introduced to Ethiopia from various countries abroad for the purpose of commercial sugarcane production. The majority of them were newly released hybrid varieties from sugarcane breeding institutes in different countries. Sampling was done randomly with approximately 10 m distance from each other for assuring varietal purity. Uppermost young leaf samples were taken from three months old randomly selected three sugarcane plants from each population that were commercially grown in the field. The leaf samples were dried with silica gel put separately in plastic bag for genomic DNA extraction. The seven populations were selected based on their population size and origins of the genotype that was introduced to Ethiopia. The genotypes were introduced from Barbados, Cuba, France, India, South Africa, Sudan, and USA. The genotypes were grouped into four major continental regions viz Africa (South Africa and Sudan), North America (Barbados, Cuba, and USA), Asia (India),

and Europe (France). Lists of individual sugarcane genotypes used in this study are shown in appendix 1.

Table 1. Number of genotypes of sugarcane used for molecular diversity analysis

No.	Populations	Letter code	No. of genotypes
1	Barbados	B	13
2	Cuba	C	11
3	France	F	22
4	India	I	9
5	South Africa	SA	8
6	Sudan	S	10
7	USA	U	9
	Total		**82**

4.2. Genomic DNA extraction

According to Gilbert *et al*. (1999), bulk sampling approach was chosen as it permits representation of a sample by optimum number of plants to yield optimal amounts of DNA. Therefore, in this study, approximately equal number of three silica-gel-dried leaf tissue per sample were bulked and ground with Mix and Mill (MM 400) grinding machine.

Genomic DNA was extracted in 2 ml Eppendorf tubes using about 0.3 to 0.6 g powder leave samples following Borsch *et al*. (2003) with slight modification (i.e. second extraction). Then, the powder was dissolved with 700 µL Cethyl Trimethyl

Ammonium Bromide (CTAB) extraction buffer (2 % CTAB, 5 M NaCl, 1 % PVP, 0.5 M EDTA at pH 8.0, 1 M Tris-HCl at pH 8.0, and 98 % Mercapto-ethanol) pre-warmed in water bath at 65 ^{0}C. The mixture was incubated for up to 30 minutes by thoroughly mixing at about 10 minute's interval in electrical constant temperature water bath. Then, centrifuged at 13,000 rpm for 7 minutes and the upper layer (supernatant) was discarded to ensure the complete separation of the genomic DNA from secondary metabolites (first extract) and the same are repeated to attain quality DNA in the second extract. The supernatant of crude DNA extraction (second extract) was transferred to new 1.5 ml Eppendorf tubes using cut blue pipette tips then, 600 µL chloroform was added immediately to the tube with supernatant and shaken thoroughly by turning inversing the Eppendorf caps for approximately 5 minutes and centrifuged at 13,000 rpm for 7 minutes; the supernatant transferred (only clear liquid) to new Eppendorf tubes with cut blue pipette tips and this step was repeated for better clean up of the genomic DNA and to remove all impurities. Then, cooled Isopropanol (4 ^{0}C), approximately 2/3 of the solution volume was added to the solution and shaken carefully and allowed to freeze for more than 2 h at -20 ^{0}C. After centrifuge for 15 minutes at 13,000 rpm, the liquid was aspirated using yellow tips without touching the pellet and 200 µL of 70 % ethanol was added to the pellet and centrifuged for 15 minutes at 13,000 rpm in a cooled centrifuge and aspirated ethanol by yellow tips and dried the DNA-pellet at room temperature for approximately 15 minutes and dissolved in 100 µL TE (1x, p.a. grade) and stored at 4 ^{0}C. Then, 50 µL cooled 7.5 M NH_4Ac-solution was added and mixed carefully. Cool 100 % ethanol (300 µL) was also added and mixed carefully and frozen for more than 2 h at -20 ^{0}C,

centrifuged for 35 minutes at 13,000 rpm and aspirated the fluid carefully. Then, 200 µL of 70 % ethanol was added and the inner cap surface by rinsed turning the cap and centrifuged for 15 minutes at 13,000 rpm and the liquid aspirated and pellet dried at room temperature. Then, pellet was dissolved in 100 µL TE (1x, p.a. grade) and these steps were repeated with 50 µL, 3 M NaAc-solution instead of 7.5 M NH_4Ac-solution (4 ^{0}C, half the volume) then, centrifuged for 15 minutes at 13,000 rpm and aspirated the liquid and dried the pellet at room temperature. Finally, the DNA-pellet was re-suspended in 100 µL TE (1x, p.a. grade) buffer (1 M Tris-HCl and 0.5 M EDTA at pH 8.0) and stored at -20 ^{0}C until the start of PCR work with the DNA stored at 4 ^{0}C.

4.3. Concentration and quality checking of DNA

DNA concentration and quality was checked using the nanodrop measurement (NanoDrop^{TM}2000/2000c) with a range of 53.7-1960.7 ng/µl concentration and 1.7-2.1 ratio purity (appendix 2). The genomic DNA was also tested using 1 % agarose gel by applying 2 µL genomic DNA loaded after mixing with 2 µL 6X loading dye to check the concentration and quality of DNA on gel matrix using gel documentation system (Biosens SC750) under UV light. Genomic DNAs were then diluted for ISSR-PCR to an approximate concentration of 20 ng/µl.

4.4. Screening and optimization of ISSR primers

A total set of 15 ISSR primers were acquired from Genetic Research Laboratory (originally obtained from University of British Colombia, Primer kit UBC 900) of the Microbial, Cellular and Molecular Biology Department, College of Natural Sciences, Addis Ababa University. These primers were selected based on published

experimental results in sugarcane (Da Costa *et al*., 2011) and related *Saccharum* species. For efficient molecular characterization of sugarcane genotypes, initially two individuals were randomly selected from three representative populations to screen the primers for their polymorphism and reproducibility at 45 ^{0}C, 48 ^{0}C, 53 ^{0}C, and 55 ^{0}C annealing temperature. Band intensity and reproducibility of all conditions were compared and optimized.

Then, based on amplified PCR reactions of six representative sugarcane samples, eight di-, one tri-, one tetra-, and two penta-nucleotide primers were finally selected for further molecular analysis based on their reproducibility and polymorphism. All di-nucleotide repeats have anchored 3'sequence and the remaining tri-, tetra-, and penta-nucleotide repeats have unanchored 3' sequence. The total primers screened to amplify 82 genotypes of sugarcane are listed in Table 2, while the 12 selected primers after screening are listed in Table 3.

Table 2. Total ISSR primers used for initial screening and PCR optimization

No.	Primers code	Sequence (5' - 3')	Repeat motifs	Amplified products	Reproducibility	Remark
1	810	(GA)8 T	di-	present (good)	reproducible	Selected
2	811	(GA)8 C	di-	present (good)	reproducible	Selected
3	812	(GA)8 A	di-	present (good)	reproducible	Selected
4	814	(CT)8 A	di-	absent	×	Rejected
5	815	(CT)8 G	di-	present (good)	reproducible	Selected
6	822	(TC)8 A	di-	present (good)	reproducible	Selected
7	824	(TC)8 G	di-	present (good)	reproducible	Selected
8	834	(AG)8 YT	di-	present (good)	reproducible	Selected
9	848	(CA)8 RG	di-	present (good)	reproducible	Selected
10	866	(CTC)6	tri-	present (good)	reproducible	Selected
11	873	(GACA)4	tetra-	present (good)	reproducible	Selected
12	874	(CCCT)4	tetra-	absent	×	Rejected
13	878	(GGAT)4	tetra-	absent	×	Rejected
14	880	(GGAGA)3	penta-	present (good)	reproducible	Selected
15	881	(GGGTG)3	penta-	present (good)	reproducible	Selected
		Total = 15			**12 Selected**	

× = *reproducibility test was not performed because of no initial amplification at all.*

Table 3. Selected ISSR primers used for total amplification of all DNA samples

Primers code	Sequence (5' - 3')	Repeat motifs	Tm	GC %	Ta
810	(GA)8 T	di-	50	47.0	45
811	(GA)8 C	di-	52	52.9	48
812	(GA)8 A	di-	50	47.0	45
815	(CT)8 G	di-	52	52.9	48
822	(TC)8 A	di-	54	50.0	48
824	(TC)8 G	di-	52	52.9	48
834	(AG)8 YT	di-	52	44.4	45
848	(CA)8 RG	di-	54	50.0	48
866	(CTC)6	tri-	60	66.6	55
873	(GACA)4	tetra-	48	50.0	45
880	(GGAGA)3	penta-	48	60.0	45
881	(GGGTG)3	penta-	54	80.0	53

Tm = melting temperature, Ta = annealing temperature, Y = (C, T), R = (A, G)

4.5. PCR amplification and gel electrophoresis

The polymerase chain reaction (PCR) was conducted in Biometra 2003 T3 Thermo cycler machine. PCR amplification was carried out in a 26 µL final reaction volume containing 2 µL diluted template DNA, 16.2 µL ddH_2O (ultra pure), 1 µL dNTPs (20 mM each), 2.5 µL Taq reaction buffer (10X buffer), 3.5 µL $MgCl_2$ (25 mM), 0.4 µL primer (20 pmol/µl), and 0.4 µL Taq DNA polymerase (5 u/µl). 12 ISSR primers

were selected having good band intensity and reproducible results with PCR condition operated at 94 ^{0}C initial denaturation for 4 minutes followed by 40 cycles of 94 ^{0}C denaturation for 15 seconds, 45 ^{0}C/ 48 ^{0}C/ 53 ^{0}C/ 55 ^{0}C for 1 minute primer annealing (depending on primer GC content and Tm), 72 ^{0}C for 1 minute and 30 seconds extension and final extension at 72 ^{0}C for 7 minutes for Taq DNA polymerase at 99 ^{0}C lid temperature. A negative control, in which the template DNA was replaced by water, was also included in each round of reactions to check for absence/presence of contamination. PCR products were stored at 4 ^{0}C until loading on gel for electrophoresis.

For gel electrophoresis of PCR products, 0.86 g (1.72 %) agarose was dissolved in 50 ml, 1X TBE buffer and melted by microwave oven for about 2 minutes and 30 seconds by gentle shacking with interval of 30 seconds. Then, 2 μL of ethidium bromide (10 mg/ml) was added to melted agarose after it was cooled. The agarose was casted for 30 minutes and 12 μL of PCR products were mixed with 2 μL of 6X loading dye on parafilm per sample and loaded to respective wells. Fifteen wells comb was used for each gel slab. The first lane was loaded with 100 bp DNA ladder by loading 2 μL (peq gold range mix) with loading dye in that well as a size standard and the last lane was loaded by a control. The electrophoreses were done at constant voltage of 100 V for about three hours depending on molecular size of amplified product during optimization and stained in staining solution containing 50 μL of ethidium bromide in 450 ml distilled water for 30 minutes. It was then destained in equal amount of distilled water for 30 minutes and after destained the banding patterns were visualized under UV light and photographed using canon camera in the

gel documentation system (Biosens SC750) and documented for band scoring and analysis.

4.6. Band scoring and data analysis

The ISSR band profiles were considered as dominant markers and each locus was counted as a bi-allelic locus with one amplifiable and one null allele. ISSR bands were counted manually after the capturing of gel image with gel documentation system and scored as binary data; '1' for band present, '0' for band absent, and '?' for missing data. Only unambiguously amplified ISSR bands were scored, while faint bands were excluded from scoring. Based on recorded bands, different molecular software was used for analysis of binary data matrix. Percent of polymorphic loci (PPL), Nei's (1973) gene diversity (h), Shannon's information index (I), Gene differentiation between populations were estimated from the coefficient of gene differentiation (G_{ST}), and gene flow (Nm).

POPGENE version1.32 software (Yeh *et al.*, 1999) was used to calculate the following genetic diversity parameters: percent of polymorphic loci (PPL), gene diversity (h), and Shannon's information index (I).

Jaccard's similarity coefficient (J) was calculated from NTSYS- pc version 2.02 (Rohlf, 2000) and Free Tree 0.9.1.50 (Pavlicek *et al.*, 1999) and in turn used to calculate similarity between pairs of populations with the formula:

$$S_{ij} = \frac{a}{a + b + c}$$

Where, ‘a’ is the total number of bands shared between individual i and j,

‘b’ is the total number of bands presents in individual i but not in individual j, and

‘c’ is the total number of bands present in individual j but not in individual i.

Analysis of molecular variance (AMOVA) was used to estimate genetic variance components for the ISSR phenotypes using Areliquin version 3.01 (Excoffier *et al*., 2006) computer software to describe the genetic structure at population level.

To examine the genetic relationship among individuals, pairwise similarity indices were estimated using the Jaccard coefficient of similarity (Jaccard, 1908) from NTSYS- pc version 2.02 (Rohlf, 2000) that is used for Unweighed Paired Group Method with Arithmetic Mean (UPGMA) tree construction. Likewise, Free Tree 0.9.1.50 (Pavlicek *et al.,* 1999) was used to construct Neighbor Joining (NJ) tree (Saitou and Nei, 1987). Unlike that of UPGMA which assumes equal rates of evolution (molecular clock assumption), NJ assumes variations in the rate of change and it is often preferred than UPGMA. In addition to NJ and UPGMA tree, two dimensional (2D) and three dimensional (3D) Principal Coordinate Analysis (PCoA) were used to reveal patterns of variation among individual samples based on Jaccard’s similarity coefficient using PAST software version 1.18 (Hammer *et al.,* 2001). The first three axes were later used to plot the three dimensional (3D) PCoA with STATISTICA version 6.0 software (Hammer *et al.,* 2001).

5. RESULTS

5.1. ISSR primer variation and level of polymorphism

In this study, out of the fifteen (15) ISSR primers tested initially, twelve (12) primers (eight di-, one tri-, one tetra-, and two penta-nucleotide repeats) gave clear, reproducible, and polymorphic bands and also ISSR banding profiles of sugarcane genotypes using primer 810, 811, 866, 873, 880, and 881 was shown Figure 3, 4, and 5. The other three ISSR primers failed to give amplification products, or gave poor or non-reproducible bands which were then excluded from the analysis. Thus, the only 12 ISSR primers were selected and used for further molecular analysis of 82 introduced sugarcane genotypes in Ethiopia. A total of 149 clear and scorable bands were amplified, with an average of 12.41 bands per primer. From the total amplified loci, 124 were polymorphic among the sugarcane genotypes, with an average of 10.33 ISSR polymorphic loci per primer. All primers amplified fragments, with a number of amplicons varying from 10 (primers 811 and 812) to 18 (primer 880) fragments, with molecular weight ranging from 200 to 4000 bp. The number of polymorphic loci ranged from 7 (primer 812) to 16 (primer 880) attributing to 83.22 % total polymorphism at species level (Table 4).

There was no single or unique band that was specific to any individual genotype. Percent of polymorphic loci (PPL) was varying from 70 (primer 812) to 90.91 % (primers 810 and 824). Primer 812 resulted in the least polymorphic loci (7) with 70 % of polymorphism overall individual comparison, while primer 810 and 824

generated the highest percent of polymorphic loci (90.91 %) with 10 polymorphic loci out of 11.

In addition to highest PPL diversity parameter revealed by 810 and 824 (90.91 %), the highest Nei's gene diversity (0.399) and Shannon's information index (0.563) were exhibited by primer 881 followed by primer 810 (0.388 and 0.559, Nei's gene diversity and Shannon's information index, respectively). As PPL observation, primer 812 showed least Nei's gene diversity (0.213) and Shannon's information index of 0.329. The overall Nei's gene diversity (h) and Shannon's information index (I) for all primers were 0.310 and 0.458, respectively (Table 4).

Table 4. Lists of all primers used in the analysis, primer sequence, estimated band size range, number of amplified loci (NAL), number of polymorphic loci (NPL), percent of polymorphic loci (PPL), level of gene diversity (h), and Shannon's information index (I).

