Genetic Diversity and Population Structure Analysis of Vernonia [*Vernonia galamensis* (Cass.) Less L.] Germplasm in Ethiopia using ISSR Markers

GETE ESHETU

A Thesis Submitted to

The department of Applied Biology

School of Applied Natural Science

Presented in Partial Fulfillment of the Requirement for the Degree of Master's in Applied Biology (Specialization in Biotechnology)

Office of Graduate Studies

Adama Science and Technology University

October, 2021

Adama, Ethiopia

Genetic Diversity and Population Structure Analysis of Vernonia [*Vernonia galamensis* (Cass.) Less L.] Germplasm in Ethiopia using ISSR Markers

GETE ESHETU

Advisor: Alemneh Mideksa (PhD)

Co-Advisors: Kassahun Tesfaye (PhD)

Tilahun Mekonnen (PhD)

A Thesis Submitted to Department of Applied Biology

School of Applied Natural Science

Presented in Partial Fulfillment of the Requirement for the Degree of Master's in Applied Biology (Specialization in Biotechnology)

Office of Graduate Studies

Adama Science and Technology University

October, 2021

Adama, Ethiopia

Declaration

I hereby declare that this Master Thesis entitled “Genetic Diversity and Population Structure Analysis of Vernonia [*Vernonia galamensis* (Cass.) Less L.] Germplasm in Ethiopia using ISSR Markers” is my original work. It has not been submitted for the award of any academic degree, diploma or certificate in any other university. All sources of materials that are used for this thesis have been duly acknowledged through citation.

Name…… Signature…… Date……

Recommendation

we, the advisor(s) of this thesis, hereby certify that I/we have read the revised version of the thesis entitled "Genetic Diversity and Population Structure Analysis of Vernonia [*Vernonia galamensis* (Cass.) Less L.] Germplasm in Ethiopia using ISSR Markers'', prepared under our guidance by Gete Eshetu submitted in partial fulfillment of the requirements for the degree of Mater's of Science in Biotechnology. Therefore, we recommend the submission of revised version of the thesis to the department following the applicable procedures.

Major Advisor	Signature	Date
Co- advisor	Signature	Date
Co-advisor	Signature	Date

Approval of Board of Examiners

I/we, the advisors of the thesis entitled "Genetic Diversity and Population Structure Analysis of Vernonia *[Vernonia galamensis* (Cass.) Less L.] Germplasm in Ethiopia using ISSR Markers" developed by Gete Eshetu hereby certify that the recommendation and suggestions made by the board of examiners are appropriately incorporated into the final version of the thesis.

Major Advisor	Signature	Date
Co-advisor	Signature	Date

We, the undersigned, members of the Board of Examiners of the thesis by **Gete Eshetu** have read and evaluated the thesis entitled "Genetic Diversity and Population Structure Analysis of Vernonia [*Vernonia galamensis* (Cass.) Less L.] Germplasm in Ethiopia using ISSR Markers "and examined the candidate during open defense. This is, therefore, to certify that the thesis is accepted for partial fulfillment of the requirement of the degree of Master of Science in Applied Biology (Biotechnology).

Chairperson	Signature	Date
Internal Examiner	Signature	Date
External Examiner	Signature	Date

Final approval and acceptance of the thesis is contingent upon submission of its final copy to the Office of Postgraduate Studies (OPGS) through the Department Graduate Council (DGC) and School Graduate Committee (SGC).

Department Head	Signature	Date
School Dean	Signature	Date
Office of Postgraduate Studies, Dean	Signature	Date

Acknowledgement

First of all, I would like to thank the Almighty GOD for His works of mercy in my life, may his holy name be worshipped forever and ever.

I would like to thank my research main advisor, Dr. Alemneh Mideksa, for his skillful guidance, encouragement up to the accomplishment of this thesis. I would like to express my sincere thanks to my co-advisor, Dr. Kassahun Tesfaye, Director General of Ethiopian Biotechnology Institute (EBTi) for allowing me to work in his Plant Genetics Research Laboratory at College of Natural and Computational Sciences, Addis Ababa University, Ethiopia.

My special appreciation goes to my co-advisor Dr. Tilahun Mekonnen for his constructive guidance, endless and technical support during the laboratory work, interpretation of the results and review of the thesis. I would like to thank Adama Science and Technology University for their support by facilitating condition for continuity of my post graduate study. I would also like to thank Addis Ababa university, Institute of Biotechnology for providing the molecular laboratory facilities to conduct this research.

I am also deeply grateful to Mrs. Wubishet Mamo for his thoughtfulness, encouragement, advice and support during laboratory work, data analysis and for cooperation in every aspect. So many people have made meaningful contribution to achieve my goal. I am thankful to all friends who I have met along the way even if your names do not appearhere.

Last but not least, I would like to pass my heartfelt thanks to my family without their support, encouragement and unconditional love, my ambition would not have been realized.

TABLE OF CONTENTS

Contents **Pages**

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX

LIST OF ABBREVIATIONS AND ACRONYMS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
CTAB	Cetyl-trimethyl Ammonium Bromide
DNA	Deoxyribonucleic acid
DNTPS	Deoxyribonucleotide triphosphate
Fst	Fixation index
I	Shannon information index
ISSR	Inter simple sequence repeat
Na	Number of allele
Ne	Number of effective allele
NJ	Neighbor joining
PCoA	Principal coordinates analysis
PCR	Polymerase chain reaction
PPL	Percent of polymorphic loci
PVC	Polyvinylchloride
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
SNNPR	South nation, nationality and people's region
SSR	Simple sequence repeat
UPGMA	Unweighted pair group with arithmetic mean

ABSTRACT

Vernonia galamensis L. is a wild plant from the family Asteraceae which is endemic to East Africa. This crop has the potential to become a new oil crop for industrial uses with very high content of natural epoxidized oil, which can be used for manufacturing of polyvinyl chloride, petrochemicals, cosmetic and pharmaceutical products. Knowledge of the genetic structure of the crop is important for its improvements, conservation and wise-utilization. Therefore, the present study was designed to investigate the genetic diversity and population structure of 60 V. galamensis L. germplasm in Ethiopia representing eight populations using eight inter simple sequence repeat markers. Awide range of diversity indices including number of alleles, effective number of alleles, Shannon's diversity index, percentage of polymorphic loci and polymorphic information content, were computed to determine the extents of genetic variation within and among populations. Analysis of molecular variance was performed to see genetic variation within and among the populations. All the markers used in the present study were polymorphic and highly informative. Polymorphic information content ranged from 0.5 to 0.71 with overall mean of 0.51. A high within-populations genetic diversity was confirmed with gene diversity values ranging from 0.28 to 0.43 with an overall mean of 0.36. Analysis of molecular variance revealed that most (65%) of the total genetic variation was accounted for within populations genetic variation and 35% was among population variation. Clustering and Principal co-ordinate analyses did not sharply group the materials following their geographical areas of sampling. Population structure analysis revealed the presence of two sub-populations with lower degree of genetic admixture. In this study Inter simple sequence repeat markers were observed to be an appropriate molecular marker to analysis genetic diversity and population structures of Vernonia galamensis in Ethiopia. Among the studied populations, East Shoa and East Hararghe populations showed relatively higher gene diversity indicating that these sites could be targeted for breeding and conservation of Vernonia galamensis. Therefore, the present study had successfully disclosed the genetic diversity and population structure of Vernonia galamensis germplasms in Ethiopia.

Keywords: Clustering, Conservation, Genetic diversity, ISSR marker, Polymorphism, *Vernonia galamensis*

1.INTRODUCTION

Vernonia (*Vernonia galamensis* (Cass.) Less.) (2n = 18) is the most important wild oil crop that belongs to Asteraceae (Compositae), a diverse group of angiosperms (Rahman *et al.,* 2008). This species, often called ironweed and its genus has two major centers of origin, South America and tropical Africa (Kedir, 2019). Most of the *Vernonia* species occur in South America although more than 300 species are occurred in Africa countries such as Ethiopia, Madagascar, Cape Verde, Eritrea, Mozambique, Kenya and northern Tanzania (Perdue *et al.*, 1986). The highest diversity of *V. galamensis* is mostly occurring in Ethiopia and Kenya (Baye & Becker, 2005).

In Eastern and South Eastern Ethiopia, the plant is found in farmlands, in the forests, and in the compounds of mosques and churches. The species is diversified in areas with well-drained soil, well defined seasons with relatively warm climate, and it occurs in areas with annual rainfall as low as 250 mm (Kedir 2009). Nowadays, *V. galamensis* is grown in many parts especially in eastern, southern and southeastern parts of Ethiopia.

Economic importance of *V. galamensis* comprises manufacturing of polyvinylchloride, paints, adhesives, insecticides, and structural polymers for the production of plastic materials and petrochemicals (Shimelis & Hugo, 2010). The seed of this crop produces useful natural epoxy fatty acids that are better than artificial epoxy oils and contain high amount of vernolic acid (Shimelis *et al.,* 2008), which has many applications in the chemical, pharmaceutical and agro-processing industries. In addition to industrial use, Vernonia oil is used as traditional medicine, food, and possessing antitumor and antiviral activities and also the leaves have been smoked in place of tobacco in Ethiopia. In Tanzania, the leaves are cooked in porridge, or drunk as a tea to treat chest pain and in Kenya the plant is used to treat stomach pain (Van der Vossen *et al*., 2007)

Studying the genetic diversity of *V. galamensis* is useful for its effective utilization, conservation, improvement and management (Idrees & Irshad, 2014; Govindaraj *et al*., 2015; Bhandari *et al*., 2017). Genetic diversity analysis could be conducted based on morphological, biochemical, molecular markers or their combination (Govindaraj *et al*., 2015). Morphological characterization that depends on phenotypic traits like colour, shape, size etc is considered to be less reliable as the traits are highly influenced by environmental factors or management

practices. Similarly, the lower level of polymorphism, technical complexity, developmental stage and tissue specificity made biochemical or protein markers less applicable for genetic diversity analysis (Chittaranjan & Albert, 2008).

Nowadays, DNA markers are the method of choice for genetic analysis as they are highly polymorphic, stable and detectable in all plant tissues regardless of the growth, differentiation, development or status of cells, and not influenced by environmental factors (Bornet *et al.* 2002). Molecular marker are functional instruments for examining genetic difference and offer a well-organized means to connect genotypic and phenotypic difference (Varshney *et al.* 2005). One of the most convenient and popular methods to identify and study of intraspecific genetic polymorphism is interring simple sequence repeats (ISSR) PCR based technology. The technique is very reproducible due to the utilization of longer primers, which permit for top annealing temperatures (Saleh *et al.*, 2016). It targets the simple sequence repeats and amplify the region between two closely spaced and oppositely oriented SSR loci to scan polymorphisms among germplasm without prior knowledge of the DNA sequence.

In Ethiopia, *V. galamensis* is neglected, limited information is available, is only considered as a wild weed and colonizing disturbed areas with bare agricultural lands. Even their production and management system is restricted, their potential value for industries has been underutilized and underexploited (Kedir, 2019). Molecular diversity analysis on the crop is greatly limited. Therefore, the present study was targeted to investigate the genetic diversity and population structure of *V. galamensis* collected from selected regions of Ethiopia using ISSR markers to generate baseline information useful for its improvement, conservation and wise-uses.

1.2. Statement of the problem

Vernonia galamensis is a weedy plant that naturally grows in marginal and disturbed areas. This makes it ideal for developing countries within the tropics because it reduces possible competition for land with other food crops, creates diversification of products, serves as a break crop to prevent pest damage, as an insurance against crop failure and as a new cash crop. This species has a major potential as an industrial source of vernolic acid due to a high demand of natural epoxidized oil (Shimelis *et al.,* 2008). Because of the high oil and vernolic acid content and its relatively low shattering nature, subsp. *galamensis* var. *ethiopica* has been the focus of research (Van der Vossen & Mkamilo, 2007).

In spite of its importance, the species is underestimated and underexploited. No much attention has been given to its improvement and conservation. Furthermore, detailed information about the production system,genetic variability and genetic potential are major limitation of *V. galamensis* variety Ethiopica. So far, limited molecular diversity studies have been reported for *V. galamensis* (Ramalema *et al.,* 2010;Nwakanma *et al.,* 2018; Kedir, 2019; Alemneh *et al.*,2020) However, there were no ISSR markers report(s) on *V. galamensis* genetic diversity in Ethiopia and elsewhere in the world. This study is, therefore, aimed at assessing the genetic diversity of *V. galamensis* germplasm using ISSR markers.

1.3. Objectives

1.3.1. General objective

To assess the genetic diversity and population structure of *Vernonia galamensis* germplasm collected from selected areas of Ethiopia using inter simple sequence repeat (ISSR) markers.

1.3.2. Specific objectives

- To investigate the genetic diversity of *V. galamensis* in Ethiopia using inter simple sequence repeat (ISSR) markers
- To identify the population structure of *V. galamensis* in Ethiopia based on inter simple sequence repeat marker data.
- To evaluate the potential of ISSR technique to detect the level of diversity present in samples collected from different regions of Ethiopia.

1.3. Significance of the study

The result of this study:

- Could provide information on the general genetic diversity of *V. galamensis* germplasm which may assist in the identification and selection of genetic material for conservation and improvement
- Can indicate the possible application of inter simple sequence repeat markers to genetic diversity of *vernonia galamensis.*
- Can be used as a reference for other researchers to do similar works.

2.LITERATURE REVIEW

2.1. Vernonia: origin and distribution

Genus *Vernonia* have two major centers of origin, South America and tropical Africa (Kedir, 2019. The great diversity is found in East African countries and only one variety occurs in West Africa (Van der Vossen & Mkamilo, 2007). Those genus comprises greater than1,000 species and among this more than 300 species are described from Africa with about one-third in Madagascar and about 50 in Ethiopia. *Vernonia galamensis* variety *ethiopica* was first identified in 1964 by Perdue in eastern Ethiopia on the Harar-Jijiga road (9°14 N and 42°35 E at 1700 m above sea level) and later distributed to other African countries like Eritrea, Kenya and Tanzania (Thompson, *et al.,* 1994) Recent studies on herbarium materials in Addis Ababa, Ethiopia, and Kew, London, indicated the existence of about 40–50 species in Ethiopia and the species grown in many parts of the country especially around the city of Harar (Naliaka, 1990).

2.2. Taxonomy

According to the taxonomic revision of the species, *galamensis* is recognized to include six subspecies such as *galamensis, mutomoensis, nairobensis, afromontana, gibbosa*, and *lushotoensis*. Moreover, subspecies *galamensis* has four botanical varieties, namely variety *galamensis, petitiana, australis, ethiopica* (Ramalema *et al*., 2010). Based on its taxonomic classification *V. galamensis* variety ethiopica was used in the current study.

Taxonomic classification of *V. galamensis* can be placed as:

Kingdom: Plantae (plants)

Division: Entophyte (flowering plants)

Sub division: Angiosperm

Class: Magnoliopsida

Order: Asterales

Family: Asteraceae

Sub family: Cichorioideae

Tribe**:** Vernonieae

Subtribe**:** Vernoniinae

Genus**:** *Vernonia*

Species: V. galamensis

2.3. Morphology of *V. galamensis*

The morphological characters of the *V. galamensis* plant was described comprehensively by Perdue *et al.,* (1986), and reported that *V. galamensis* is an herbaceous, usually annual, varying from small ephemerals 20 cm tall with a single flower head which grow up to 5 m tall with many flower heads. It is potentially grown as a seed oil crop in tropical and subtropical environments with frost-free and short-day length for flower initiation and development. It is also noted that the stems branch only after the first flower head is formed and the inflorescence consists of a terminal flower head with lateral flower heads from the uppermost axils. *V. galamensis* leaves is alternate, semi-transparent, 0.6-5cm wide and 25 cm long (Kedir, 2009).

2.4. Ecology of *V. galamensis*

Vernonia galamensis is endemic to East African countries, and naturally grows as a weed in the fields or in woodlands and adapted to the semi-arid tropics of dry land. The species is diversified in areas with well-drained soil, well defined seasons and relatively warm climate (Baye, 2002; Baye and Becker, 2005). It also requires a rainy season that provides sufficient moisture to permit the main flower heads to develop (Baye *et al.,* 2001). However, it can grow under low rainfall as 250 mm and marginal conditions that is suitable for dryland farming. But, a long

rainy season permits developing secondary flower heads, resulting in non-uniformity of maturation and a risk of seed shattering (Hadebe *et al.,* 2017). In Eastern and south eastern Ethiopia, the plants are found in farmlands, in the forests, and in the compounds of mosques and churches. The plants tolerate substantial shading, high temperature and full sun as long as the soil moisture is sufficient (Shimelis *et al.,* 2008).