ISSR primer	Sequence (5'-3')	Band size (bp)	NAL	NPL	PPL	h±SD	I±SD
810	(GA)8 T	750-4000	11	10	90.91	0.388±0.148	0.559±0.202
811	(GA)8 C	1000-4000	10	9	90.00	0.291±0.165	0.445±0.221
812	(GA)8 A	300-1000	10	7	70.00	0.213±0.191	0.329±0.270
815	(CT)8 G	300-2000	14	12	85.71	0.336±0.177	0.492±0.243
822	(TC)8 A	200-1000	13	11	84.62	0.339±0.176	0.495±0.243
824	(TC)8 G	1500-4000	11	10	90.91	0.314±0.131	0.478±0.184
834	(AG)8 YT	1000-4000	13	10	76.92	0.269±0.207	0.401±0.282
848	(CA)8 RG	750-3500	11	9	81.82	0.282±0.201	0.421±0.271
866	(CTC)6	400-1000	11	9	81.82	0.248±0.158	0.387±0.226
873	(GACA)4	300-1000	14	10	71.43	0.261±0.214	0.384±0.297
880	(GGAGA)3	750-4000	18	16	88.89	0.341±0.174	0.499±0.234
881	(GGGTG)3	200-1500	13	11	84.62	0.399±0.178	0.563±0.251
Mean		**620-2583**	**12.41**	**10.33**	**83.14**	**0.306±0.176**	**0.454±0.243**
Overall		**200-4000**	**149**	**124**	**83.22**	**0.310**	**0.458**

SD = Standard Deviation

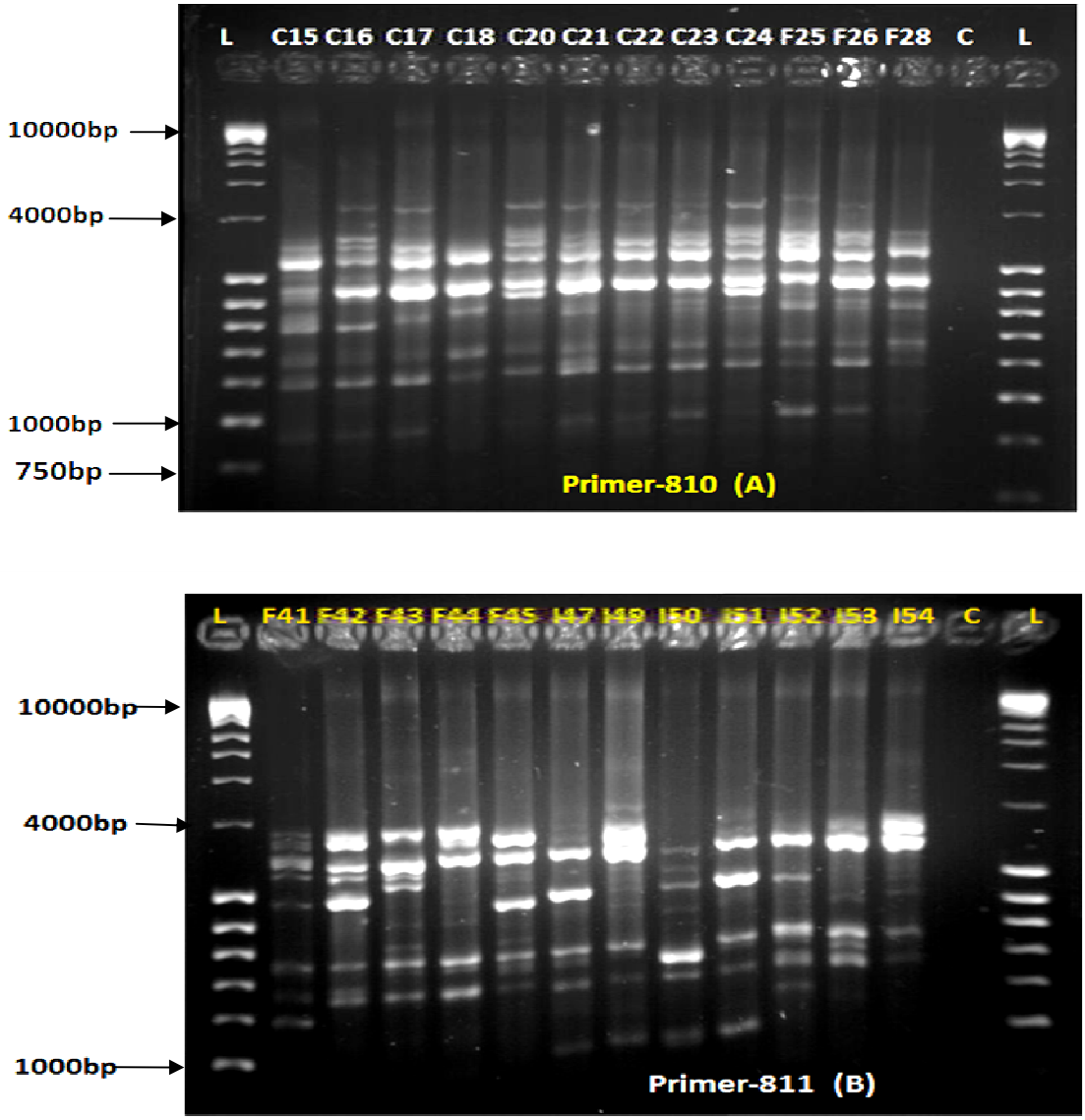


Figure 3. ISSR profiles of sugarcane genotypes generated by di-nucleotide primers (810 (A)) and (811 (B)) using agarose gels. L = DNA Ladder (100 bp); C = Control; C, F, and I = refer to populations of Cuba, France, and India, respectively, and the number system refers to individual code in each population.

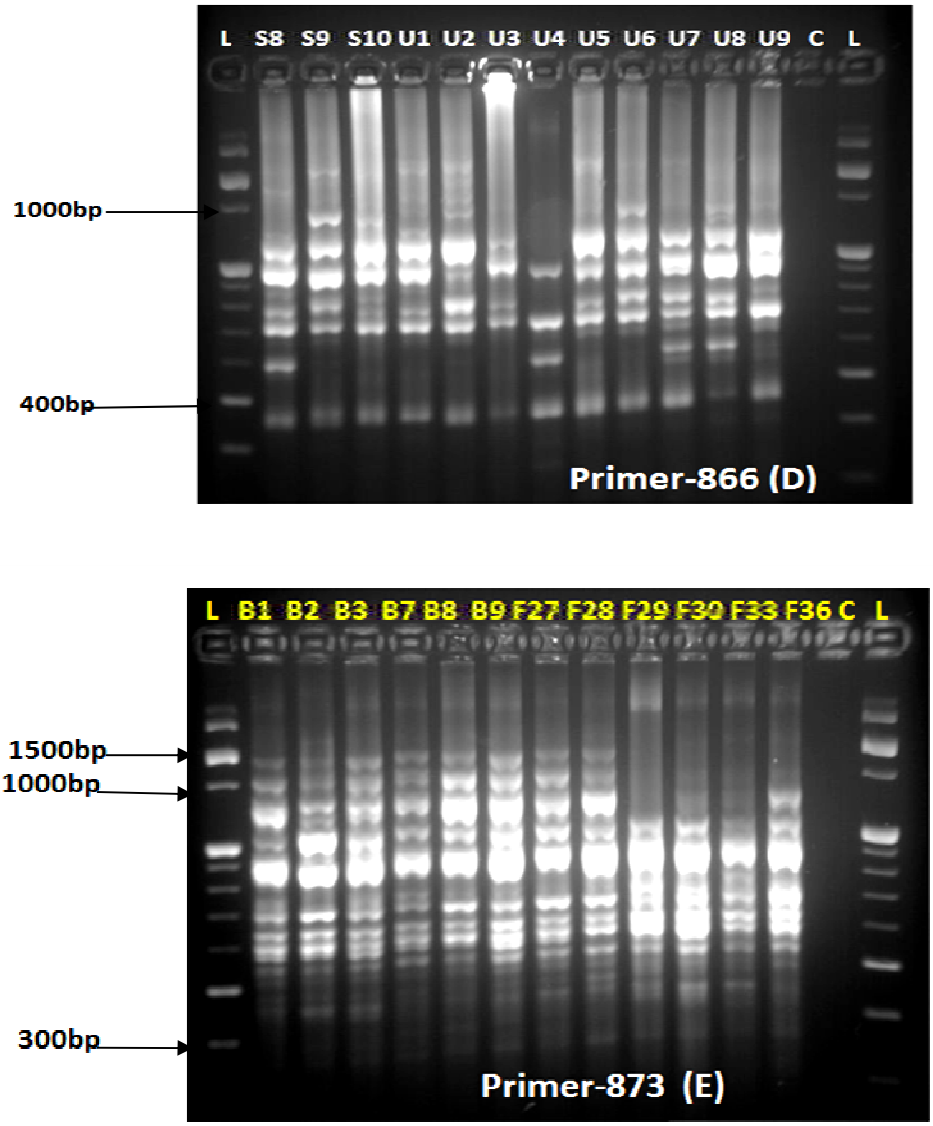


Figure 4. ISSR profiles of sugarcane genotypes generated by tri-nucleotide primer [866 (D)] and tetra-nucleotide primer [873 (E)] using agarose gels. L = DNA Ladder (100 bp); C = Control; S, U, B, and F = refer to populations of Sudan, USA, Barbados, and France, respectively, and the number system refers to individual code in each population.

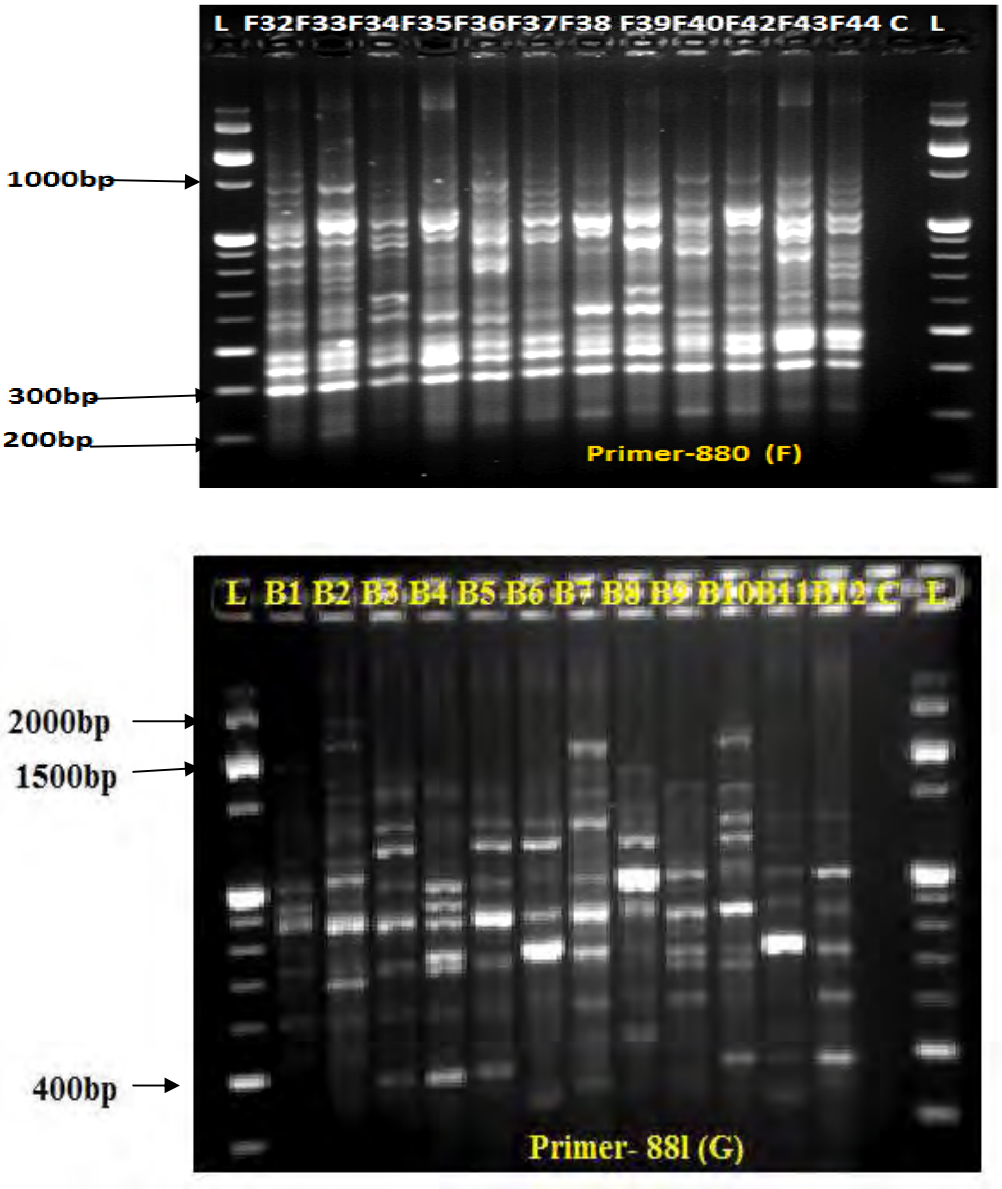


Figure 5. ISSR profiles of sugarcane genotypes generated by penta-nucleotide primers [880 (F] and 881 [G)] using agarose gels. L = DNA Ladder (100 bp); C = Control; B and F = refer to populations of Barbados and France, respectively, and the number system refers to individual code in each population.

5.2. Level of genetic diversity in sugarcane populations

Within populations, the number and percent of polymorphic loci (data generated by all primers) ranged from 43 (28.86 %) for introduced genotypes from Barbados to 71 (47.65 %) for France with a mean of 38.35 % which indicates genotypes from France as having highest genetic diversity, while genotypes from Barbados with lowest genetic diversity. Beside this, high level of polymorphism was obtained when sugarcane populations were analyzed by groups based on geographical origins of the population. Accordingly, North American populations showed highest percent polymorphism (71.14 %) followed by African populations (57.72 %) and European population (47.65 %), while the least level of polymorphism were detected for Asian population only 40.27 % (Table 5).

Based on the analyses of the only eight di-nucleotide primers, the highest polymorphism was observed in France and India populations (48.39 %) followed by genotypes from Cuba (46.24 %). Among population analyses, genotypes from Barbados (34.41 %) showed the least polymorphism, while based on grouping, highest polymorphism was detected for populations from North America (74.10 %) followed by populations from Africa (60.22 %). However, populations from Asia and Europe showed the least polymorphism (48.39 %). The analysis from the single tri-nucleotide primer (866) used for all populations depicted 45.45 % of polymorphism for France genotypes and 9.09 % of polymorphism for Sudan genotypes. Similarly, single tetra-nucleotide primer (873) was used for all populations' analyses and based on this, France genotypes showed the highest percent of polymorphism (64.29 %),

while population of Barbados and Sudan showed the least polymorphism (14.29 %). The overall level of polymorphism for only single tri-nucleotide primer (866) and tetra-nucleotide primer (873) attributed to 81.82 % and 71.43 %, respectively. The penta-nucleotide repeat primers (880 and 881) revealed highest polymorphism in Sudan genotypes with 61.29 %, followed by genotype from Cuba with 58.06 % of polymorphism. In terms of grouping using penta-primers, highest polymorphism was detected in population from North America (83.87 %), while Asian population showed least polymorphism (25.81 %). The two penta-nucleotide primers contributed 87.10 % of polymorphism, while the only eight di-nucleotide primers contributed 83.87 % of polymorphism for all sugarcane populations (Table 5).

The Nei's gene diversity index (h) reflected the diversity among the germplasm collections or groups genotypes. The Shannon's information index (I) was also used to evaluate the genetic diversity within and between the populations or groups genotypes. The higher the index, the higher is the genetic diversity. Based on the overall dataset, Nei's gene diversity (h) ranged from 0.097 for Barbados to 0.171 for France with a mean of 0.137 and Shannon's information index ranged from 0.147 for Barbados to 0.255 for France with a mean of 0.205 (Table 6 and 7).

Nei's gene diversity (h) value generated by only eight di-primers for France and India genotypes were highest with values of 0.173, followed by Cuba genotype with a value of 0.168. The lowest gene diversity was observed in Barbados (0.115). In terms of grouping, the highest gene diversity was observed in genotypes of North America (0.284) followed by population from Africa (0.211), while the least gene diversity

was obtained in population from Asia and Europe (0.173). Based on this dataset, Shannon's information index for France population was highest with a value of 0.260, followed by India and Cuba population with value of 0.259 and 0.249, respectively, but the least Shannon index was observed in population of Barbados (0.174). The overall gene diversity and Shannon's information index detected by eight di-nucleotide primers were 0.306 and 0.455, respectively. Similarly, the analysis based on the two penta-nucleotide primers, the highest gene diversity and Shannon's information index were obtained from genotypes of Sudan with values of 0.251 and 0.364, respectively, followed by Cuba with values of 0.209 and 0.308, respectively, while Barbados population showed the least gene diversity and Shannon index (0.072 and 0.111, respectively). The overall gene diversity and Shannon index analyzed by only two penta-nucleotide primers were 0.365 and 0.526, respectively (Table 6 and 7). The overall percent of polymorphism (PPL), gene diversity (h), Shannon's information index (I), gene flow (Nm), and coefficient of gene differentiation (G_{ST}) for all study populations of sugarcane with 12 ISSR primers of the current study was 83.22 %, 0.310, 0.458, 0.401, and 0.555, respectively (Table 5, 6, and 7).