2.5. Economic importance of *V. galamensis*

V. galamensis has a major potential as an industrial source of vernolic acid due to a high demand of its natural epoxidized oil (Shimelis *et al.*, 2008). *V. galamensis*, has several advantages over many other oil seeds. Primarily, the plant high oil yield which is more than 40% oil content of total seed weight. Next, the plant is resistant to insect pests, disease and drought. Thirdly, *V. galamensis* grows in areas with low annual rainfall as little as 200mm of seasonal rainfall which makes it suitable cash crop for farmers of arid and semi-arid areas (Baye & Sileshi2002).

The industrial use of *V. galamensis* seed oil (vernolic acid) is important to reduce emissions of volatile organic compounds, and thus is environmentally friendly, it is also less expensive (Bhardwaj *et al.,* 2000). In contrast, chemically synthesized epoxidized soybean and linseed oil are highly viscous, expensive (need high epoxidation process) and contains volatile organic compounds (VOC) with high emission to the environment during processing and use. In addition, artificial epoxidized oil is semi-solid at 10°C and may not be poured at 0°C while vernolic acid are often poured even below the melting point (Mebrahtu *et al.,* 2009; Shimelis & Gwata, 2013). The low viscosity of vernonia oil makes it a good solvent in paint manufacture, while its highly reactive epoxy group makes it to chemically bind in the dried paint rather than evaporating to the atmosphere (Shimelis & Hugo, 2011).

A recent work by Getachew Ali (2008) on Ethiopian varieties of *V. galamensis* shows average oil content of 35.21% with high variability among various germplasm. The presence of epoxy group, the low viscosity and polymerizing characteristics of this oil makes it valuable as a solvent in industrial coatings and paints. The product of this plant (Vernonia oil) could also be used as a natural source of plasticizers and stabilizers for producing

polyvinylchloride (PVC) plastic, which is currently manufactured from petroleum. Similarly uses as lubricants, adhesives, protective coatings, cosmetics, detergents, a raw product for nylons and much more (Jaworski & Cahoon, 2003).

2.6. Plant genetic diversity

Genetic diversity can be expressed as any variation in the nucleotides, genes, chromosomes, or genomes of a species at the level of individual, population, species and sub-species for a given time (Fu, 2015). Genetic diversity reflects the presence of various alleles within the gene pool, and hence, different genotypes within populations. Genetic diversity of crops plays an important role in sustainable development and food security in such a way that highly diversified species have probability to survive in any environmental condition (Esquinas & Alcazar 2005). Knowledge of the genetic diversity and population structure of germplasm collections is an important foundation for crop improvement and a key component of effective conservation and breeding strategies (Thomson *et al.*, 2007).

Genetic diversity can be measured by counting the number of different alleles in a gene pool. Populations with greater genetic diversity are experienced with high fertility, low mortality and able to evolve in response to a changing environment than those with low genetic diversity even in good environmental conditions (Bhandari *et al.,* 2017). The assessment of genetic diversity within and between plant populations is routinely performed on the basis of morphological, biochemical, and molecular markers (Shimelis *et al.,* 2008; Govindaraj *et al.,* 2015).

2.7. Markers and their application in genetic diversity assessment

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 & 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, and production of improved varieties (Henry, 2001).

There are three major types of genetic markers: morphological markers (also called "classical" or "visible" markers) which are phenotypic traits, biochemical markers, which are called isozymes, including allelic variants of enzymes and DNA markers (or molecular markers) which reveal sites of variation in DNA .Each type of marker system has its own advantages and disadvantages and type of marker is selected based on the species types, availability and cost benefit analysis (Semagn *et al.,* 2006).

2.7.1. Genetic diversity study by using molecular markers

Molecular markers are fragments of DNA sequence that are associated to a part of the genome carrying gene responsible for a trait (Adebayo *et al*.,2015). Molecular markers have several advantages over the phenotypic markers that are challenging and time-consuming to choose by plant breeders (Kedir, 2019).

These DNA markers are not influenced by environmental conditions and are detectable at all plant growth stages. They are important to study genetic diversity of the plant and manage plant genetic resources and also used to detect tolerance or susceptibility for abiotic stresses in crop plants, (Song *et al*., 2003), The most common types of dominant DNA markers that are used for genetic diversity studies are Random amplified length polymorphism (RAPD), Inter simple sequence repeat (ISSR), Amplified fragment length polymorphism (AFLP), polymorphism (De Vicente & Fulton, 2003). Molecular markers are commonly used to characterize genetic diversity within or between populations or groups of individuals because they typically detect high levels of polymorphism (Dalirsefat *et al.*, 2009).

2.8. Non-PCR based molecular markers

2.8.1. Restricted fragment length polymorphism (RFLP)

Restricted fragment length polymorphism was hybridization based molecular marker technique. In case of these markers, the genomic DNA is digested and a series of fragments having varied length produced by restriction enzymes (Preetha & Raveendren, 2008). These markers enabled DNA variations to be tested as substitutions of a single base in the recognition sequence of a restriction enzyme that altered the length of resultant restriction fragments (Schlotterer, 2004). RFLP markers have wide range of application due to their co-dominance nature, high reproducibility and cheap methods toperform (Misra, 2009).

Conversely, RFLP requires the presence of high quantity and quality of DNA which is time consuming, laborious, uses probe that is difficult to handle and dispose (Said and Kassahun 2012), Beside this (RFLP) markers, have low level of DNA polymorphism especially in a self-pollinated crop that make many researches have turned to PCR based markers (Burow *et al.*, 2012).

2.9.PCR- based molecular markers

PCR was the major breakthrough for Molecular markers. Any genomic region could be amplified and analyzed in many individuals without the requirement for cloning and isolating large amounts of ultrapure genomic DNA (Schlotterer, 2004). PCR based approach uses only small quantities of DNA, avoid DNA blotting and use of radioactivity and are amenable to automation (Nagaraju *et al.,* 2001). This technology has led to the development of a number of simple and quick techniques such as RAPD, ISSR and AFLP. Most of the PCR based markers with other novel approaches have been extensively utilized in genetic diversity assessment (Montes *et al.,* 2009).

2.9.1. 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). RAPD markers are useful DNA based method for initial assessment of genetic variation, especially the assessment of genetic diversity in plant species (Song, 2005; Fedlu Hassen *et al.*, 2007). Since RAPD technique is simple, no requirement of sequence information for primer

design and no probe development to precede it is extensively used in genetic variability analysis of various plant and animal species (Arya *et al.,* 2011). Lower reproducibility because of low stringency condition is the main disadvantage of RAPD. Therefore, constant reaction conditions are crucial for reliable and reproducible results (Weising *et al*., 2005).

2.9.2. Amplified fragment length polymorphism (AFLP)

Amplified fragment length polymorphism is a PCR-based marker that used in genetic research and DNA finger printing. The technique uses restriction enzymes to cut genomic DNA, followed by legation of adaptors to the sticky ends of the restriction fragments (Srivastava & Misra, 2009). This marker is used to distinguish closely related individuals in sub species level (Althoff *et al.,* 2007). AFLP gives high level of resolution to allow delineation of complex genetic structures, to differentiate individuals in a population gene flow experiments and also to register plant varieties (Misra *et al.,* 2010). The advantages of AFLP are highly reliable and reproducible, it does not require any DNA sequence information from the organism under study, it is information-rich due to its ability to analyze a large number of polymorphic loci simultaneously (effective multiplex ratio) with a single primer combination on a single gel as compared to RAPDs, RFLPs and microsatellites, co-migrating AFLP amplification products are mostly homologous and locus specific with exceptions in polyploid species (Kumar *et al*., 2014).

However, limitations of AFLP are that it requires a greater number of steps to produce the result, requires template DNA free of inhibitor compounds that interferes with the restriction enzyme, the technique requires the use of polyacrylamide gel in combination with $AgNO_3$ staining, radioactivity, or fluorescent methods of detection, which will be more expensive and laborious than agarose gels, involves additional cost to purchase both restriction and ligation enzymes as well as adapters. Most AFLP loci are dominant, which does not differentiate dominant homozygotes from heterozygotes (Misra *et al*., 2014). This reduces the accuracy of AFLP markers in population genetic analysis, genetic mapping, and marker assisted selection (MAS) (Govindaraj *et al*., 2015).