Table 5. Percent of polymorphic loci (PPL), number of polymorphic loci (NPL) genetic differentiation (G_{ST}), and gene flow (Nm) estimate in sugarcane populations. The analysis was carried out at country level and grouping was made based on geographic origin.

	Percent of polymorphic Loci (PPL)				Over all PPL	NPL	G_{ST}	Nm
Populations	di-primers	tri-primer	tetra-primer	penta-primers	All primers	All primers		
Barbados	34.41	18.18	14.29	22.58	28.86	43	-	-
Cuba	46.24	18.18	21.43	58.06	44.30	66	-	-
France	48.39	45.45	64.29	38.71	47.65	71	-	-
India	48.39	18.18	35.71	25.81	40.27	60	-	-
S. Africa	36.56	36.36	21.43	22.58	32.21	48	-	-
Sudan	39.78	9.09	14.29	61.29	39.60	59	-	-
USA	35.48	18.18	28.57	45.16	35.57	53	-	-
Mean	41.32	23.37	28.57	39.17	38.35	57.14	-	-
Overall	**83.87**	**81.82**	**71.43**	**87.10**	**83.22**	**124**	**0.555**	**0.401**
Groups								
Africa	60.22	36.36	28.57	70.97	57.72	86		
N. America	74.10	45.45	42.86	83.87	71.14	106		
Europe	48.39	45.45	64.29	38.71	47.65	71		
Asia	48.39	18.18	35.71	25.81	40.27	60		
Mean	57.79	36.36	42.85	54.84	54.19	80.75		
Overall	**83.87**	**81.82**	**71.43**	**87.10**	**83.22**	**124**		

Table 6. Nei's gene diversity (h) in sugarcane populations. The analysis was carried out at country level and grouping was made based on geographic origin.

	Nei's gene diversity (h)				Overall (h)
Populations	di-primers	tri-primer	tetra-primer	penta-primers	All primers
Barbados	0.115	0.075	0.053	0.072	0.097
Cuba	0.168	0.087	0.095	0.209	0.164
France	0.173	0.146	0.220	0.145	0.171
India	0.173	0.054	0.144	0.108	0.148
S. Africa	0.125	0.156	0.076	0.092	0.116
Sudan	0.136	0.031	0.044	0.251	0.144
USA	0.119	0.046	0.111	0.168	0.123
Mean	0.144	0.085	0.106	0.149	0.137
Overall	**0.306**	**0.248**	**0.260**	**0.365**	**0.310**
Groups					
Africa	0.211	0.140	0.125	0.285	0.213
N. America	0.284	0.193	0.189	0.291	0.270
Europe	0.173	0.146	0.220	0.145	0.170
Asia	0.173	0.054	0.144	0.108	0.148
Mean	0.210	0.133	0.169	0.207	0.200
Overall	**0.317**	**0.252**	**0.259**	**0.366**	**0.317**

Table 7. Shannon's information index (I) in sugarcane populations. The analysis was carried out at country level and grouping was made based on geographic origin.

	Shannon's information index (I)				Overall (I)
Populations	di-primers	tri-primer	tetra-primer	penta-primers	All primers
Barbados	0.174	0.109	0.079	0.111	0.147
Cuba	0.249	0.122	0.136	0.308	0.241
France	0.260	0.223	0.334	0.217	0.255
India	0.259	0.086	0.212	0.157	0.221
S. Africa	0.189	0.225	0.116	0.135	0.174
Sudan	0.205	0.048	0.069	0.364	0.214
USA	0.181	0.076	0.164	0.249	0.185
Mean	0.216	0.127	0.158	0.220	0.205
Overall	**0.455**	**0.387**	**0.384**	**0.526**	**0.458**
Groups					
Africa	0.314	0.208	0.180	0.415	0.314
N. America	0.416	0.279	0.270	0.438	0.397
Europe	0.260	0.223	0.334	0.217	0.255
Asia	0.259	0.086	0.212	0.157	0.221
Mean	0.312	0.199	0.249	0.306	0.296
Overall	**0.468**	**0.393**	**0.383**	**0.529**	**0.467**

5.3. Analysis of Molecular Variance (AMOVA)

Levels of genetic variation can be attributed to within and among population components. Analysis of molecular variance (AMOVA) was used to partition the existing genetic variation into different components. In the present study, analysis of molecular variance was carried out in two ways; one was done using the populations grouped into their geographical regions and the other was done for the overall loci of the entire groups (i.e., using the seven countries group as it is without grouping) by considering them as one geographic region.

Partitioning of genetic diversity by analysis of molecular variance using grouped populations revealed that out of the total genetic diversity, most of the ISSR diversity was distributed between individual genotypes within the populations (61.66 %), with the remaining diversity being distributed among populations within groups (34.88 %) (i.e., populations within geographical regions) and among groups (3.46 %) (Table 8). Similarly, partitioning of genetic diversity by analysis of molecular variance without grouping populations revealed that out of the total genetic diversity, most of the genetic diversity is due to differences between individual genotypes within the populations/ country based group (63.56 %), while the remaining is due to differences among populations (36.44 %) (Table 9). In both cases, Fixation index (F_{ST}) indicated about 38.3 % for grouping and 36.4 % without grouping among population genetic differentiation with significant components of molecular variations ($P = 0.00$).

Table 8. Partitioning of the genetic variation into among groups, among populations within groups, and within populations by analysis of molecular variance (AMOVA) based on 149 ISSR bands generated from sugarcane.

Source of variation	df	Sum of squares	Variance components	Percentage of variation	Fixation Index (F_{ST})	P-value*
Among groups	3	366.553	0.825 Va	3.46		0.00
Among populations within groups	3	268.471	8.328 Vb	34.88	0.383	0.00
Within populations	75	988.821	14.72 Vc	61.66		0.00
Total	81	1623.846	23.874	100		

*df = degree of freedom, * significance tests (1023 permutations)*
N.B: (Va) represents variance components due to among groups, (Vb) represents variance components due to among populations within groups, and (Vc) represents variance components due to within populations variations.

Table 9. Partitioning of genetic variations into within and among populations by analysis of molecular variance (AMOVA) based on 149 ISSR bands generated from sugarcane.

Source of variation	df	Sum of squares	Variance components	Percentage of variation	Fixation Index (F_{ST})	P-value*
Among populations	6	114.252	1.44530 Va	36.44	0.364	0.00
Within populations	75	189.065	2.52087 Vb	63.56		
Total	81	303.317	3.96617	100		

*df = degree of freedom, * significance tests (1023 permutations)*
N.B: (Va) represents variance components due to among populations and (Vb) represents variance components due to within population variations.

5.4. Genetic similarity among the sugarcane populations

On the basis of Jaccard's similarity coefficient, pairwise comparison of the seven populations of sugarcane, similarity ranged from 0.532 to 0.700 with mean of 0.616. The highest similarity was found between Sudan and South Africa populations (0.700) followed by similarity between USA versus Sudan and USA versus South Africa with 0.653 and 0.651 similarity coefficients, respectively. However, the maximum genetic distance was observed between population from USA and Barbados followed by USA and Cuba with similarity coefficient of 0.532 and 0.543, respectively (Table 10). The average genetic distance among the seven sugarcane country was 38.4 % only but most of the populations showed an average of 61.6 % genetic similarity.

UPGMA based dendrogram for seven sugarcane population generated two distinct major clusters with 0.61 similarity coefficients (Figure 6). The first bottom major cluster contains populations of USA, Sudan, and South Africa, while the second top major cluster contained populations of Barbados, Cuba, France, and India. Population based UPGMA clustering of an overall analysis showed no strong clustering of populations with respect to their geographical origin, indicating there was widespread practice of clone exchange between sugarcane growing regions.

Table 10. Similarity matrix among the sugarcane populations obtained by Jaccard's similarity coefficient

Populations	Barbados	Cuba	France	India	South Africa	Sudan	USA
Barbados	1.000						
Cuba	0.621	1.000					
France	0.589	0.625	1.000				
India	0.614	0.615	0.617	1.000			
South Africa	0.584	0.569	0.609	0.555	1.000		
Sudan	0.562	0.568	0.646	0.601	**0.700**	1.000	
USA	**0.532**	0.543	0.595	0.594	0.651	0.653	1.000

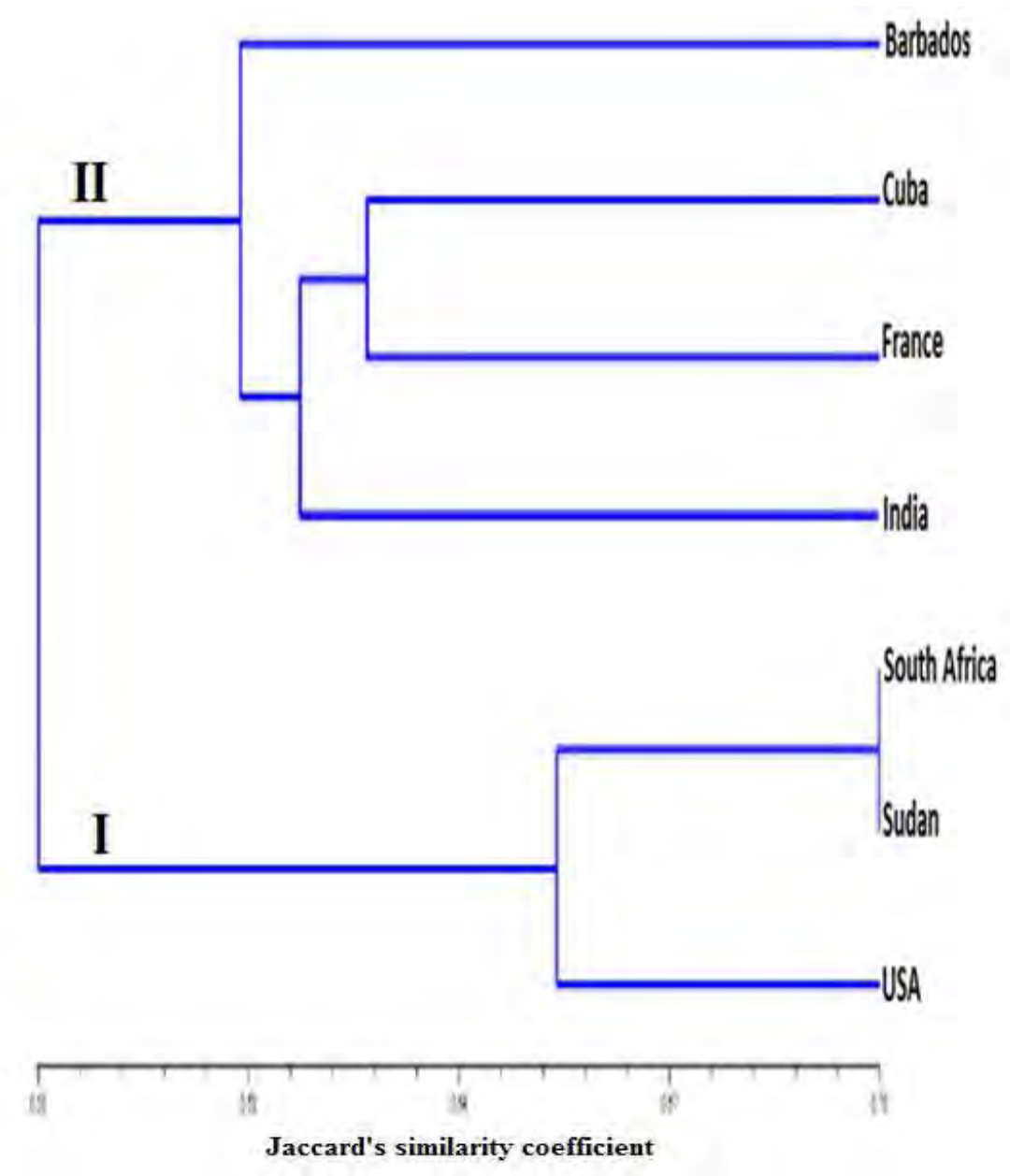


Figure 6. UPGMA based dendrogram for the seven populations of sugarcane.

5.5. Clustering analysis

UPGMA (Unweighted Pair Group Method with Arithmetic Mean) and NJ (Neighbor Joining) tree construction methods was used to construct dendrogram on the basis of Jaccard's similarity coefficient to explore the relationship among genotypes of sugarcane based on 149 ISSR bands amplified by a total of 12 ISSR primers (eight di-, one tri-, one tetra-, and two penta-nucleotide primers). The dendrogram derived from UPGMA analysis of the whole ISSR data with 82 sugarcane genotypes showed two major distinct clusters (major-I and II) and two sub-clusters (sub-I and II) within each major cluster at around 0.60 similarity coefficient and with similarity values ranging from 0.59 to 0.90 with mean of 0.74 for all individuals (Figure 7). The first major cluster (major-I) was composed of two sub-clusters; the first sub-cluster (sub-I) was dominated by all individuals from Barbados (B) and one individual each from Cuba (C) and USA (U), while the second sub-cluster (sub-II) was highly dominated by individuals from Cuba (C), India (I), and some of individuals from France (F) with one individual from South Africa (SA) and they formed their own separate groups within this cluster. The second major cluster (major-II) also branched into two sub-clusters; the third sub-cluster (sub-III) was comprised individuals only from France (F), while the fourth sub-cluster (sub-IV) was dominated by individuals from South Africa (SA), Sudan (S), and USA (U) and clearly formed their own separate groups within the clusters. Individual based UPGMA clustering of an overall analysis showed strong clustering of individuals with respect to their populations except few intermixed individuals from other population.

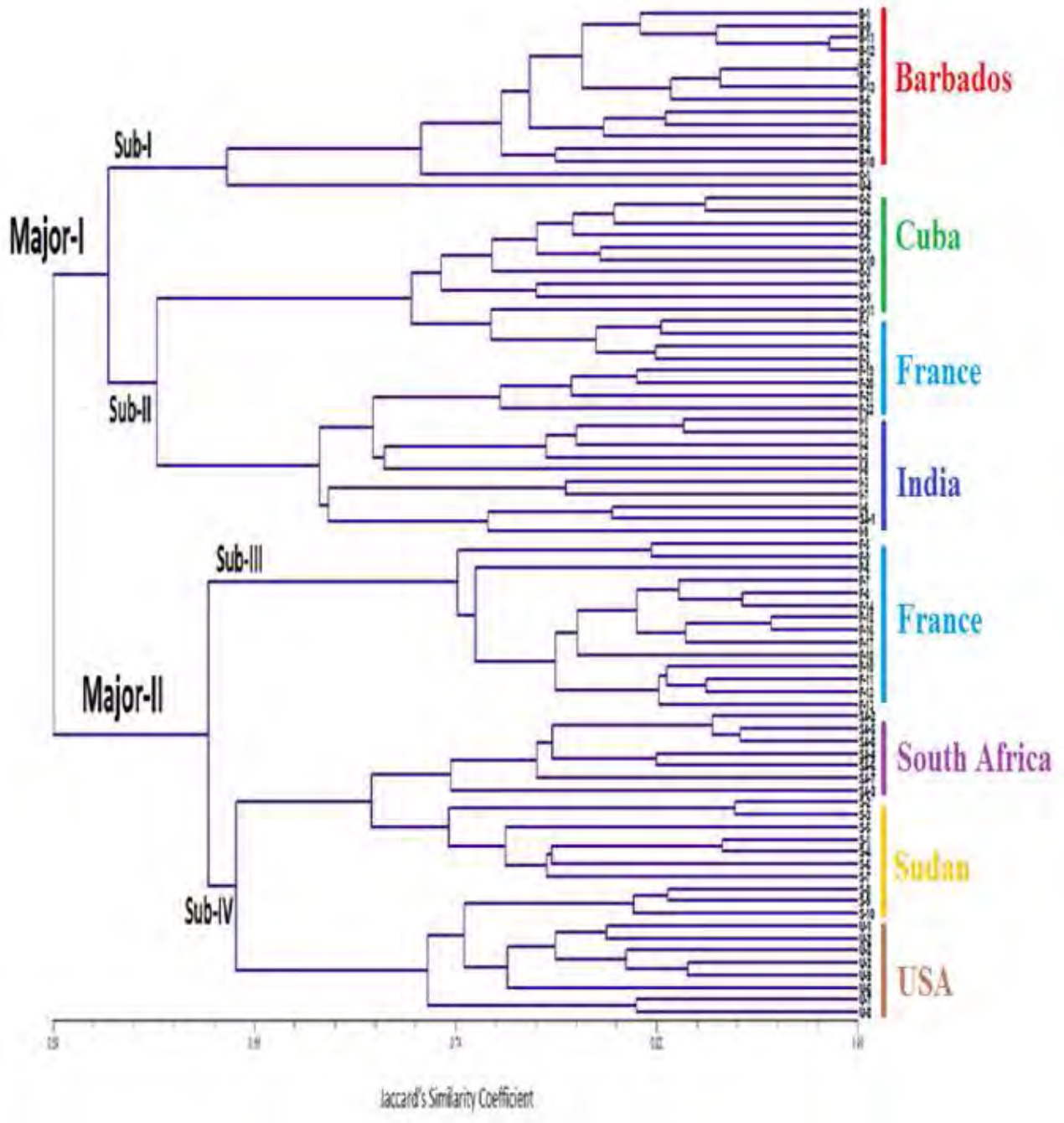


Figure 7. Dendrogram obtained with UPGMA based on Jaccard's similarity coefficients for 82 genotypes of sugarcane using 12 ISSR primers; Key: B = Barbados, C = Cuba, F = France, I = India, SA = South Africa, S = Sudan, U = USA.