2.9.3. Simple sequence repeat (SSR)

SSR is one of the most powerful molecular marker and characterized by a high degree of repetition and occurs at many thousand loci in the nuclear genome (Singh *et al.,* 2010). SSR polymorphism reflects differences in simple repetitive sequences of defined regions of the genome and the repeated sequence consisting of di-, tri-, and tetra nucleotide repeats. SSR markers are frequently amplified by polymerase chain reaction (PCR), using two unique sequences which are complementary to the flanking regions as primers (Govindaraj *et al.,* 2015). PCR reactions for SSRs are performed within both forward that anneal at the 5′ and reverse primers anneal at 3′ ends of the template DNA. The polymorphisms are identified by building PCR primers for the DNA flanking the short tandem repeats region. SSR markers are consistently distributed throughout the genome, highly polymorphic, easily automated, good analytic resolution and highly reproducible which make them a preferred choice of markers (Matsuoka *et al.,* 2002). SSR are most broadly used for individual genotyping, germplasm evaluation, genetic diversity studies, genome mapping, phylogenetic and evolutionary studies.

However, build out of microsatellites requires extensive knowledge of DNA sequences, and sometimes they underestimate genetic structure measurements; hence they are developed mainly for cultivated species, rather than wild species (Mondini *et al.,* 2009). Beside this, Initial isolation of useful SSR loci are often time consuming and expensive process (Chittaranjan & Albert, 2008).

2.9.4. Single nucleotide polymorphism (SNP)

SNPs are a single nucleotide base difference between two DNA sequences or individuals that exist in coding**,** non-coding and in the intragenic regions between genes at different frequencies in different chromosome (Kumar *et al*., 2014). These markers are the most abundant source of polymorphisms and widely distributed throughout the genomes with a variable distribution among species. SNPs can be identified by using either microarrays or denaturing high-performance liquid chromatography (DHPLC) machines. They are used for a wide range of purposes, including rapid identification of crop cultivars and construction of ultrahigh-density genetic maps. They provide valuable markers for the study of agronomic or adaptive traits in plant species, using strategies based on genetic mapping or association genetics studies and show high stability and easiness of standardization (Werner *et al*., 2004).

Even though, SNP's have several advantages over other technology, there are limitations to the discovery of SNP's in the non-model organism. This is due to the expenses and technical difficulties involved in the currently available SNP isolation strategies (Bali *et al.,* 2018).

2.8. 5. Inter simple sequence repeat (ISSR)

Inter simple sequence repeat (ISSR) markers are region of the DNA located at specific region between two inverted simple repeat regions when primers are annealed on two microsatellites sequence of DNA strand. Among various PCR-based markers, ISSR markers have been revealed to be useful as novel DNA markers in studies on purposing crop improvement such as genomic fingerprinting, phylogenetic analysis and gene tagging (Isshiki *et al*., 2008). ISSR markers amplify regions between adjacent and inversely oriented microsatellites using di-, tri-, tetra- and penta nucleotide. The PCR reaction generates band of particular size for each locus that scored for a private allele in binary way as 1 or 0 (presence or absence of product), respectively. Inter simple sequence repeat (ISSR) sequences as molecular markers which will result in the detection of polymorphism and also a replacement approach to review SSR distribution and frequency. ISSR primers with the advantage that the information of the target DNA sequence is not required (Shen et al., 2010).

The ISSR is technically simple; more specific than RAPD to review genetic diversity do to its longer SSR based primers (16.25bp). It has been used for evaluating the genetic variation and genetic relatedness of crop species with higher primer annealing temperature (45°C-60°C) which enable for higher strength and bad reproducibility. The major limitations of other molecular markers like low reproducibility of RAPD, high cost of AFLP, the necessity to understand the flanking sequenceto develop specific primers for SSR polymorphism and enormous quantity of genomic DNA for AFLP (Saleh *et al.,* 2016) are overcome by this (Anne, 2006).

The generation of ISSR markers makes use of microsatellite sequences that are highly variable and ubiquitously distributed across the genome, had higher reproducibility compared to RAPDs and, take less time and low cost compared to AFLPs. Moreover, ISSR Markers are used as markers for genetic variation studies (Shafiei-Astani *et al.,* 2015), DNA fingerprinting (Shen *et al.,* 2006) and ISSR markers produce amplified products of 200–2000 bp long and are used for

genetic characterization of varied plant species such as *Hagenia abyssinica* (Tiliye *et al*., 2007), coffee (*Coffee- arabica* L.), (Esayas *et al.*, 2005; Kassahun *et al.*, 2014), rice (*Oryza- sativa*) (Gezahegn *et al*., 2010), anchote (*Coccinia abyssinica* (Lam.) Cogn.) and (Abreham *et al.*, 2014).

3.MATERIALS AND METHODS

3.1. Plant materials and sampling strategy

The study involved a total of 60 *V. galamensis* germplasm representing eight populations (Table 1 and Appendix 3) that were collected from major growing regions of Ethiopia. Seed of the study material were mainly collected from Oromia, Amhara and, Southern Nations, Nationalities and Peoples Regional State (Figure 1). For use in genomic DNA extraction, healthy young leaves were collected from six weeks old seedling (pooled from 3 plants per accession) that were grown at Adama science and Technology University. The collected leaf samples were dried with silica gel in zip locked plastic bag at room temperature and preserved until genomic DNA extraction was executed. The dried leaf was transported to Addis Ababa University, Institute of Biotechnology, plant genetics research laboratory, for genotyping.

Table 1. Number of *V. galamensis* genotypes used for molecular diversity analysis

No.	Populations	No. of genotypes
1	Borena	8
2	Derashie	7
3	East Shoa	7
4	East Hararghe	8
5	Konso	7
6	South Wollo	8
7.	West Arsi	7
8.	West Gojjam	8
	Total	**60**