Clustering based on Neighbor Joining (NJ) also resulted in similar clustering patterns as that of UPGMA with two major clusters (major cluster-I and II) without considering two outliers of France individuals (F-15 and F-7). There was no any cutoff value but moderate intermixing was observed in NJ clustering than with UPGMA clustering (Figure 8). The first major cluster (major-I) branched into two sub-clusters (sub cluster-I and II). From this cluster, sub-cluster-I (sub-I) was dominated by individuals from Barbados (B), Cuba (C), India (I), and some of individuals from France (F) and they formed their own clear separate groups within the cluster as UPGMA analysis. Sub-cluster-II (sub-II) was dominated by individuals from USA (U), Sudan (S), and South Africa (SA); particularly this sub cluster (sub-II) was comprised some sort of intermixed individuals from one another. However, the second major cluster (major-II) was dominated by only individuals from France (F). In terms of grouping intensity, however, individuals from Barbados (B), Cuba (C), France (F), and India (I) tend to form strong grouping with respect to their population than individuals from USA (U), Sudan (S), and South Africa (SA) in NJ analysis.

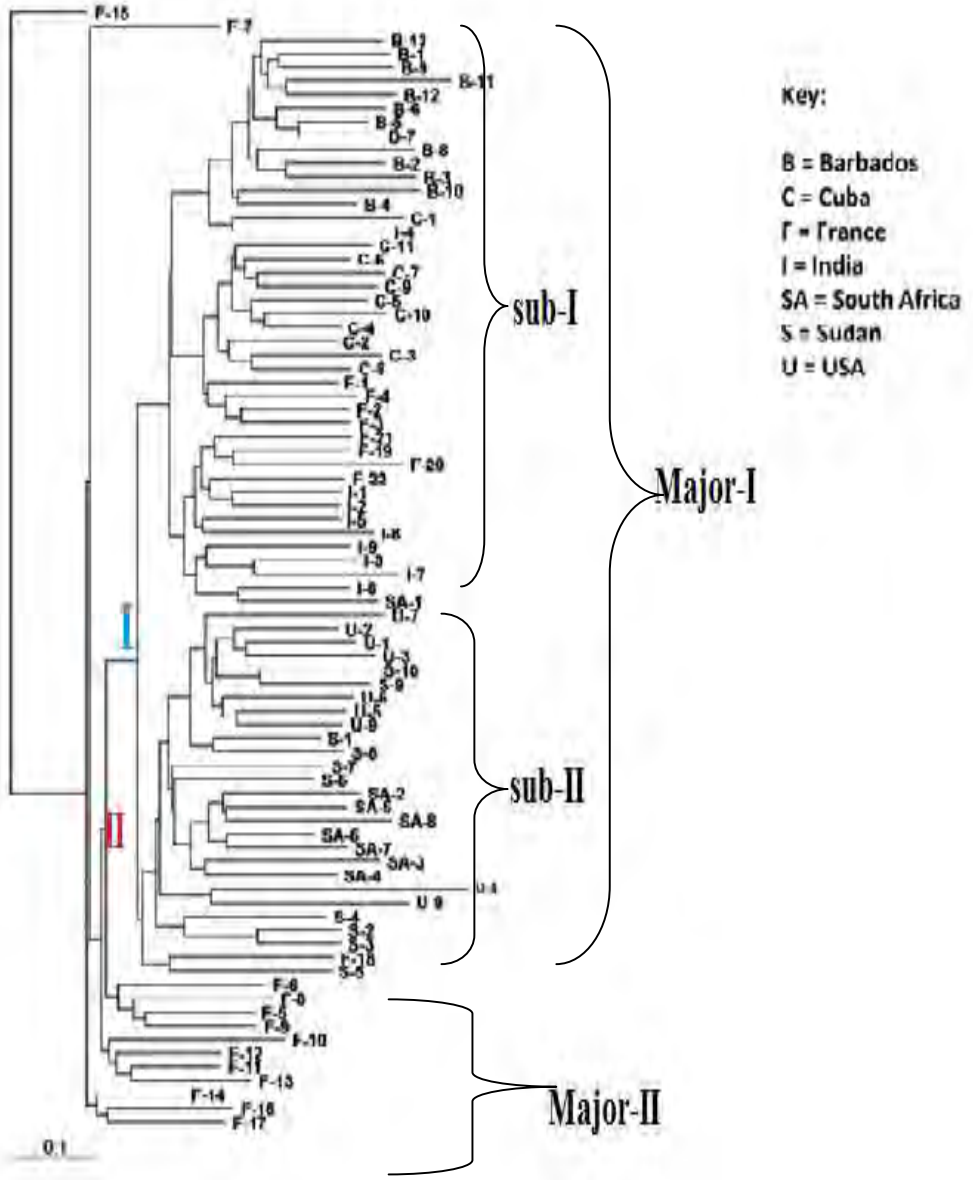


Figure 8. NJ based analysis of 82 genotypes of sugarcane using 12 ISSR primers after pairwise comparisons using Jaccard's similarity coefficient.

5.6. Principal Coordinate Analysis (PCoA)

In addition to UPGMA and NJ analyses, two dimensional (2D) and three dimensional (3D) Principal Coordinate Analysis (PCoA) were used to reveal patterns of variation among sugarcane genotypes on the basis of Jaccard's similarity coefficient using PAST and STATISTICA software packages. Based on the overall ISSR data obtained, Principal Coordinate analysis was performed utilizing the first three coordinates of the PCo having Eigen-values of 4.45, 3.04, and 2.31, which accounted for 7.81 %, 5.35 %, and 4.06 % variation, respectively. It was used to show the grouping of genotypes using both two dimensional and three dimensional PCoA analyses.

The analysis of this study showed that the genotypes of the seven sugarcane populations examined tend to form grouping based on their geographic region. However, with an overall analysis, the two dimensional (2D) representation of PCoA revealed somewhat poor patterns of grouping on the basis of population and the individuals were intermixed from one another except genotypes from Barbados (B). Some of the individuals from France (F) also formed distinct grouping, while the three dimensional (3D) representation showed better grouping patterns than two dimensional (2D) and most of the genotypes were made some sort of clear grouping pattern with respect to their population. Genotypes from Barbados (B) and France (F) showed clear grouping on 2D analysis, here also genotypes from France formed their own grouping on 3D PCoA. Likewise, genotypes of Cuba and USA formed their own grouping relative to others. Similar to 2D PCoA observation, individuals of India (I),

South Africa (SA), and Sudan (S) populations were intermixed with other populations in 3D PCoA (Figure 9-10).

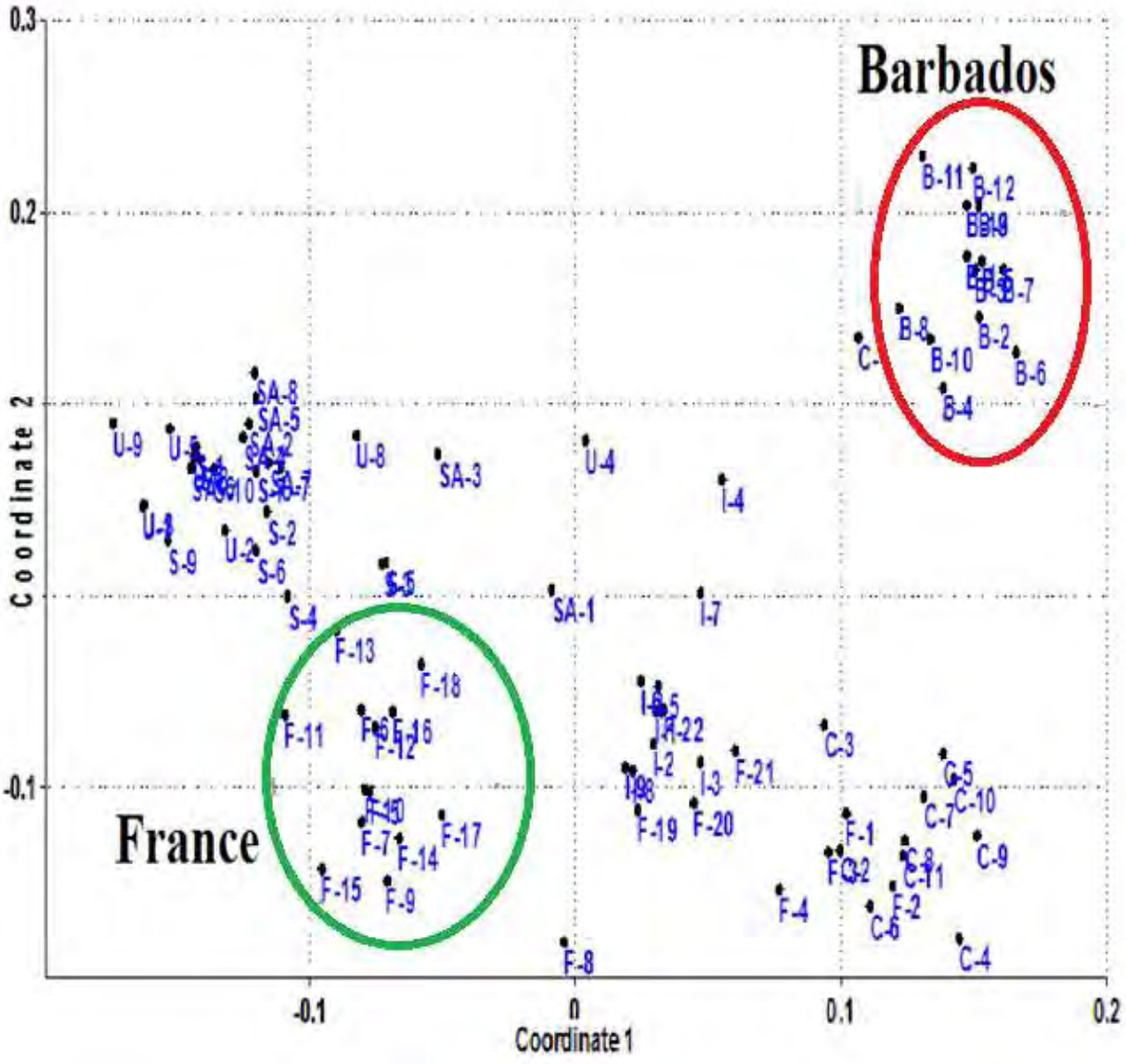


Figure 9. Two dimensional (2D) representation and relationships of 82 genotypes of seven sugarcane populations on Principal Coordinate analysis based on Jaccard's similarity coefficients. Key: B = Barbados, C = Cuba, F = France, I = India, SA = South Africa, S = Sudan, U = USA.

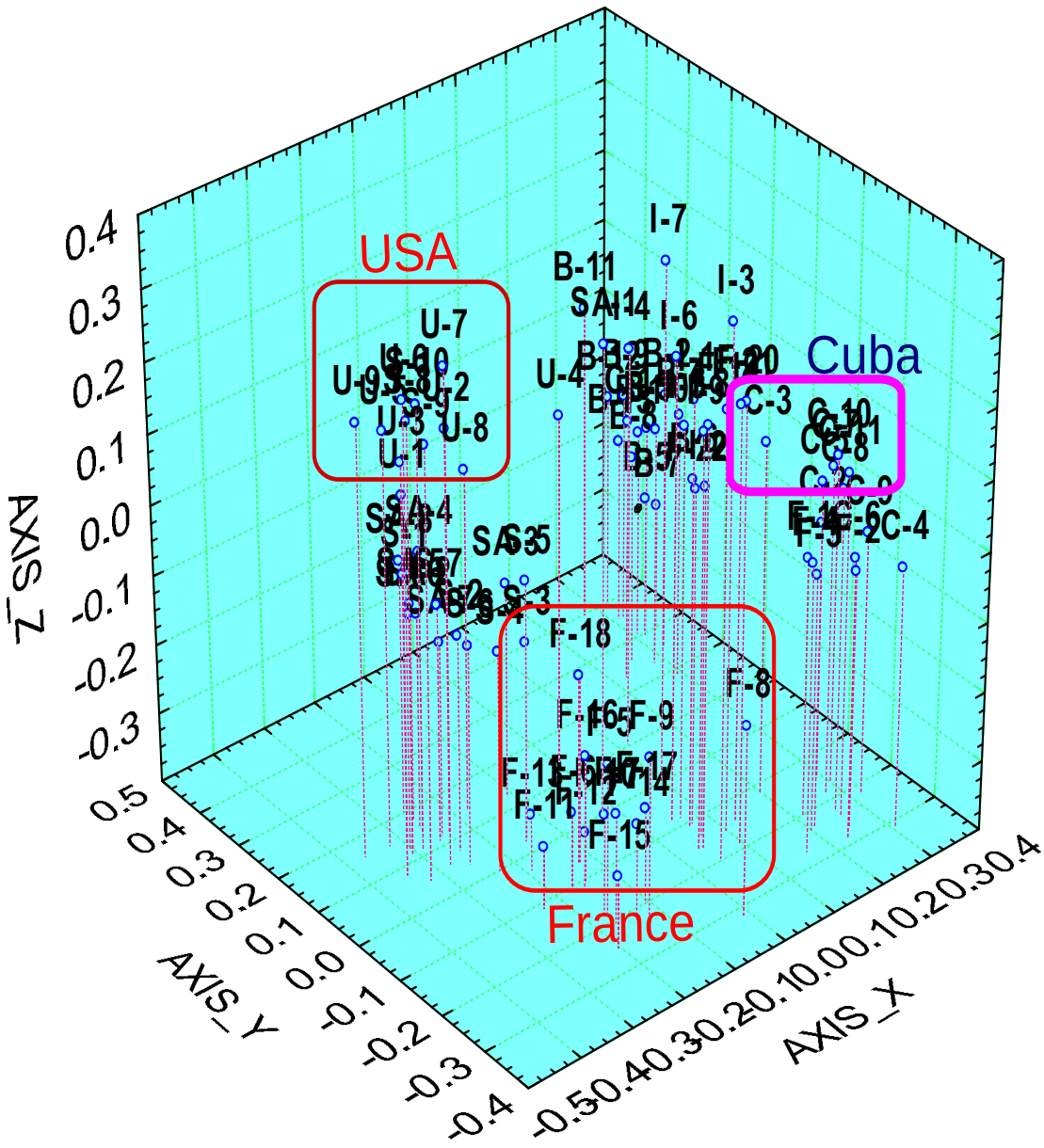


Figure 10. Three dimensional (3D) representation of Principal Coordinate analysis of genetic relationships among 82 genotypes of sugarcane as revealed by ISSR marker generated by 12 primers, Key: B = Barbados, C = Cuba, F = France, I = India, SA = South Africa, S = Sudan, U = USA.