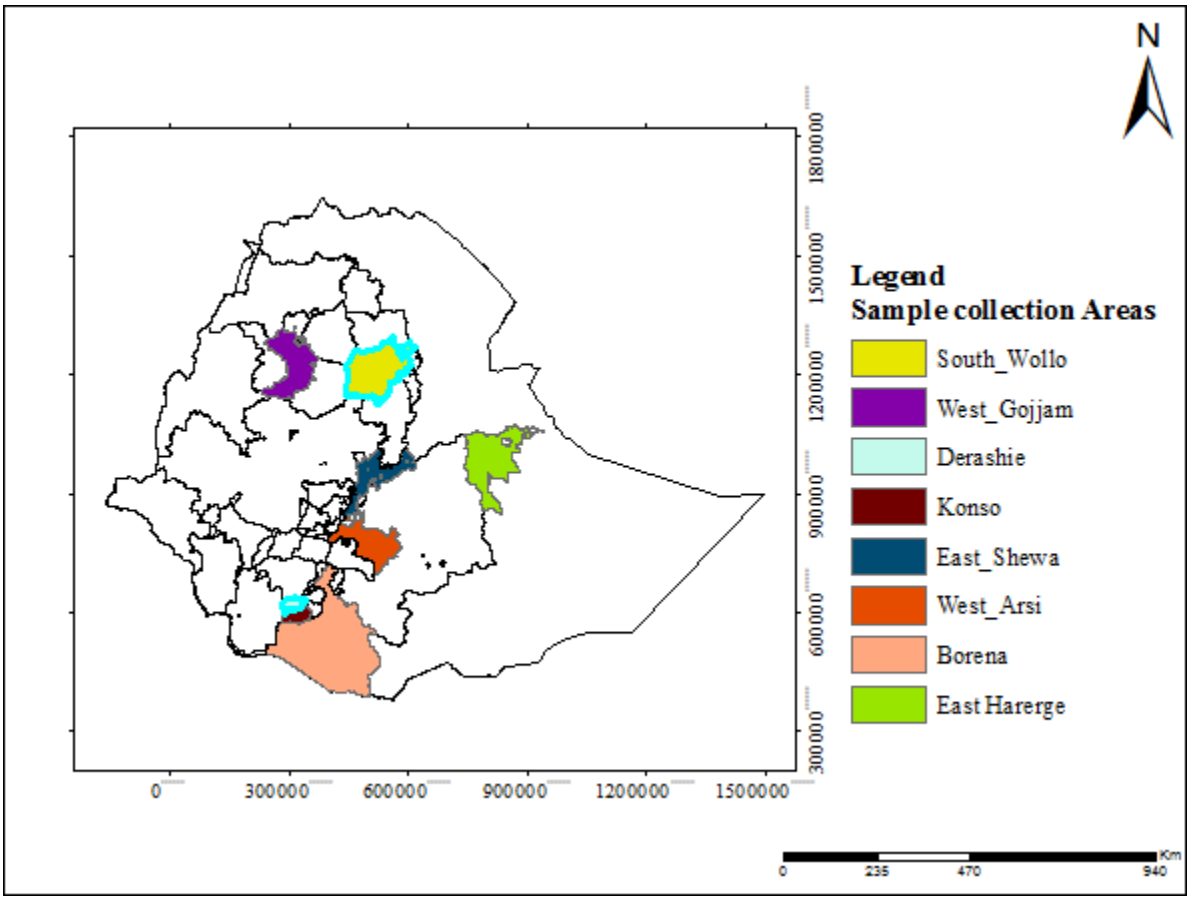


Figure 1. Map of *V. galamensis* collection areas in Ethiopia

3.2. Genomic DNA extraction

Genomic DNA extraction and downstream activities were carried out at Institute of Biotechnology Ababa University. Genomic DNA extraction was conducted according to modified cetyl tri methyl ammonium bromide (CTAB) protocol (Govers, 2005), (Appendix 4). Then, the extracted DNA quality was determined in 1% agarose gel. The DNA quantity was confirmed using Nano Drop (NanoDrop™2000/2000c) spectrophometer (Termo Scientific™. USA). Then, the final concentration of each genomic DNA samples was adjusted to 100 ng µL-1 for use in PCR amplification and stored at -20°C until use.

3.3. Primers selection and optimization

A total of thirteen available ISSR primers which were originally obtained from University of British Colombia (UBC) were screened for their polymorphism and reproducibility on eight randomly selected samples (Table 2). The primers annealing temperatures were gradient optimized.

Table 2. List of the 13 primers, sequence, amplification pattern, repeat motifs and annealingtemperature, used for optimization and screening.

Primers code	Sequence (5' - 3')	Amplification quantity	Repeat Motifs	Ta	Remark
812	(GA) 8A	Reproducible and polymorphic	di-	45^0C	Selected
818	(CA) 8G	Reproducible and polymorphic	di-	48^0C	Selected
835	(AG) 8 CC	Reproducible and polymorphic	di-	48^0C	Selected
841	(GA) 8 TC	Reproducible and polymorphic	di-	48^0C	Selected
844	(CT) 8 GC	Reproducible and polymorphic	di-	48^0C	Selected
845	(CT) 8 GG	Reproducible and polymorphic	di-	48^0C	Selected
848	(CA) 8 AG	Reproducible and polymorphic	di-	48^0C	Selected
881	(GGGTG)3	Reproducible and polymorphic	penta-	48^0C	Selected
826	(AC) 8 CC	No Banding	di-	45^0C	Not selected
814	(CT) TA	No Banding	di-	48^0C	Not selected
809	(AG) GG	No Banding	di-	50^0C	Not selected
824	(TC) G	No Banding	di-	50^0C	Not selected
860	(TG) GA	No Banding	di-	45^0C	Not selected
Total	**13**		**Selected**	**8**	

Ta = annealing temperature in degree Celsius

3.4.PCR Amplification and gel Electrophoresis

The polymerase chain reaction was conducted using biometra 2003 Thermo cycler machine. PCR Amplification was carried out in a 20 µl final reaction volume containing 2µl template DNA (100 ng μL^{-1}), 11.7µl dd H_20 (ultra-pure), 0.5µl dNTPS, 2.5µl PCR buffer (10x), 2.5µl $MgCl_2$ (25 mM), 0.5µl primer (20 pmol/µl), and 0.3µl Taq polymerase. The PCR program was set at an initial denaturation step for 4 minutes at 94^oC followed by 40 cycles of 1-minute denaturation at 94^oC, primers annealing at 45/48^oC (depending on the primers) for 30s, primer extension at 72^oC for 1-minute 30s and final extension at 72^oC for 7 minutes followed with holding temperature at 4°C. The PCR product was stored at 4^oc until loaded on the gel for electrophoresis. The products were fractionated in 2% agarose gel using 1 X TBE buffer at 100 V for 2 hrs. The gel was stained in gel red (2 µl), visualized under UV light and subsequently photographed using a BIO-RAD Gel Doc TM EZ Imaging System. A 100 bp size ladder was used to estimate the size of the amplified products.

3.5. Data scoring and analysis

Each ISSR bands were treated as a unit character and clear bands were counted manually after taking gel image with gel documentation system and scored as binary data; '1' for band present, '0' for band absent, and '?' for missing data. The resulted binary data matrix was subjected to different molecular software to execute the various genetic diversity parameters. Number of alleles (Na), number of effective alleles (Ne), Shannon's information index (I) gene diversity (h), genetic identity and genetic distance (GD), analysis of molecular variance (AMOVA), percentage of polymorphic loci (% PPL) and principal coordinate analysis (PCoA) were calculated using GenAlex version 6.5 software (Peakall & Smouse, 2012). Locus based diversity indices including major allele frequency (MAF) and polymorphic information contents (PIC) were analyzed using power marker version 3.25 software (Liu & Muse, 2005). Population differentiation (Fst) and gene flow (Nm) were determined using POPGENE version 1.32 (Yeh *et al.,* 1999).

A genetic dissimilarity matrix parameter (Unweighted Pair Group Method with Arithmetic mean (UPGMA) was calculated using Darwin software version 6.0 (Perrier & Jacquemoud-Collet, 2006). Population structure and admixture patterns were determined using the admixture model based on the Bayesian algorithm implemented in STRUCTURE software version 2.3.3 (Pritchard *et al.,* 2000). The admixture model with correlated allele frequencies was used, assuming that the genome of each individual resulted from the mixture of K ancestral populations. To estimate the true number of population cluster (K), a burn-in period of 30,000 was used in each run, and data were collected over 50,000 Markov Chain Monte Carlo (MCMC) replications for K = 1 to K = 8 using 20 iterations for each K. The optimum K value was predicted following the simulation method of Evanno *et al*. (2005) using the web-based STRUCTURE HARVESTER version 0.6.92 (Earl, 2012). A bar plot for the optimum K was determined using Clumpak beta version (Kopelman *et al*., 2015).

4.RESULTS

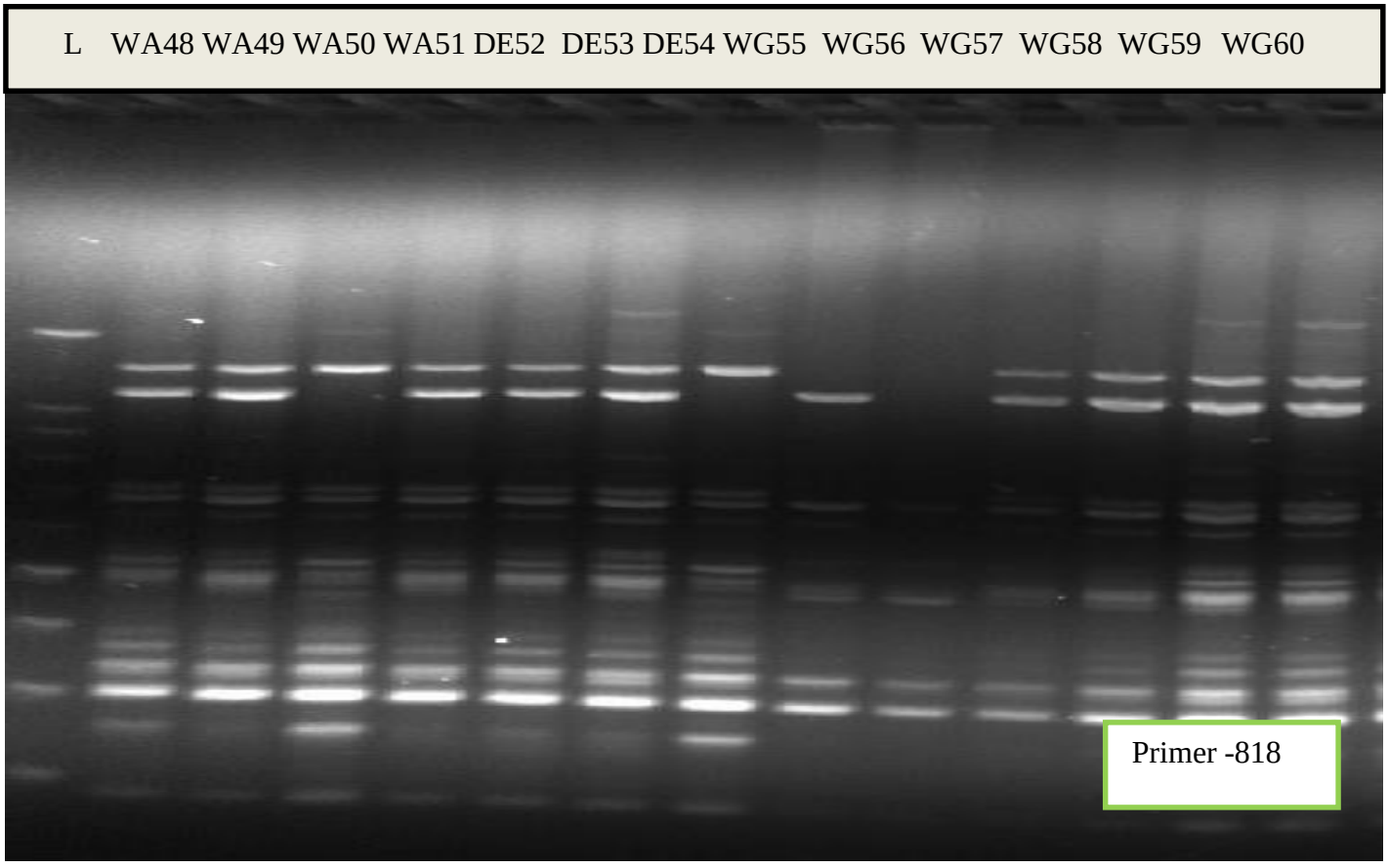
4.1. ISSR primers variation and level of polymorphism

In this study, eight (8) polymorphic markers (seven di-, and one penta-nucleotide repeats) were used for molecular profiling of the 60 *Vernonia galamensis* germplasm. Accordingly, the 8 ISSR primers generated a total of 83 clear and scorable bands with an average of 10.4 bands per primer. The highest number of bands (13) was recorded for primer 881 while the lowest bands (8)were generated by primer 845 (Table 2). The effective number of alleles per marker ranged from 3.41 (881) to 7.22 (841) with an average of 5.6. The highest major allele frequency (0.57) was recorded for the primer 881, while the lowest (MAF =0.20) was obtained by primer 848 with the overall mean of 0.35. The PIC values for the ISSR loci ranged from 0.50 (for the marker 835) to 0.71 for marker 881, with an average mean value of 0.51 (Table 3). The gene diversity index (h) for all primers ranged from 0.27 (for the markers 812 and 818) to 0.63 (for the marker 881), with an average value of 0.39 alleles per locus (Table 3). In addition, Shannon information index ranged from 0.24 (for the primer 812) to 0.61 (for the primer 881, with an average of 0.40 (Table 3). The lowest genetic differentiation (Fst=0.04) and the highest gene flow (Nm= 3.71) were scored for primer 835 (Table 3). Conversely, the highest genetic differentiation value of 0.27 and the lowest gene flow (Nm= 0.98) were recorded for the primer 848 (Table 3).

Table 3. Summary of genetic diversity indices for the 8 ISSR loci across the sixty *V. galamensis germplasm* in Ethiopia

Primer code	Na	Ne	MAF	h	I	PIC	Fst	Nm
812	9.0	4.81	0.24	0.27	0.24	0.56	0.18	2.66
818	12.0	5.77	0.21	0.27	0.43	0.59	0.15	1.97
835	10.0	6.80	0.45	0.42	0.33	0.50	0.04	3.71
841	11.0	7.04	0.53	0.37	0.54	0.53	0.11	1.53
844	11.0	4.63	0.33	0.31	0.44	0.51	0.24	2.09
845	8.0	5.34	0.24	0.44	0.30	0.51	0.18	1.45
848	9.0	7.22	0.20	0.50	0.34	0.65	0.27	0.98
881	13.0	3.41	0.57	0.63	0.61	0.71	0.09	1.07
Average	**10.4**	**5.6**	**0.35**	**0.39**	**0.40**	**0.51**	**0.16**	**1.93**

Key: Na=Number of allele, Ne=number of effective allele, MAF = Major allelic frequency, h= gene diversity; I = Shannon information index, Fst = Fixation index, Nm = gene flow



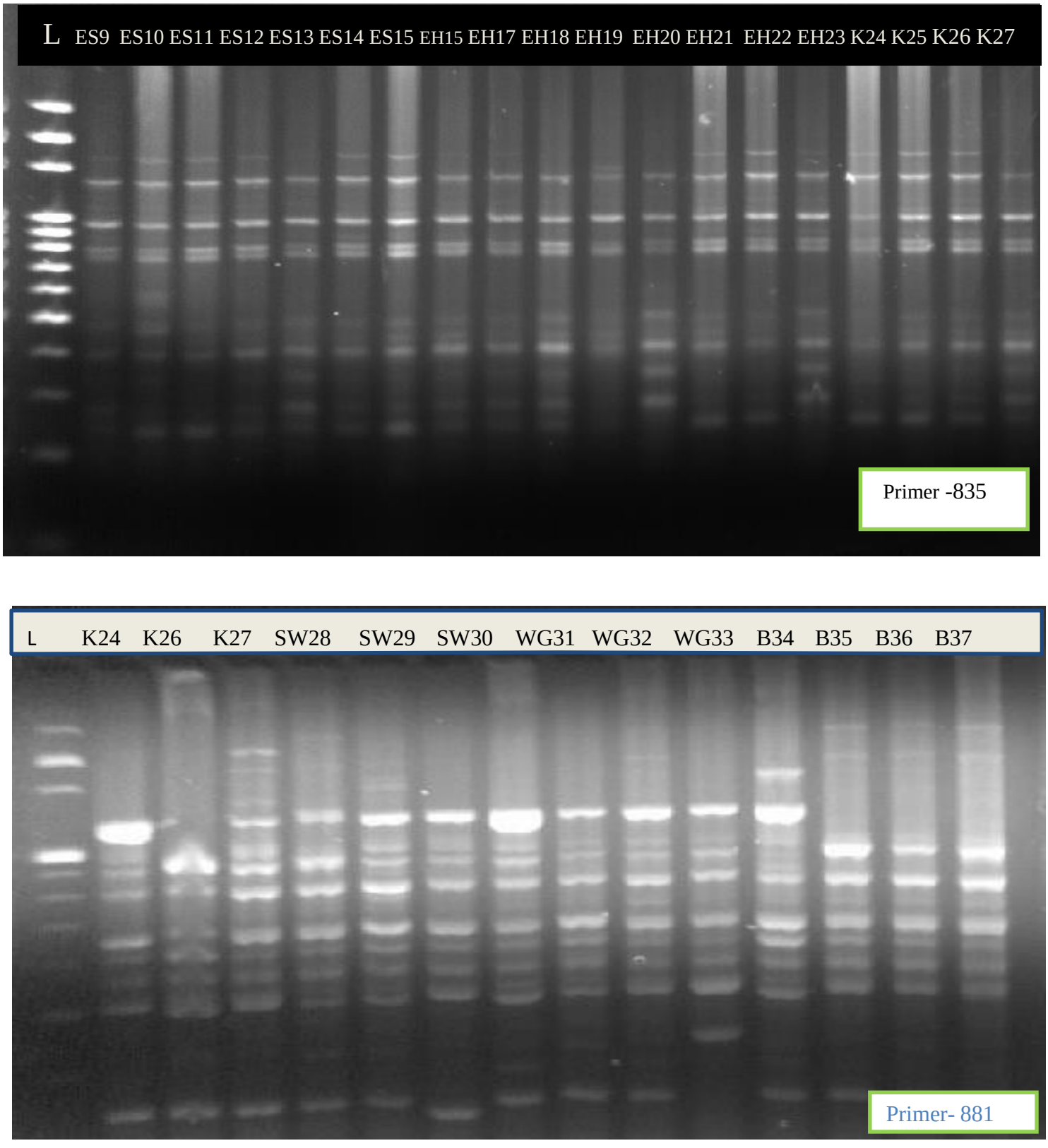


Figure 2. ISSR profiles of *V. galamensis* genotypes generated for primer 818,835 and 881 respectively, L = DNA Ladder (100 bp). letters in lanes indicate the germplasm. ES=East Shoa, EH= East Hararghe, B= Borena k= Konso DE=Derashie, WG=West Gojjam, WA=West Arsi, SW=South Wollo

4.2. Genetic diversity within and among the populations

Table 4 presents a summary of the different genetic diversity estimates over all the loci across populations. Among the assessed *V. galamensis* populations, East Shoa populations exhibited relatively the highest gene diversity (h = 0.29) and Shannon diversity index (I = 0.43), followed by the population of East Hararghe that resulted (h = 0.28; I = 0.41) (Table 3). Nei's unbiased measure of genetic diversity (Uh) ranged from 0.22 (South Wollo) to 0.34 (East Shoa) with an overall mean of 0.28 (Table 4). The percentage of polymorphic loci of the populations ranged from 50.60 (South Wollo) to 75.90 (East Shoa).