6. DISCUSSION

6.1. ISSR polymorphism and application for genetic diversity study in sugarcane

In this study, ISSR markers were utilized to investigate genetic variability and relationships within and among the seven groups of introduced sugarcane genotypes in Ethiopia. This report makes this study the first investigation to assess genetic diversity within and among sugarcane genotypes using molecular marker from Ethiopia.

The number of polymorphic loci ranged from 7-16 fragments per primer displaying 83.22 % of polymorphism with high gene diversity (h = 0.310) and Shannon's information index (I = 0.458). However, in terms of percent polymorphism, gene diversity, and Shannon's information index, per classes of primer, penta-nucleotide primers were found to be superior (PPL = 87.10 %, h = 0.365, I = 0.526) followed by the di-nucleotide primers (PPL = 83.87 %, h =0.306, I = 0.455). These all clearly indicate that the extent and distribution of penta- and di-nucleotide repeats polymorphism in sugarcane genome used for this study is very high and every individual sugarcane genotype is almost unique for these primers.

The di-nucleotide primers 810 and 824 with (GA) and (TC) repeats anchored with (T) and (G) respectively, detect high level of polymorphism (90.91 %) among the sugarcane populations followed by primer 811 (90.00 %) with (GA) repeats anchored with (C). Similarly, Reddy *et al*. (2002) point out that, generally primers with (AG), (GA), (CT), (TC), (AC), and (CA) repeats show higher polymorphism in plants than

primers with other di-nucleotide repeats, while tri- and tetra-nucleotide repeats are less frequent in plants and their use in ISSRs is lower than the di-nucleotide repeats. Furthermore, the choice of appropriate primer repeat motives in ISSR fingerprint is critical to detect high level of polymorphism and reveal relationship within and among populations. Similar results were reported by Da Costa *et al*. (2011) (GA), (CA), (AG), (CT), and (TC) repeats based primers have shown to amplify clear bands and reveal high level of polymorphism in sugarcane using ISSR marker. Blair *et al*. (1999) and Reddy *et al*. (2002) reported that the (AG) and (GA) based ISSR primers have been shown to amplify clear bands in rice. Joshi *et al*. (2000) also reported comparative assessments of ISSR primer repeat patterns of di-, tri-, tetra- and penta-nucleotides in the genus *Oryza* and found that di-nucleotide repeats of (AG)8 and (GA)8 with a number of anchors gave the best polymorphic and informative patterns. Therefore, such preliminary information on SSR motif distribution and levels of variation could be useful in characterizing widely distributed polymorphism of microsatellite markers in target taxa (Kassahun Tesfaye *et al*., 2014).

The overall analysis of the current study showed that the ISSR markers detected high level of polymorphism, gene diversity, and Shannon's information index at species level (PPL = 83.22 %, h = 0.310, and I = 0.458). This indicates the existence of high level of genetic diversity within and among introduced sugarcane genotypes. Similar findings were reported for other accessions during the evaluation of the genetic diversity of sugarcane varieties from India with ISSR markers by Srivastava and Gupta (2008) with 78.48 % of the bands being polymorphic. Moreover, Smiullah *et*

al. (2013) detected 85.25 % polymorphism with SSR markers on accession from Pakistan.

The result obtained in the current study bear higher polymorphism levels than those previously reported by other authors using RAPD markers (Nair *et al*., 2002; Kawar *et al*., 2009). The high level of genetic polymorphism detected by ISSR markers in our study was expected due to the use of newly released hybrid varieties from different countries of sugarcane breeding institutes and also considering allopolyploid nature of sugarcane, generally attributed to the interspecific hybridization crosses used during breeding programs that generated the actual breeding varieties (Da Costa *et al*., 2011; Pandey *et al*., 2011). The nature of ISSR, targeting regions especially rich in microsatellites may also justify the higher level of polymorphism, since those regions are known to accumulate a larger number of mutations in DNA during replication and unequal crossing-over (Schlotterer and Tautz, 1992). The complexity of the sugarcane genome and the abundance of microsatellite repeats make also this a priority (Pandey *et al*., 2011). Therefore, this study has identified polymorphic ISSR markers for genotyping sugarcane genotypes and the data proved useful for examining genetic relationships and diversity among the introduced sugarcane genotypes in Ethiopia.

6.2. Diversity among sugarcane genotypes in Ethiopia

As pointed out by Ullah *et al*. (2013), a basic understanding of the genetic diversity that exists in the germplasm available for breeding is essential to the success of a breeding program. This knowledge is useful in the utilization and management of genotypes and indeed genes in the breeding gene pool. Moreover, rapid advances in the field of molecular biology and its allied sciences made the use of molecular markers a routine practice providing plant breeders a precise tool in analyzing genetic diversity for plant improvement. Hence, for achieving improved productivity in sugarcane crops, it is essential to maintain a high degree of genetic diversity among the commercial varieties and breeding populations since sugarcane is a clonally propagated crop, and creation of new genotypes is done through sexual crossing (Que *et al*., 2014).

Choosing parental lines is the most crucial step in any sugarcane improvement program. There has never been single incidence of developing a sugarcane cultivar out of a poor cross (Ming *et al*., 2006). Therefore, genetic diversity analysis of sugarcane germplasm based on molecular evaluation and characterization is the basis for effective germplasm utilization. A high genetic diversity and complementarity between two parental lines are crucial for producing high quality seedling populations of hybrid progeny (Ming *et al*., 2006).

In the present study, the ISSR markers found to be an appropriate molecular marker system for generating the detailed intraspecific genetic diversity data to evaluate the extent and distribution of genetic diversity within and among sugarcane genotypes

introduce from abroad. Based on this, all the diversity parameters confirm that there is high genetic diversity in sugarcane genotypes (PPL = 83.22 %, h = 0.310, I = 0.458).

With regard to intra varietal diversity, France cultivars have been found to be a more diverse population as compared to other populations studied in the overall analysis (PPL = 47.65 %, h = 0.171, I = 0.255) followed by Cuba population (PPL = 44.30 %, h = 0.164, I = 0.241), while the least variable population was from Barbados (PPL = 28.86 %, h = 0.097, I = 0.147). This could be associated with the original sources of genotypes and wider genetic bases among individuals used for breeding program in France and Barbados. Moreover, the introduced genotypes in Ethiopia are from various breeding institutes globally with different genetic origin.

Rodriguez *et al*. (2005) found high percentage of similarity among the sugarcane cultivars with AFLP markers and suggests that the sugarcane germplasm collections from Mexico present a genetically narrow base. This is likely due to the fact that their parental breeding lines, used to develop these cultivars, are the same or are very close to each other. Theoretically, a hybridization event followed by selfing would be predicted to decrease the number of polymorphic loci in subsequent generations by 50 % (Abdalla *et al*., 2001). To this respect D'Hont *et al*. (1995) reported, because of the small number of clones of these species used in the primary crosses, the genetic base of modern hybrid varieties appears to be narrow and could be the reason for the present slow progress in sugarcane breeding. However, in this study, analysis of ISSR diversity identified genetically unique genotypes from France and Cuba that are

potentially important source of diverse sugarcane germplasm along with divergent individuals.

Similarly, analysis on the basis of geographical regions, viz Africa, North America, Europe, and Asia, showed a range of value of percent polymorphism, gene diversity, and Shannon information's index. Among the four geographic groups, North America attributed to the highest genetic diversity (PPL = 71.14 %, h = 0.270, I = 0.397) as compared to others and Africa had also high genetic diversity (PPL = 57.72 %, h = 0.213, I = 0.314), ranking the second. This high level of genetic diversity might be explained probably due to breeding scheme used. This is meaning that the use of either diverse or closely related parental lines for developing improved varieties in different countries or due to large sample size.

The overall analyses of percent of polymorphism, gene diversity, Shannon index, gene flow (Nm), and genetic differentiation (G_{ST}) for all study populations with 12 ISSR primers of the current study were 83.22 %, 0.310, 0.458, 0.401, and 0.555, respectively, as mentioned above. In this study, the Nei's gene diversity (h = 0.31) is relatively closer to what was recently found in sugarcane varieties by Que *et al*. (2014) in their effort to study genetic analysis of diversity within a Chinese local sugarcane germplasm based on start codon targeted polymorphism (h = 0.35) but much higher than which was reported by You *et al*. (2013) (h = 0.17) using SSR markers. Lu *et al*. (1994b) and Janno *et al*. (1999) also reported that due to its polyploid nature, interspecific origin, and vegetative mode of propagation, high levels of heterozygosity were detected among sugarcane cultivars using RFLP markers.

Similar patterns of molecular diversity were also detected using AFLP (Lima *et al*., 2002) and SSR markers (Pandey *et al*., 2011).

Similarly, our study revealed high genetic diversity among the sugarcane genotypes commonly used in crossing programs and this opens up an excellent opportunity for their utilization in sugarcane breeding programs since intercrossing of elite hybrids is a main component in varietal development and such studies based on molecular marker information can be helpful in selecting diverse parental combinations giving maximum diversity with respect to sugar content.

6.3. Genetic distance among sugarcane genotypes and its implications for varietal improvement

The genetic distances that exist among various genotypes included in this study were determined through Jaccard's coefficients of similarity. Accordingly, genotypes from USA and Barbados were the most divergent and would be appropriate for crossing program as the genetic bases were wider. If genotypes with wider genetic bases were crossed, one would expect a higher genetic diversity among the cross progeny and generations derived from the crosses (Que *et al*., 2014).

The result from Jaccard's pairwise similarity showed that Sudan and South Africa populations were the closest with similarity coefficient of 0.70, followed by USA with South Africa and USA with Sudan with similarity coefficient of 0.651 and 0.653, respectively. In contrast, USA and Barbados were observed to be the most distantly related populations with 0.53 (from Jaccard's pairwise similarity) followed by USA and Cuba with 0.54 similarity coefficient. In fact, USA, Barbados, and Cuba

populations are geographically very far apart but these populations were assumed to be grouped in the same regional continent (i.e., North American continent). In this respect, geographical distance was not correlated with similarity of genetic diversity.

The result in this study is in agreement with the studies of Abreham Bekele *et al*. (2014) with ISSR marker who did not find correlations between geographic distances and genetic differences among anchote populations. Similarly, both Almaz Negash *et al*. (2002), using ALFP and Genet Birmeta *et al*. (2004), using RAPD didn't find any correlation between genetic diversity and agro-ecological zone in their studies on enset. Accordingly, in this study, genetic distance was not entirely correlated with geographic distance. Hence, this could be due to the widespread practice of exchanging commercial and parental clones in between sugarcane growing regions for the purpose of commercial production and breeding activities.

In general, populations with high intra-population genetic diversity were genetically distantly related to populations with low intra-population diversity. Crossing between any distantly related populations is therefore expected to yield more heterotic and vigorous plants constituting much of the different traits contained in the two parental lines (Edossa Fikiru *et al*., 2007). Hence, hybridization or crossing between distantly related populations of the present study could be an appropriate strategy for the ongoing sugarcane breeding programs in Ethiopia.

6.4. Genetic differentiation and structure among sugarcane genotypes

According to Schaal *et al*. (1998) and Weising *et al*. (2005), the extent of genetic variation in a species and its distribution among and within populations is determined by the interactions of various factors, including the long-term evolutionary history of the species (habitat fragmentation and population isolation), genetic drift, mating system, gene flow, mutation, and selection. The general picture arising from these comparisons is that long-lived perennial and outcrossing species retain most of their variation within populations whereas annual and selfing species allocate more variation among populations. For dominantly inherited DNA markers, genetic differentiation among populations is often estimated with G_{ST} according to Nei (1973). However, a multi-locus approach, AMOVA (analysis of molecular variance) is nowadays even more widely used than G_{ST} for the partitioning of genetic variation (Excoffier *et al*. 1992). According to De Vicente and Fulton, (2003) the values of F_{ST} from AMOVA can be grouped into four categories with respect to genetic differentiation; very high ($F_{ST} > 0.25$), high (0.15 - 0.25), intermediate (0.05 - 0.15), and low (0.0 - 0.05).

In this study, the estimate of population differentiation in sugarcane using F_{ST} and G_{ST} were 0.36 and 0.55, respectively, but the value of G_{ST} may not be as strong as F_{ST} as pointed out by Culley *et al*. (2002) since G_{ST} is dependent on sample sizes and number of populations, in addition to its reliance on Hardy-Weinberg genotype proportions, and conditions that may be violated when analyzing small isolated populations, while the value of F_{ST} is more recommended in various literatures

(Laurentin, 2009) particularly for multi-locus markers since F_{ST} is the degree of gene differentiation among populations in terms of allele frequencies (De Vicente and Fulton, 2003). However, in both cases i.e., using F_{ST} and G_{ST} the present study revealed relatively high level of genetic differentiation among populations. This could be attributed to mutation of SSR loci (annealing site for ISSR primers), random genetic drift, and differential selection pressure (by the environment) on the loci assessed (Edossa Fikiru *et al*., 2007).

Using SSR marker on sugarcane cultivars from China, the analysis of molecular variance (AMOVA) revealed that the majority of variation detected in the SSRs was within populations (90.5 %) with only 9.5 % of the variation was attributed to differences among populations (You *et al*., 2013). Que *et al*. (2014) also reported that a high level of genetic variation was found within populations rather than among populations of sugarcane using start codon targeted (SCoT) polymorphism. Similar finding was reported by Glynn *et al*. (2009) using SSR marker on sugarcane cultivars from USA, the values of AMOVA across the populations indicated that intrapopulation genetic variation (96.6 %) was much higher than that of interpopulation (3.4 %).

The AMOVA results obtained in the current study do not contradict with the above findings. It is a prevalent view that outcrossing and long-lived perennial species retain higher genetic variation within populations than among populations. Though sugarcane is a cross-pollinating species (Berding *et al*., 2004). In accordance with this, AMOVA analysis in the present study revealed higher within population genetic

variation (63.56 %) than among population genetic variation (36.44 %). Similarly, for grouping, AMOVA showed a higher genetic variation within populations (61.66 %) but lower variation exists among populations within groups (34.88 %) and among groups (3.46 %) so that the majority of the ISSR diversity was distributed among individual genotypes within the populations. This is due to gene flow during hybridization or crossing program in an effort to develop improved sugarcane varieties for sustaining the fast and ongoing expansion of sugar factories worldwide.

The value of gene flow (Nm) estimate from POPGENE software package indicates the approximate number of individuals migrating from one population to the other. In general, if Nm < 1, the genetic drift would lead to genetic differentiation among populations then, local differentiation of populations will result but if Nm > 1, gene flow can resist the effect of genetic drift within populations and prevent the differentiation of populations then, there will be no significant differentiation among populations (Slatkin, 1985). The fact that estimates of Nm was < 1 suggests that gene flow among populations is insufficient to counter the effects of random genetic drift and genetic differentiation results from genetic drift only if Nm < 1, but not if Nm > 1 (Slatkin, 1985). As reported by Slatkin (1985) Nm values can be grouped into three categories: high (Nm ≥ 1.00), intermediate (0.25 - 0.99), and low (0.00 - 0.24).

In our study, relatively high genetic differentiation and intermediate level of gene flow (Nm = 0.401) was detected indicating that genetic drift has greatly affected the genetic composition of individual populations. However, the limited gene flow among populations could be attributed to geographical isolation/barrier, clonal way of

propagation, and limited number of parental clones used in crossing programs (Nair *et al*., 1999; Nair *et al*., 2002; Selvi *et al*., 2003; Berding *et al*., 2004; Rodriguez *et al*., 2005). This limited gene flow, however, obtained largely from the interspecific crosses involving *Saccharum officinarum* and the wild species *S. spontaneum* in sugarcane breeding program since their interspecific F1 hybrids could form complete pairing in meiosis thus resulting in gene flow between them. It is also believed that the high genetic diversity observed within sugarcane populations studied, might be due to preferential adaptive gene complexes being evolved during long evolutionary period in the regions adapted to environmental changes. Moreover, introduction of genotypes from various breeding institutions and lineages could be another explanation for the observed variation.

6.5. Genetic relationships in sugarcane genotypes

UPGMA, NJ, and PCoA clustering analyses were employed to better visualize the genetic relationships of the 82 genotypes of sugarcane in Ethiopia that originated from different countries abroad. On the basis of UPGMA, most of the genotypes of the respective population were observed to form clear grouping with respective populations except few intermixing with other populations (Figure 7). All genotypes from Barbados, Cuba, and India grouped in one major cluster and the genetic similarity values among genotypes were relatively high, while genotypes from France was clustered in both major clusters since these genotypes had high genetic diversity than others. Similarly, genotypes from USA, Sudan, and South Africa formed different groups involving most of the genotypes from the same populations.

Generally, the UPGMA analysis in this study revealed that most of the genotypes that originated from the same geographic region showed close genetic relationships and were grouped into the same cluster. This result is congruent with the studies of Chen *et al*. (2009) who used SSR markers and found that based on Jaccard's similarity coefficient 40 sugarcane cultivars from different geographical regions tend to cluster with their respective origins. In NJ analysis, the majority of the genotypes of the respective population were observed to form similar grouping patterns with that of UPGMA but moderate intermixing with other group and also some of the individuals from each population appear to have longer branches which indicate that they are more variable. This is in accordance with Weising *et al*. (2005), UPGMA assumes a rigid molecular clock, which means that the evolutionary rates along all branches of

the tree need to be identical, while neighbor joining (NJ) algorithm produces additive trees and does not assume identical evolutionary rates along all branches.

From the two distinct major clusters of UPGMA population based dendrogram (Figure 6), the first bottom cluster contains populations of USA, Sudan, and South Africa. The highest within cluster similarity was, however, found between Sudan and South Africa followed by USA versus Sudan and USA versus South Africa. Surprisingly, these populations are geographically unrelated but made a cluster in one major group. This might be because South Africa extensively uses cultivars from Sudan and USA as commercial and parental clones in South Africa Sugarcane Research Institute (SASRI) and vice versa. South Africa relied on imported varieties for its initial sugarcane production (Zhou, 2013). When imported varieties succumbed to diseases and pests and were not adapted to local conditions then the South African sugar industry established the South Africa Sugarcane Research Institute (SASRI) with the mandate to develop improved sugarcane varieties (Zhou, 2013).

The second top major cluster contained populations of Barbados, Cuba, France, and India. This is mainly due to the involvement of sugarcane genotypes in breeding programs for improving some of the traits of commercially exploited varieties so that to achieve such objectives there might be parental line exchange throughout sugarcane growing regions in the world. On the other hand, countries may use the cross progenies that are released by Center for Research in Regional Planning and Development (CRAD) in France and Indian Institute of Sugarcane Research (IISR). As a general point of view, it was not surprising that geographically different

populations made a cluster with other populations since most of the cultivars that were bred after the turn of the 20th century are interspecific hybrids between *S. officinarum* and *S. spontaneum*. However, cultivars inherited a greater proportion of the *S. officinarum* genome as 'nobilization' involved several backcrosses to the *S. officinarum* parent during which this parent transmitted the chromosome to its cross progeny (D'Hont *et al*., 1996). Hence, it might be the cross progeny clustered with their progenitor parents from other regions or the parents may have clustered with their cross progeny from distantly related populations. This is in agreement with the results obtained by microsatellite markers (Selvi *et al*., 2003) wherein the highest genetic similarity was observed between the cross progeny and their progenitor parents and they cluster in the same group despite of their geographic origin. Rodriguez *et al*. (2005) also reported that one explanation for the high levels of similarity within sub-groups and groups is that the lineages have been subjected to a greater degree of intercultivar gene flow.

7. CONCLUSION

This study is the first of its kind that utilized ISSR marker technique to assess genetic diversity and relationships among introduced genotypes of sugarcane in Ethiopia. However, the scanty information available on the genetic diversity within and between sugarcane varieties has been based mainly on morphological characteristic. This necessitated the estimation of genetic diversity based on molecular markers which could provide more accurate information to sugarcane breeder. This data will support the exploitation of sugarcane germplasm on molecular basis. Furthermore, the level of polymorphism observed proved that the ISSR marker system was robust at amplifying markers on sugarcane and did so according to the complex genome structure interconnecting with interspecific origin. Hence, genetic diversity could be evaluated using ISSR markers for the trait(s) of interest and genotypes or species displaying unique diversity selected for germplasm enhancement.

Results from all diversity parameters showed that France and Cuba populations exhibited high genetic diversity, while Barbados and South Africa populations had lowest genetic diversity and should get special attention in order to balance genetic resources of such beneficial genes across the various agroecology and genetic origin.

AMOVA of the present study showed that intra-population/country diversity was significantly higher than inter-population/country diversity. From Jaccard's pairwise similarity, the highest genetic similarity was found between Sudan and South Africa, while the most distantly related populations were USA and Barbados. Clustering analysis with UPGMA, NJ, and PCoA suggested that there were some intermixed

individuals from one to other populations and no entire strong correlation was found between geographic and genetic distance.

Finally, ISSR markers could potentially be useful in the characterization and management of germplasm where the aim is to enhance the germplasm for specific traits and future breeding efforts involving crosses between and within the groups identified in this study may provide useful strategies for combining beneficial genes and alleles in developing improved sugarcane varieties.

8. RECOMMENDATIONS

Based on the findings obtained in this study, the following recommendations were forwarded.

- Sugarcane breeding program should focus on divergent genotypes from France and Cuba populations and also from distantly related populations of USA and Barbados.
- Genotypes from Barbados were observed to show low genetic diversity and hence, special attention should be give to conserve these genotypes and/or study their unique character.
- This study is the first attempt to reveal existence of genetic diversity in sugarcane at molecular level using ISSR markers. However, a study (using other markers or combination of markers) such as the codominant markers like SSR are so indispensible for comparative analysis for better inference and estimate of gene flow.
- Characterization of more populations is also very important and an in depth study should be made to indicate level of diversity among populations and agro-ecological zones.
- Local landraces should also be studied with ISSR and SSR markers.
- Penta-nucleotide and di-nucleotide primers could be used as a preliminary data for further molecular analysis.

9. REFFERENCES

Abdalla, A. M., Reddy, O. U. K., El-Zik, K. M., Pepper, A. E. (2001). Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor. Appl. Genet.* **102:** 222-229.

Abreham Bekele, Tiliye Feyissa, and Kassahun Tesfaye (2014). Genetic diversity study of anchote (*Coccinia abyssinica* (Lam.) Cogn.) using Inter Simple Sequence Repeat (ISSR) markers. *Genet. Resour. Crop Evol.* **42:** 1105-1111.

Acquaah, G. (2007). *Principles of Plant Genetics and Breeding* 1st edn. Blackwell Publisher Ltd, London, UK, pp. 564.

Akhtar, M., Jemil, M. and Ahmad, S. (2006). Agronomic traits and morphological characteristics of some exotic varieties of sugarcane. *Pak. J. Agr. Res.* **19:** 70-77.

Almaz Negash, Admasu Tsegaye, Treuren, R.V. and Visser, B. (2002). AFLP analysis of enset clonal diversity in South and Southwestern Ethiopia for conservation. *Crop Sci.* **42**: 1105-1111.

Alwala, S., Suman, A., Arro, J. A., Veremis, J. C. and Kimbeng, C. A. (2006). Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. *Crop Sci.* **46:** 448-455.

Arif, I. A., Bakir, M. A., Khan, H. A., Al Farhan, A. H., Al Homaidan, A. A., Bahkali, A. H., Al Sadoon, M. and Shobrak, M. (2010). A brief review of molecular techniques to assess plant diversity. *Int. J. Mol. Sci.* **11:** 2079-2096.

Artschwager, E. and Brandes, E. W. (1958). *Sugarcane (Saccharum officinarum L.) Origin, Classification, Characteristics, and Descriptions of Representative Clones.* United

States Department of Agriculture, Washington, D.C., USA, pp. 21.

Bagali, P. G., Antony Herold Prabhu, P. D., Raghavendra, K., Hittalmani, S. and Vadivelu, J. S. (2010). Application of molecular markers in plant tissue culture. *AsPac. J. Mol. Biol. Biotech.* **18:** 85-87.

Barber (1931). History and development of sugarcane as a world crop. *J. Sugarcane* **1:** 1-5.

Bardakci, F. (2001). Random amplified polymorphic DNA (RAPD) markers. *Turk J. Biol.* **25:** 185-196.

Berding, N., Hogarth, M. and Cox, M. (2004). Plant improvement of sugarcane. **In:** *Sugarcane*, pp. 20-46, (James, G., ed.). Blackwell Science Ltd, London, UK.

Besse, P., Mclntyre, C. L. and Berding, N. (1997). Characterisation of *Erianthus* sect. Ripidium and *Saccharum* germplasm (Andropogpnease Saccharinae) using RFLP markers. *Euphytica* **93:** 283-292.

Blair, M. W., Panaud, O. and McCouch, S. R. (1999). Inter simple sequence repeats (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). *Theor. Appl. Genet*. **98:** 780-792.

Borsch, T., Hilu, K.W., Quandt, D., Wilde, V., Neinhuis, C. and Barthlott, W. (2003). Non coding plastid trnT-trnF sequences reveal a well resolved phylogeny of basal angiosperms. *J. Evol. Biol*. **16:** 558-576.

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. *Am. J. Hum. genet*. **32:** 314-331.

Carrera, A. D., Poverene, M. M. and Rodriguez, R. H. (1996). Isozyme variability in *Helianthus agrophyllus*: Its application in crosses with cultivated sunflower. *Helia.* **19:** 19-28.

Cheavegatti-Gianotto, A., Abreu, H. M., Arruda, P Filho, J. C. B., Burnquist, W. L., Creste, S., Di Ciero, L., Ferro, J. A., Figueira, A. V. O., Filgueiras, T. S., DeSá, M. F. G., Guzzo, E. C., Hoffmann, H. P., Landell, M. G. A., Macedo, N., Matsuoka, S., Reinach, F. C., Romano, E., Da Silva, W. J., Filho, M. C. S. and Ulian, E. C. (2011). Sugarcane (*Saccharum officinarum*): A reference study for the regulation of genetically modified cultivars in Brazil. *Trop. Plant Biol.* **4:** 62-89.

Chen, P. H., Pan, Y. B., Chen, R. K., Xu, L. P. and Chen, Y. Q. (2009). SSR marker-based analysis of genetic relatedness among sugarcane cultivars (*Saccharum* spp. hybrids) from breeding programs in China and other countries. *Sugar Technol*. **11:** 347-354.

Cho, Y. G., McCouch, S. R., Kuiper, M., Kang, M. R., Pot, J., Groenen, J. T. M. and Eun, M. Y. (1998). Integrated map of AFLP, SSLP, and RFLP markers using a recombinant inbred population of rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **97:** 370-380.

Cordeiro, G. M. (2001). Molecular marker systems for sugarcane germplasm analysis. **In:** *Plant Genotyping: the DNA Fingerprinting of Plants*, pp. 129-146, (Henery, R. J., ed.). CABI Publisher, London.

Culley, T. M., Wallace, L. E., Gengler-Nowak, K. M. and Crawford, D. J. (2002). A comparison of two methods of calculating G_{ST}, a genetic measure of population differentiation. *Am. J. Bot.* **89:** 460-465.

Cunff, L. L., Garsmeur, O., Raboin, L. M., Pauquet, J., Telismart, H., Selvi, A., Grivet, L.,

Philippe, R., Begum, D., Deu, M., Costet, L., Wing, R., Glaszmann, J. C. and D'Hont, A. (2008). Diploid/polyploid syntenic shuttle mapping and haplotype-specific chromosome walking toward a rust resistance gene (bru1) in highly polyploid sugarcane (2n 12x 115). *Genet*. **180:** 649-660.

Da Costa, M. L. M., Amorim, L. L. B., Onofre, A. V. C., Tavares de Melo, L. J. O., de Oliveira, M. B. M., de Carvalho, R. and Benko-Iseppon, A. M. (2011). Assessment of genetic diversity in contrasting sugarcane varieties using Inter Simple Sequence Repeat (ISSR) markers. *Am. J. Plant Sci*. **2:** 425-432.

Daniels, J. and Roach, B. T. (1987). Taxonomy and evolution. **In:** *Sugarcane Improvement through Breeding*, pp. 7-84, (Heinz, D. J., ed.). Elsevier, Amsterdam, Netherlands.

De Vicente, M. C. and Fulton, T. (2003). Molecular marker technology in studies on plant genetic diversity. IPGRI, Rome, Italy and Institute for Genetic Diversity, Ithaca, New York, USA.

D'Hont, A., Rao, P. S., Feldmann, P., Grivet, L., Islam-Faridi, N., Taylor, P. and Glaszmann, J. C. (1995). Identification and characterization of sugarcane intergeneric hybrids, *Saccharum officinarum* X *Erianthus arundinaceus*, with molecular marker and DNA in situ hybridization, *Theor. Appl. Genet*. **91:** 320-326.

D'Hont, A., Gravet, L., Fildman, P., Rao, S., Berding, N. and Glasman, J. C. (1996). Characterization of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. *Mol. Gen. Genet*. **250:** 405-413.

D'Hont, A., Ison, D., Alix, K., Roux, C. and Glaszmann, J. C. (1998). Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome* **41:** 221-225.

Duttamajumder, S. K., Sharma, A. K. and Prakash, B. (2011). *Indian Institute of Sugarcane Research*. Army Printing Press, New Delhi, India, pp. 28.

Edossa Fikiru, Kassahun Tesfaye, and Endashaw Bekele (2007). Genetic diversity and population structure of Ethiopian lentil (*Lens culinaris* Medikus) landraces as revealed by ISSR marker. *Afr. J. Biotech.* **6:** 1460-1468.

Ejeta, G., Goldsbrough, P. B., Tuinstra, M. R., Grote, E. M., Menkir, A., Ibrahim, Y., Cisse, N., Weerasuriya, Y., Melak-Berhane, A. and Shaner, C. A. (1999). Molecular marker applications in sorghum. **In:** *Application of Molecular Markers in Plant Breeding*, pp. 81-89, (Haussmann, B. I. G., Geiger, H. H., Hess, D. E., Hash, C. T. and Bramel-Cox, P., eds.). ICRISAT, Patancheru, Andhra Pradesh, India.

Ethiopian Investment Agency (EIA) (2008). Investment opportunity profile for sugarcane plantation and processing in Ethiopia. Addis Ababa, Ethiopia.

Eksomtramage, T., Paulet, F., Noyer, J. L., Feldmann, P. and Glaszmann, J. C. (1992). Utility of isozymes in sugarcane breeding. *Sugarcane* **3:** 14-21.

Ethiopian Sugar Development Agency (ESDA) (2009). Annual Review. Addis Ababa, Ethiopia.

Ethiopian Sugar Development Agency (ESDA) (2013). Annual Review. Addis Ababa, Ethiopia.

Ethiopian Sugar Development Agency (ESDA) (2014). Annual Review. Addis Ababa, Ethiopia.

Esayas Aga, Endashaw Bekele, and Bryngelsson, T. (2005). Inter simple Sequence Repeat (ISSR) variation in forest coffee trees (*Coffea arabica* L.) populations from Ethiopia. *Genet.* **124:** 213-21.

Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genet*. **131:** 179-191.

Excoffier, L., Laval, G. and Schneider, S. (2006). Arlequin Ver 3.01, An Integrated Software Package for Population Genetics Data Analysis, Computational and Molecular Population Genetics Lab (CMPG), Institute of Zoology, University of Berne, Switzerland, *http://cmpg.unibe.ch/software/arlequin3*

Fang, D. Q. and Roose, M. L. (1997). Identification of closely related citrus cultivars with Inter Simple Sequence Repeat markers. *Theor. Appl. Genet*. **95:** 408-417.

Genet Birmeta, Nybom, H. and Endashaw Bekele (2004). Distinction between wild and cultivated enset (*Ensete ventricosum*) gene pools in Ethiopia using RAPD markers. *Hered.* **140:** 139-148.

Gezahegn Girma, Kassahun Tesfaye, and Endashaw Bekele (2010). Inter Simple Sequence Repeat (ISSR) analysis of wild and cultivated rice species from Ethiopia. *Afr. J. Biotech.* **9:** 5048-5059.

Gilbert, J. E., Lewis, R. V., Wilkinson, M. J. and Caligari, P. D. S. (1999). Developing an appropriate strategy to assess genetic variability in plant germplasm collections. *Theor. Appl. Genet.* **98:** 1125-1131.

Glaszmann, J. C., Fautret, A, Noyer, J. L., Feldmann, P. and Lanaud, C. (1989). Biochemical genetic markers in sugarcane. *Theor. Appl. Genet*. **78:** 537-543.

Glynn, N. C., McCorkle, K., Jack, C. and Comstock, J. C. (2009). Diversity among mainland USA sugarcane cultivars examined by SSR genotyping. *J. Am. Soci. Sugarcane Technol.* **29:** 36-52.