Table 4. Summary of allelic patterns and diversity indices across populations over the eightISSR loci

Population	N	Na	Ne	I	H	Uh	PPL
Borena	8.00	1.31	1.35	0.31	0.21	0.24	56.63
East Shoa	7.00	1.60	1.50	**0.43**	**0.29**	0.34	**75.90**
East Hararghe	8.00	1.57	1.48	0.41	0.28	0.32	73.49
Konso	7.00	1.30	1.45	0.40	0.26	0.30	71.08
Derashie	7.00	1.39	1.38	0.33	0.23	0.26	60.24
South Wollo	8.00	1.22	1.33	**0.28**	**0.19**	0.22	**50.60**
West Arsi	7.00	1.30	1.30	0.29	0.20	0.22	53.01
West Gojjam	8.00	1.53	1.45	0.39	0.25	0.30	71.08
Average	7.5	1.40	1.41	0.36	0.24	0.28	64.01

Where N= Number of samples, Na= Number of alleles, Ne = Number of effective alleles, I = Shannon's Information Index, PPL= Percentage of polymorphic loci. h = Diversity uh = Unbiased diversity

4.3. Genetic similarity among the *V. galamensis* populations

Table 4 presents a summary of the pairwise genetic similarity of the eight populations computed based on Jaccard's similarity index. The largest (0.98) pairwise similarity coefficient was between Konso and Borena populations while smallest pairwise similarities (0.53) was between West Arsi and West Gojjam) populations (Table 5). The second highest genetic similarity was observedbetween the populations of West Arsi and East Shoa (0.90) (Table 5).

Table 5. Similarity matrix among the *V. galamensis* populations attained by Jaccard's similarity coefficient based on ISSR data.

	BOR	EHG	KON	ESH	DER	SW	WA	WG
BOR	1.00							
EHG	0.79	1.00						
KON	**0.98**	0.83	1.00					
ESH	0.78	0.78	0.63	1.00				
DER	0.67	0.68	0.71	0.81	1.00			
SW	0.68	0.70	0.75	0.74	0.72	1.00		
WA	0.76	0.73	0.72	0.90	0.65	0.83	1.00	
WG	0.75	0.78	0.70	0.74	0.69	0.68	**0.53**	1.00

Key: BOR=Borena, ESH = East Shoa, WA = West Arsi, EH = East Hararghe, WG = West Gojjam, SW = South Wollo, KON = Konso, DER = Derashie

4.4. Analysis of molecular variance (AMOVA)

Analysis of molecular variance (AMOVA) revealed presence of significant (p = 0.001) and higher genetic differentiation (0.16) with high gene flow (1.9) and the variability within populations accounted (65%) of the total genetic variation and 35% of the total variation was due to the among populations genetic variation (Table 6).

Table 6. Analysis of molecular variance (AMOVA) showing distribution of genetic variation within and among populations of *V.galamensis* using ISSR markers

Source of variation	**Df**	**SS**	**MS**	**Est. Var.**	**% Variation**	**F-Statistics**	**P-value**
Among Pops	7	407.84	58.26	6.23	35%	0.16	0.001
Within Pops	52	602.96	11.60	11.60	65%		0.001
Total	59	1010.80		17.82	100%		
Nm		1.93					

Df = Degrees of Freedom SS = Sum of Squares; MS = Mean Square; Est. Var. = Estimated Variability, Nm=haploid populations

4.5. Cluster and principal co-ordinates analysis

4.5.1. Cluster analysis

The neighbor-joining-cluster analysis divided the 60 individuals into three main clusters (Cluster C1, C2 and C3) (Figure 2). The first cluster (C1) consisted of 15 samples, and most of the samples were from west Arsi and west Gojjam this population may appear together due to long distance transfer of pollen & Migration of individuals through different pollinators. The second cluster (C2) was considered as major group that contained of 33 samples. The group composed of individuals collected from Konso, Derashie and South Wollo. Cluster three (C3) was considered to be the smallest group that composed of 12 accessions that were obtained from Borena and East Shoa (Figure 2). The clustering analysis confirmed the presence of intermixing of the individuals of the variouspopulations except Borena and West Gojjam populations, where they form their own groups without any intermixing within each other and with other populations at the sub-cluster levels (Figure 2). Some of the individual accessions collected from the same region tend to spread all over tree without forming their own group, likely due to the presence of considerable gene flow among population (Nm= 1.93).

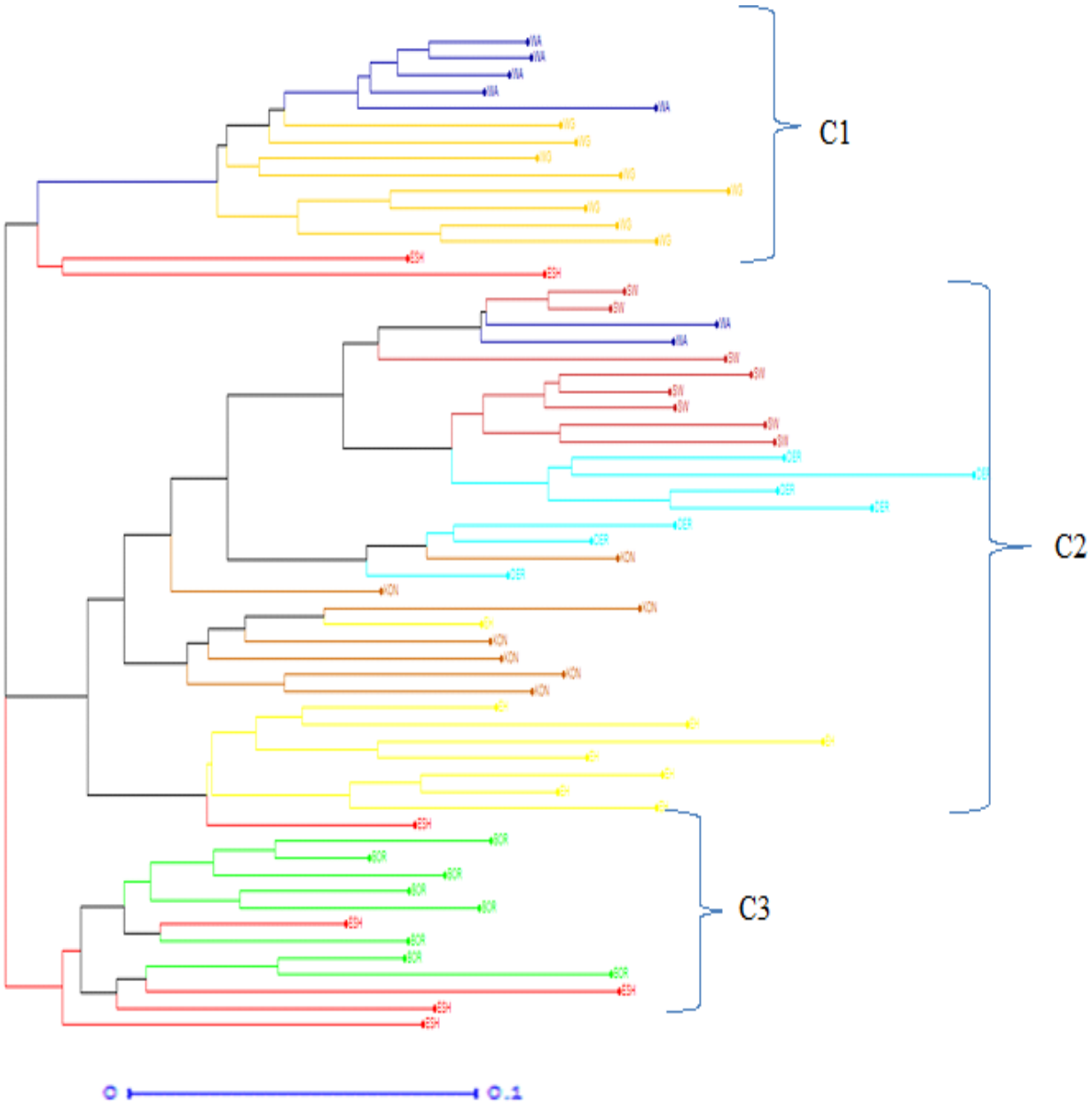


Figure 3. Neighbor-joining tree analysis of 60 *V. galamensis* individuals using 8 ISSR markers.