Grivet, L., D'Hont, A., Roques, D., Feldmann, P., Lanaud, C. and Glaszmann, J. C. (1996). RFLP mapping in cultivated sugarcane *(Saccharum* spp.): genome organization in a highly polyploid and aneuploid inter-specific hybrid. *Genet*. **142:** 987-1000.

Grivet, L., Daniels, C., Glaszmann, J. C. and D'Hont, A. (2004). A review of recent molecular genetics evidence for sugarcane evolution and domestication. *Ethnobot. Res. Appl*. **2:** 9-17.

Gupta, M., Chyi, Y. S., Romero, S. J. and Owen, J. L. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple sequence repeats. *Theor. Appl. Genet*. **89:** 998-1006.

Gupta, P. K., Varshney, R. K., Sharma, P. C. and Ramesh, B. (1999). Molecular markers and their applications in wheat breeding. *Plant Breed.* **118:** 369-390.

Gupta, P. K., Varshney, R. K. and Prasad, M. (2002). Molecular Markers: principles and methodology. **In:** *Molecular Techniques for Crop Improvement*, pp. 9-54, (Jain, S. M., Brar, D. S. and Ahloowalia, B. S., eds.). Kluwer Academic Publishers, Boston.

Ha, S., Moore, P. H., Heinz, D., Kato, S., Ohmido, N. and Fukui, K. (1999). Quantitative chromosome map of the polyploid *Saccharum spontaneum* by multicolor fluorescence in situ hybridization and imaging methods. *Plant Mol. Biol.* **39:** 1165-1173.

Hadrys, H., Balick, M. and Schierwater, B. (1992). Applications of Random Amplified Polymorphic DNA (RAPD) in molecular ecology. *Mol. Ecol*. **1:** 55-63.

Hammer, O., Harper, D. and Ryan, P. (2001). PAST: Paleontological statistics software for education and data analysis. *Paleontologia Electronica* **4:** 1-9.

Hartl, D. L. and Clark, A. G. (1997). *Principles of Population Genetics* 3rd edn. Sinauer Assoc., Sunderland, Massachsetts, pp. 525.

Hartl, L., Mohler, V., Zeller, F. J., Hsam, S. L. K. and Schweizer, G. (1999). Identification of AFLP markers closely linked to the powdery mildew resistance genes *Pm1c* and *Pm4a* in common wheat (*Triticum aestivum* L.). *Genome* **42:** 322-329.

Hayden, M. J., Stephenson, P., Logojan, A. M., Khatkar, D., Rogers, C., Elsden, J., Koebner, R. M. D., Snape, J. W. and Sharp, P. J. (2006). Development and genetic mapping of sequence-tagged microsatellites (STMs) in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet*. **113:** 1271-1281.

Hedrick, P. W. (2011). *Genetics of Populations* 3rdedn. Jones and Bartlett Publisher, LLC, New York, pp. 663.

Henry, R. J. (2001). *The DNA Fingerprinting of Plants: Plant Genotyping.* CAB1 Publisher Wallingford, UK, pp. 512.

Hokanson, S. C., Lamboy, W. F., Szewc-McFadden, A. K. and McFerson, J. R. (2001). Microsatellite (SSR) variation in a collection of Malus (apple) species and hybrids. *Euphytica* **118:** 281-294.

International Sugar Organization (ISO) (2009). Sugar year-book. London, UK.

Jaccard, P. (1908). Nouvelles researchers Sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat*. **44:** 223-270.

Jannoo, N., Grivet, L., Seguin, M., Paulet, F., Domaingue , R., Rao ,P. S., Dookun , A., D'Hont , A. and Glaszmann, J. C. (1999). Molecular investigation of the genetic base of sugarcane cultivars. *Theor. Appl. Genet*. **99:** 171-184.

Jiang, G. L. (2013). Molecular markers and marker-assisted breeding in plants. **In:** *Plant*

Breeding from Laboratories to Fields, pp. 46-83, (Jiang, G. L., ed.). BlackWell Publisher, London, UK.

Joshi, S. P., Gupta, V. S., Aggarwal, R. K., Ranjekar, P. K. and Brar, D. S. (2000). Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theor. Appl. Genet*. **100:** 1311-1320.

Karp, A., Seberg, O. and Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. *Annal Bot*. **78:**143-149.

Kassahun Tesfaye, Govers, K., Endashaw Bekele, and Borsch, T. (2014). ISSR fingerprinting of *Coffea arabica* throughout Ethiopia reveals high variability in wild populations and distinguishes them from landraces. *Plant Syst. Evol*. **300:** 881-897.

Kawar, P. G., Devarumath, R. M. and Nerkar, Y. (2009). Use of RAPD markers for assessment of genetic diversity in sugarcane cultivars. *Ind. J. Biotech.* **8:** 67-71.

Kesawat, M. S. and Das, B. K. (2009). Molecular markers: It's application in crop improvement. *J. Crop Sci. Biotech*. **12:** 169-181.

Kremer, A., Petit, R. G. and Pons, O. (1998). Measures of polymorphism within and among populations. **In:** *Molecular Tools for Screening Biodiversity, Plants, and Animals*, pp. 301-311, (Karp, A., Issac, P. G. and Ingram, D. S., eds.). Chapman and Hall, London, UK.

Lakshmanan, P., Geijskes, R. J., Aitken, K. S., Grof, C. L. P., Bonnett, G. D. and Smith, G. R. (2005). Sugarcane biotechnology: the challenges and opportunities. *In Vitro Cell Dev. Biol. Plant* **41:** 345-363.

Laurentin, H. (2009). Data analysis for molecular characterization of plant genetic resources. *Genet. Resour. Crop Evol*. **56:** 277-292.

Lima, M. L. A., Garcia, A. A. F., Oliveira, K. M., Matsuoka, S., Arizono, H., de Souza, C. L. and de Souza, A. P. (2002). Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugarcane (*Saccharum* spp.). *Theor. Appl. Genet*. **104:** 30-38.

Lu, Y. H., D'Hont, A, Walker, D. I. T, Rao, P. S., Feldmann, P. and Glaszmann, J. C. (1994b). Relationship among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. *Euphytica* **78:** 7-18.

Martin, C., Galbe, M., Wahlbom, F. C., Hahn-Hägerdal, B. and Jonsson, L. J. (2002). Ethanol production from enzymatic hydrolyzates of sugarcane bagasse using recombinant xylose-utilising *Saccharomyces cerevisiae*. *Enzyme Microb. Technol*. **31:** 274-282.

Miller, J. D., Gilbert, R. A. and Odero, D. C. (2009). Sugarcane botany: a brief view. **In:** *Sugarcane Handbook*, pp. 1-5, (Rice, R. W., ed.). Agronomy Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, USA.

Ming, R., Moore, P. H., Wu, K. K., D'Hont, A., Glaszmann, J. C., Tew, T. L., Mirkov, T. E., da Silva, J., Jifon, M., Rai, R. J., Schnell, S. M., Brumbley, P., Lakshmanan, J. C., Comstock, J. C. and Paterson, A. H. (2006). Sugarcane improvement through breeding and biotechnology. **In:** *Plant Breeding Reviews*, pp. 15-118, (Janick, J., ed.). John Willy and Sons. Inc., New York.

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

Mondini, L., Noorani, A. and Pagnotta, M. A. (2009). Assessing plant genetic diversity by

molecular tools. *Diversity* **1:** 19-35.

Moore, P. H., Paterson, A. H. and Tew, T. (2014). Sugarcane: the crop, the plant, and domestication. **In:** *Sugarcane: Physiology, Biochemistry, and Functional Biology*, pp.1-18, (Moore, P. H. and Botha, F. C., eds.). John Wiley and Sons, Inc., New York.

Morell, M. K., Peakall, R., Appels, R., Preston, L. R. and L1oyd, H. L. (1995). DNA profiling techniques for plant variety identification. *Aust. J. Exper. Agr.* **35:** 807-819.

Moreno, S., Martin, J. P. and Ortiz, J. M. (1998). Inter simple sequence repeats PCR for characterization of closely related grapevine germplasm. *Euphytica* **101:** 117-125.

Muench, D. G., Slinkard, A. E. and Scales, G. J. (1991). Determination of genetic variation and taxonomy in lentils (*Lens miller*) species by chloroplast DNA polymorphism. *Euphytica* **56:** 213-218.

Nair, N. V., Nair, S., Sreenivasaan, T. V. and Mohan, M. (1999). Analysis of genetic diversity and phylogeny in *Saccharum* and related genera using RAPD markers. *Genet. Res. Crop Evol.* **46:** 73-79.

Nair, N. V., Selvi, A., Sreenivasan, T. V. and Pushpalatha, K. N. (2002). Molecular diversity in Indian sugarcane cultivars as revealed by Randomly Amplified DNA Polymorphisms. *Euphytica* **127:** 219-225.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Nat. Acad. Sci.* **70:** 3321-3323.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Colombia University Press, New York, pp. 429.

Paglia, G. and Morgante, M. (1998). PCR-based multiplex DNA fingerprinting techniques for the analysis of conifer genomes. *Mol. Breed.* **4:** 173-777.

Pandey, A., Mishra, K. R., Mishra, S., Singh, P .Y. and Pathak, S. (2011). Assessment of genetic diversity among sugarcane cultivars (*Saccharum officinarum* L.) using Simple Sequence Repeat Markers. *J. Biol. Sci.* **11:** 105-111.

Pan, Y. B., Burner, D. M., Ehrlich, K. C., Grisham, M. P. and Wei, Q. (1997). Analysis of primer-derived non-specific amplification products in RAPD-PCR. *Biotechniques* **22:** 1071-1077.

Parker, P. G., Snow, A. A, Schug, M. D., Booton, G. C. and Fuerst, P. A. (1998). What molecules can tell us about populations: choosing and using a molecular marker. *Ecol.* **79:** 361-382.

Pavlicek, A., Hrda, S. and Flegr, J. (1999). Free tree free ware program for construction of phylogenetic trees on the basis of distance data and bootstrap/Jack Knife analysis of the tree robustness. Application in the RAPD analysis of genus *Frenkelia*. *Folia Biol.* **45:** 97-99.

Phillips, S. (1995). Poaceae. **In:** *Flora of Ethiopia and Eritrea* vol. 7, pp. 291-293, (Hedberg, I. and Edwards, S., eds.). Addis Ababa, Ethiopia and Uppsala, Sweden.

Powell, W., Machray, G. C. and Provan, J. (1996). Polymorphism revealed by simple sequence repeat (SSR) markers. *Sci*. **1**: 215-222.

Premachandran, M. N., Prathima, P. T. and Lekshmi, M. (2011). Sugarcane and polyploidy. *J. Sugarcane Res*. **1:** 1-15.

Que, Y., Pan, Y., Lu, Y., Yang, C., Yang, Y., Huang, N. and Xu, L. (2014). Genetic analysis of diversity within a Chinese local sugarcane germplasm based on start codon targeted (SCoT) polymorphism. *BioMed. Res. Int*. **2014:** 1-10.

Rao, G. P. (2007). *Economic Botany: Potato, Sugarcane, and Vegetable oils*. Sugarcane Research Station, Kunraghat, India, pp. 53.

Raoch, B. T. (1989). Origin and improvement of the genetic base of sugarcane. *Proc. Aust. Soci. Sugarcane Technol.* **1:** 34-47.

Reddy, M. P., Sarla, N. and Siddiq, E. A. (2002). Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* **128:** 9-17.

Rodriguez, A. M. H., Marco, A., Castillo, C., Ericka, P. and Flores, B. (2005). Genetic diversity of the most important sugarcane cultivars in Mexico. *Gnosis* **3:** 1-10.

Rohlf, F. J. (2000). NTSYS-pc, Numerical Taxonomy and Multivariate Analysis System, Version 2.11, Exeter software, New York.

Rosa, R. D., James, C. M. and Tobutt, K. R. (2004). Using microsatellites for paternity testing in olive progenies. *Hort. Sci*. **39:** 351-354.

Russell, J. R., Fuller, J. D., Macaulay, M., Hatz, B. G., Jahoor, A., Powell, W. and Waugh, R. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs, and RAPDs. *Theor. Appl. Genet*. **95:** 714-722.

Saal, B. and Wricke, G. (2002). Clustering of amplified fragment length polymorphism markers in a linkage map of rye. *Plant Breed*. **121:** 117-123.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N. (1985). Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. *Sci.* **230:** 1350-1354.

Saitou, N. and Nei, M. (1987). The neighbor joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*. **4:** 406-425.

Schaal, B., Hayworth, D., Olsen, K., Rauscher, J. and Smith, W. (1998). Phylogeographic

studies in plants: problems and prospects. *Mol. Ecol.* **7:** 465- 474.

Schlotterer, C. and Tautz, D. (1992). Slippage Synthesis of Simple Sequence DNA. *Nucleic Acids Res*. **20:** 211-215.

Schueler, S., Tusch, A. and Scholz, F. (2006). Comparative analysis of the within-population genetic structure in wild cherry (*Prunus avium* L.) at the self-incompatibility locus and nuclear microsatellites. *Mol. Ecol*. **15:** 3231-3243.

Selvi, A., Nair, N. V., Balasundaram, N. and Mohapatra, T. (2003). Evaluation of maize microsatellite markers for genetic diversity analysis and fingerprinting in sugarcane. *Genome* **46:** 394-403.

Semagn, K., Bjornstad, A. and Ndjiondjop, M. N. (2006). An overview of molecular marker methods for plants. *Afr. J. Biotech*. **5:** 2540-2568.

Siles, B. A, O'Neil, K. A, Fox, M. A, Anderson, D. E., Kuntz, A. F., Ranganath, S. C. and Morris, A. C. (2000). Genetic fingerprinting of grape plant (*Vitis vinifera*) using random amplified polymorphic DNA (RAPD) analysis and dynamic size-sieving capillary electrophoresis. *J. Agr. Food Chem.* **48:** 5903-5912.

Singh, R. K., Mishra, S. K., Singh, S. P., Mishra, N. and Sharma, M. L. (2010). Evaluation of microsatellite markers for genetic diversity analysis among sugarcane species and commercial hybrids. *Aust. J. Crop Sci*. **4:** 116-125.

Singh, R. K., Singh, R. B., Singh, S. P. and Sharma, M. L. (2011). Identification of sugarcane microsatellites associated to sugar content in sugarcane and transferability to other cereal genomes. *Euphytica* **182:** 335-354.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review Ecol. Syst.* **16:** 393-430.

Smiullah, Khan, F. A., Afza, A., Abdullah, Ijaz, A. and Ijaz, U. (2013). Diversity analysis of sugarcane genotypes by microsatellite (SSR) markers. *Int. J. Biotech. Mol. Biol. Res*. **4:** 105-110.

Solomon, S. (2011). Sugarcane by-products based industries in India. *J. Sugar Technol*. **13:** 408- 416.

Souza, E. and Sorrells, M. E. (1991). Relationships among 70 North American oat germplasms cluster analysis using qualitative characters. *Crop Sci*. **31:** 605-612.

Srivastava, S. and Gupta, O. S. (2008). Inter Simple Sequence Repeat profile as a genetic marker system in sugarcane. *Sugar Technol*. **10:** 48-52.

Tiliye Feyissa, Nybom H., Bartish, I. V. and Welander, M. (2007). Analysis of genetic diversity in the endangered tropical tree species, *Hagenia abyssinica* using ISSR markers. *Genet. Res. Crop Evol.* **54:** 947-958.

Ullah, S. S., Hossain, A., Hossain, M., Barman, S., Sohag, M. H. and Prodhan, S. H. (2013). Genetic diversity analysis of chewing sugarcane (*Saccharum officinarum* L.) varieties by using RAPD markers. *J. BioSci. Biotech.* **2:** 145-150.

Varshney, R. K., Marcel, T. C., Ramsay, L., Russell, J., Roder, M. S., Stein, N., Waugh, R., Langridge, P., Niks, R. E. and Graner, A. (2007). A high density barley microsatellite consensus map with 775 SSR loci. *Theor. Appl. Genet*. **114:** 1091-1103.

Verheye, W. H. (2010). Growth and production of sugarcane. **In:** *Soils, Plant Growth, and Crop Production*, pp. 1-10, (Verheye, W. H., ed.). EOLSS Publishers Company Limited, Belgium.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., de Lee, T. V., Hornes, M., Frijters, A., Pot, J.,

Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. **23:** 4407-4414.

Weber, J. L. and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet*. **44:** 388-396.

Weir, B. S. (1996). *Methods for Discrete Population Genetic Data*. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts, pp. 290.

Weising, K., Nybom, H., Wolff, K. and Meyer, W. (1995). *DNA Fingerprinting in Plants and Fungi*. CRC Press, New York, pp. 315.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). *DNA Fingerprinting in Plants: Principles, Methods, and Applications* 2nd edn. CRC Press, New York, pp. 444.

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

Winter, P. and Kahl, G. (1995). Molecular markers technologies for plant improvement. *World J. Micro. Biotech*. **11:** 438-448.

Wolfe, A. D. and Liston, A. (1998). Contributions of PCR-based methods to plant systematics and evolutionary biology. **In:** *Plant Molecular Systematics* II, pp. 125-130, (Soltis, D. E., Soltis, P. S. and Doyle, J. J., eds.). Chapman Hall, New York.

Wolfe, A. D., Xiang, Q. Y. and Kephart, S. R. (1998). Assessing hybridization in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable inter simple sequence repeat (ISSR) markers. *Mol. Ecol*. **7:** 1107-1125.

Xu, Y. (2010). *Molecular Plant Breeding*. CABI Publisher, London, UK, pp. 734.

Yeh, F. C., Yang, R. and Boyle, T. (1999). Microsoft Window-based Freeware for Population Genetic Analysis. University of Alberta, Canada.

You, Q., Xu, L., Zheng, Y. and Que, Y. (2013). Genetic diversity analysis of sugarcane parents in Chinese breeding programs using gSSR markers. *Scient. World J.* **2013:** 1-11.

Zhou, M. (2013). Conventional sugarcane breeding in South Africa: progress and future prospects. *Am. J. Plant Sci.* **4:** 189-196.

Zietkiewicz, E., Rafalski, A. and Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. *Genomics* **20:** 176-183.

10. APPENDICES

Appendix 1. Lists of sugarcane genotypes studied with their origin and nature of utilization/current status

No.	Genotype code	Locality/Origin	Nature of genotype
1	B52/298	Barbados	Commercial
2	B4906	Barbados	Candidate
3	B59212	Barbados	Old
4	B863/349	Barbados	Old
5	B60/263	Barbados	Old
6	B52/219	Barbados	Old
7	B45/54	Barbados	Old
8	B5736	Barbados	Candidate
9	B5490	Barbados	Candidate
10	B58230	Barbados	Candidate
11	B59/04	Barbados	Candidate
12	B60/63	Barbados	Candidate
13	B50210	Barbados	Candidate
14	C92/26	Cuba	New
15	C323/68	Cuba	New
16	C98/128	Cuba	New
17	C85/102	Cuba	New
18	C132/81	Cuba	New
19	C90/530	Cuba	New

20	C88/556	Cuba	New
21	C92/514	Cuba	New
22	C89/147	Cuba	New
23	C87/51	Cuba	New
24	C88356	Cuba	New
25	FG03/291	France	New
26	FG04/798	France	New
27	FG04/420	France	New
28	FG06/806	France	New
29	PSR07/84	France	New
30	Vmc96/22	France	New
31	FG04/641	France	New
32	FG04/896	France	New
33	Vmc95/252	France	New
34	FG06/750	France	New
35	FG04/275	France	New
36	PSR97/105	France	New
37	Vmc96/47	France	New
38	FG04/622	France	New
39	FG06/747	France	New
40	Vmc9660	France	New
41	FG04/607	France	New
42	Vmc96/62	France	New

43	FG06/755	France	New
44	Vmc96/55	France	New
45	FG05/424	France	New
46	Vmc95/243	France	New
47	Co680	India	Commercial
48	Co1148	India	Candidate
49	Co678	India	Candidate
50	Co434	India	Old
51	Co0238	India	New
52	Co421	India	Commercial
53	Co449	India	Commercial
54	Co810	India	Candidate
55	Co740	India	Commercial
56	N52/219	South Africa	Commercial
57	N14	South Africa	Commercial
58	N55/805	South Africa	Old
59	NCo334	South Africa	Commercial
60	NCo376	South Africa	Candidate
61	N53/216	South Africa	Commercial
62	N11	South Africa	Old
63	NCo310	South Africa	Old
64	2/333	Sudan	New
65	2/111	Sudan	New

66	2/999	Sudan	New
67	Q217	Sudan	New
68	2/888	Sudan	New
69	2/555	Sudan	New
70	Q200	Sudan	New
71	2/777	Sudan	New
72	2/222	Sudan	New
73	2/444	Sudan	New
74	CP00/1527	USA	New
75	CP97/1944	USA	New
76	CP97/1989	USA	New
77	CP47/193	USA	New
78	CP00/1301	USA	New
79	CP61/39	USA	Candidate
80	CP00/1302	USA	New
81	CP60/23	USA	Candidate
82	CP04/1566	USA	New

Key:

Candidate = refers to breeding strains, lines of hybrids that a breeder or institution are announced for official national performance testing in order to get released as a certified.

New = newly released improved varieties from breeding institute

Commercial = any genotype under cultivation

Old = any genotype out of production

Appendix 2. Concentration and quality of DNA of all sugarcane genotypes investigated

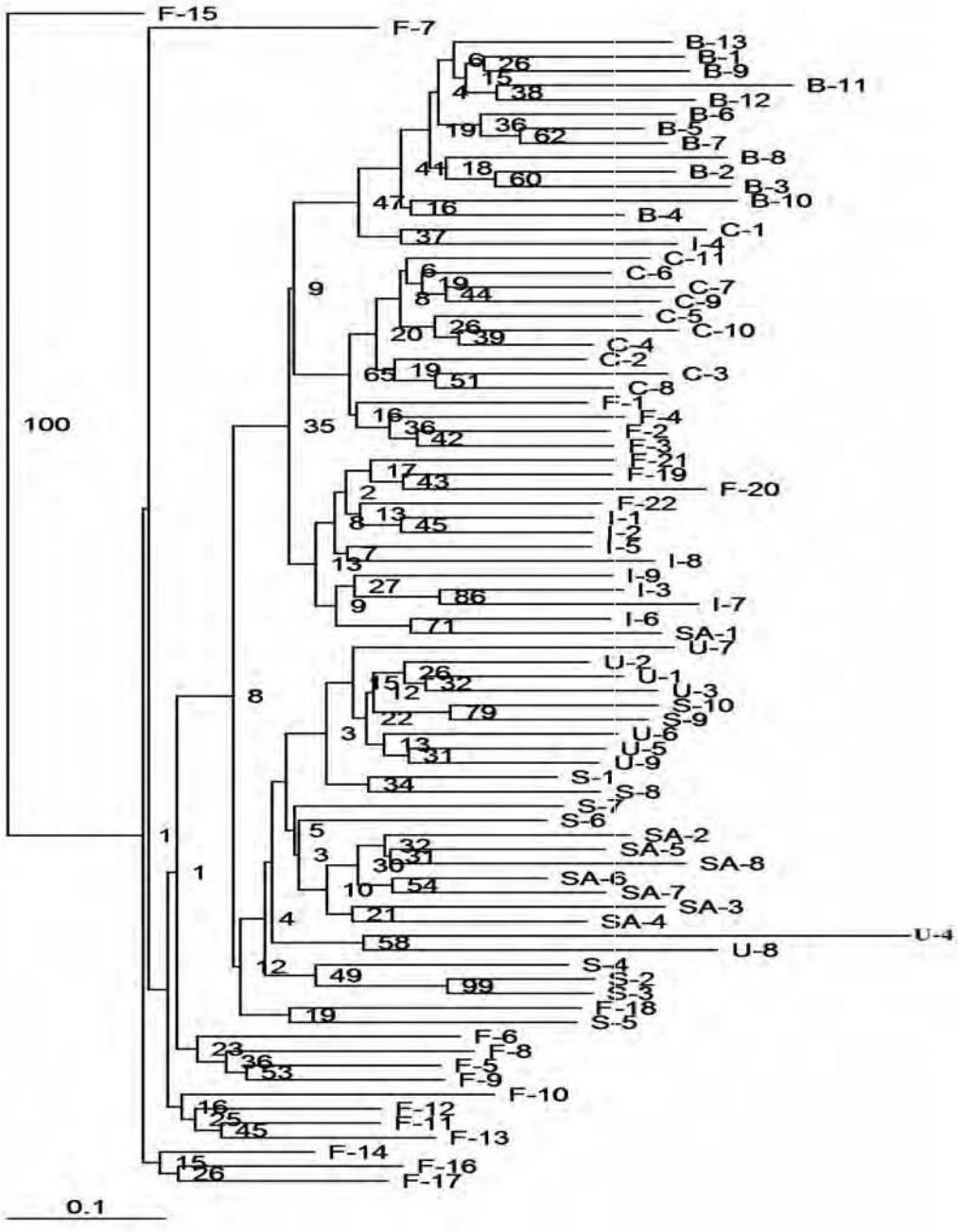
No.	Genotype code	Concentration	Unit	A260	A280	260/280	260/230
1	B52/298	352.3	ng/µl	7.046	3.453	2.04	2.25
2	B4906	219.1	ng/µl	4.382	2.056	2.13	2.47
3	B59212	251.3	ng/µl	5.025	2.396	2.1	2.35
4	B863/349	257.8	ng/µl	5.157	2.459	2.1	2.39
5	B60/263	414.1	ng/µl	8.281	4.592	1.8	1.08
6	B52/219	268.2	ng/µl	5.363	2.646	2.03	2.24
7	B45/54	235.6	ng/µl	4.712	2.3	2.05	2.28
8	B5736	463.7	ng/µl	9.273	4.42	2.1	2.34
9	B5490	1087.3	ng/µl	21.746	10.681	2.04	2.29
10	B58230	139.8	ng/µl	2.797	1.323	2.11	2.41
11	B59/04	203.3	ng/µl	4.067	1.996	2.04	1.96
12	B60/63	68.7	ng/µl	1.373	0.671	2.05	1.73
13	B50210	353.4	ng/µl	7.069	3.394	2.08	2.28
14	C92/26	448.9	ng/µl	8.977	4.414	2.03	2.23
15	C323/68	165.8	ng/µl	3.316	1.606	2.06	2.28
16	C98/128	91	ng/µl	1.82	0.877	2.08	2.51
17	C85/102	884.4	ng/µl	17.688	8.649	2.05	2.39
18	C132/81	204.4	ng/µl	4.088	1.979	2.07	2.29
19	C90/530	569	ng/µl	11.38	5.522	2.06	2.32
20	C88/556	726.7	ng/µl	14.534	7.113	2.04	2.26

21	C92/514	606.8	ng/µl	12.136	5.991	2.03	2.14
22	C89/147	728.4	ng/µl	14.568	7.191	2.03	2.26
23	C87/51	327.6	ng/µl	6.551	3.705	1.77	1.57
24	C88356	663.3	ng/µl	13.266	6.548	2.03	2.09
25	FG03/291	540.5	ng/µl	10.811	5.354	2.02	2.24
26	FG04/798	714.5	ng/µl	14.289	7.138	2	2.27
27	FG04/420	279.6	ng/µl	5.592	2.813	1.99	2.09
28	FG06/806	255.9	ng/µl	5.117	2.653	1.93	1.74
29	PSR07/84	578.3	ng/µl	11.566	6.085	1.9	2.12
30	Vmc96/22	777.3	ng/µl	15.546	7.6	2.05	2.17
31	FG04/641	1554.2	ng/µl	31.085	15.599	1.99	2.2
32	FG04/896	597.6	ng/µl	11.953	6.003	1.99	2.23
33	Vmc95/252	971.1	ng/µl	19.421	9.657	2.01	2.23
34	FG06/750	1103.4	ng/µl	22.067	10.693	2.06	2.32
35	FG04/275	867.6	ng/µl	17.352	9.081	1.91	2.13
36	PSR97/105	53.7	ng/µl	1.073	0.544	1.97	2.11
37	Vmc96/47	747.3	ng/µl	14.945	7.571	1.97	2.07
38	FG04/622	961.6	ng/µl	19.231	9.506	2.02	2.24
39	FG06/747	744.2	ng/µl	14.884	7.492	1.99	2.27
40	Vmc9660	585	ng/µl	11.7	6.507	1.8	1.3
41	FG04/607	1147.6	ng/µl	22.951	11.173	2.05	2.31
42	Vmc96/62	996.3	ng/µl	19.926	9.643	2.07	2.25
43	FG06/755	1153.8	ng/µl	23.076	11.379	2.03	2.27

44	Vmc96/55	931.2	ng/µl	18.624	9.334	2	2.14
45	FG05/424	1051	ng/µl	21.019	10.406	2.02	2.26
46	Vmc95/243	788.3	ng/µl	15.766	7.766	2.03	2.43
47	Co680	1301.3	ng/µl	26.026	13.234	1.97	2.16
48	Co1148	1317.4	ng/µl	26.347	13.063	2.02	2.23
49	Co678	1294.2	ng/µl	25.885	13.295	1.95	2.22
50	Co434	914.7	ng/µl	18.295	9.367	1.95	2.16
51	Co0238	974.1	ng/µl	19.482	9.648	2.02	2.22
52	Co421	1274.8	ng/µl	25.497	12.739	2	2.19
53	Co449	1299.9	ng/µl	25.998	12.877	2.02	2.24
54	Co810	1056.9	ng/µl	21.138	10.506	2.01	2.25
55	Co740	1077.3	ng/µl	21.546	10.638	2.03	2.22
56	N52/219	1960.7	ng/µl	39.213	18.8	2.09	2.47
57	N14	762.3	ng/µl	15.246	7.653	1.99	2.01
58	N55/805	834.6	ng/µl	16.693	8.479	1.97	2.11
59	NCo334	1009.8	ng/µl	20.196	10.387	1.94	2.19
60	NCo376	1321.5	ng/µl	26.43	13.558	1.95	2.15
61	N53/216	912	ng/µl	18.239	9.843	1.85	1.81
62	N11	955.9	ng/µl	19.117	9.904	1.93	2.2
63	NCo310	1136.2	ng/µl	22.723	11.754	1.93	2.11
64	2/333	599.1	ng/µl	11.982	6.061	1.98	2.2
65	2/111	1236.9	ng/µl	24.739	12.313	2.01	2.28
66	2/999	1173.7	ng/µl	23.475	11.935	1.97	2.23

67	Q217	963.3	ng/µl	19.266	9.749	1.98	2.21
68	2/888	1053.1	ng/µl	21.061	10.724	1.96	2.26
69	2/555	837.7	ng/µl	16.755	8.595	1.95	2.23
70	Q200	581.6	ng/µl	11.632	5.933	1.96	2.25
71	2/777	1284.8	ng/µl	25.695	13.085	1.96	2.05
72	2/222	1143.5	ng/µl	22.869	11.48	1.99	2.21
73	2/444	938.6	ng/µl	18.771	9.536	1.97	2.23
74	CP00/1527	853	ng/µl	17.061	8.734	1.95	2.24
75	CP97/1944	1255.1	ng/µl	25.101	12.71	1.97	2.26
76	CP97/1989	1197.4	ng/µl	23.947	12.389	1.93	2.04
77	CP47/193	1463.8	ng/µl	29.276	15.288	1.92	2.06
78	CP00/1301	886.1	ng/µl	17.721	8.972	1.98	2.29
79	CP61/39	1176.1	ng/µl	23.523	11.906	1.98	2.2
80	CP00/1302	589.3	ng/µl	11.787	5.907	2	2.71
81	CP60/23	918.8	ng/µl	18.377	9.514	1.93	2.2
82	CP04/1566	954.6	ng/µl	19.092	9.7	1.97	2.23

Appendix 3. Bootstrapped NJ dendrogram obtained from Jaccard pairwise similarity



Declaration

I the undersigned, declare here by that this thesis is my original work and has not been presented for a degree in any other University. All source of the material used for this thesis have been properly acknowledged.

Name	Signature	Date
Adebabay Tazeb	_____	_____

We here by certify that this thesis has been done under the guidance of us and we read and evaluated the thesis entitled as *Molecular Characterization of Introduced Sugarcane (Saccharum officinarum L.) Genotypes In Ethiopia Using Inter Simple Sequence Repeat (ISSR) Markers.* We recommend that it can be submitted as fulfilling thesis requirement.

Advisors:

	Signature	Date
Dr. Teklehaimanot Haileselassie	_____	_____

	Signature	Date
Dr. Kassahun Tesfaye	_____	_____