Key: BOR=Borena(green), EH= East Hararghe (yellow), ESH=East Shoa(red), DER= Derashie (**Aqua**), KON= Konso (Dark red), SW=South Wollo (deep red), WA= west Arsi (**blue**), WG= west Gojjam (orange)

UPGMA analysis based on the eight populations of *V. galamensis* revealed the presence of two major clusters (Cluster I and II). The first major cluster consisted of the populations of East Shoa, East Hararghe, Borena and Konso, while the second clusters contained the populations of Derashie, South Wollo, west Arsi and west Gojjam (Figure 3). Each population showed that moderate grouping based on their geographical collection areas.

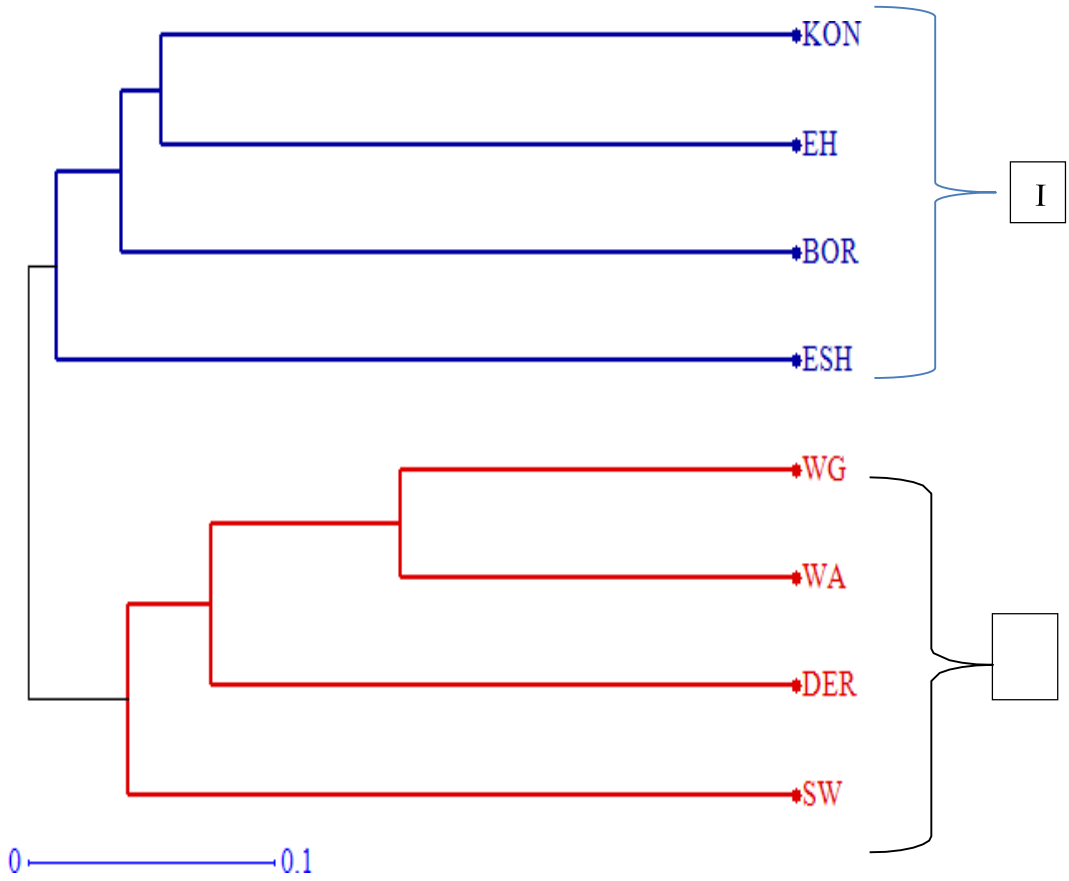


Figure 4. UPGMA based dendrogram showing genetic relationships among the eight Populations of V. galamensis using ISSR primers.

ESH=East Shoa, EH=East Hararghe, BOR=Borena, KON= Konso, SW= South Wollo, DER= Derashie, WA= West Arsi, WG= West Gojjam

4.5.2. Principal co-ordinates analysis

The principal coordinate analysis (PCoA) conducted using the 60 *V. galamensis samples* roughly grouped the populations into four groups (Figure 5). None of the clusters composed of exclusively individual of a single population. The percentages of principal coordinates accounted 35.24% of the total variation. The first, second, and third principal coordinate axes showed 14.42%, 12.78%and 8.04%, respectively. The analysis revealed that the tested populations is not tended to form grouping based on their geographic regions since some individuals were intermixed from one to another population (Figure 5).

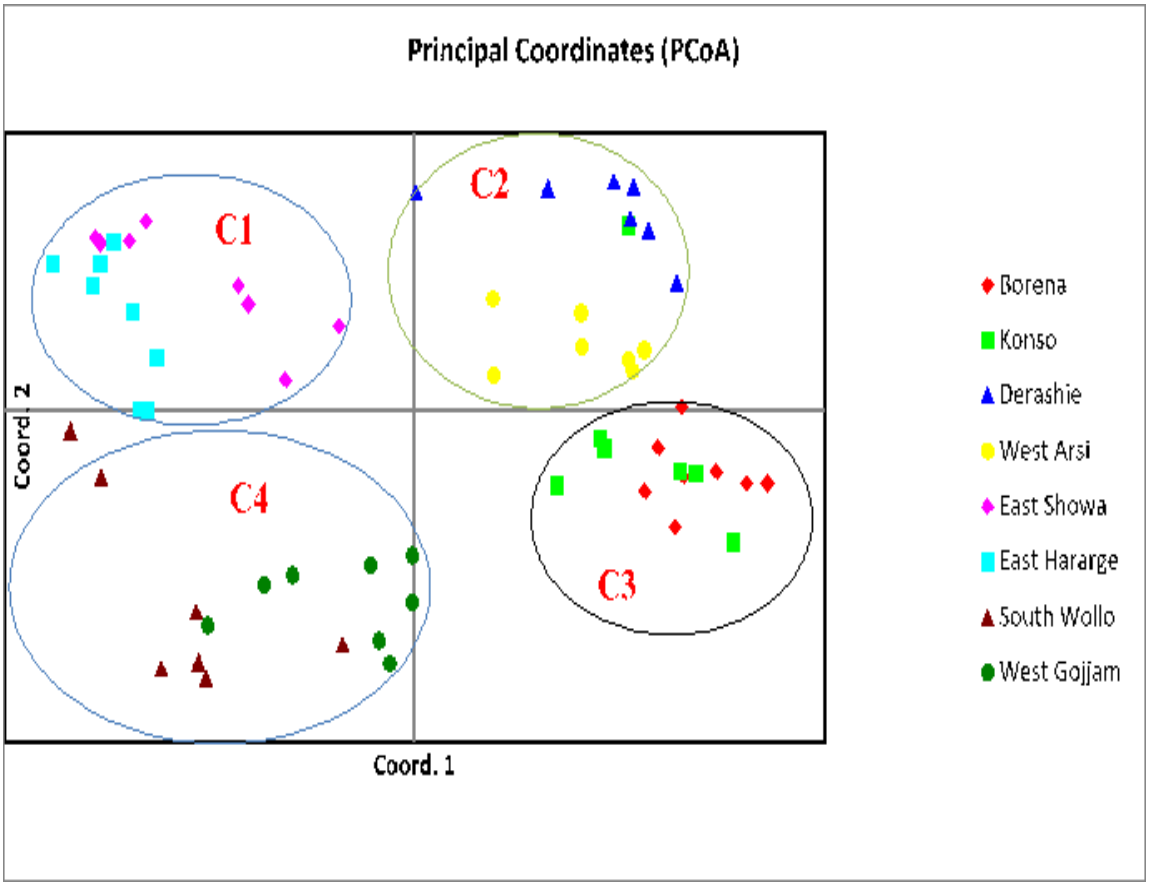


Figure 5. Principal coordinate analysis of 60 V. galamensis accessions using ISSR markers.

4.6. Population structure analysis

To provide further evidence and infer population structure, a Bayesian algorithm implemented in the STRUCTURE 2.3.4 Software analyses was used. The analysis revealed delta K became a sharp peak at K=2 (Figure 5a) confirming the presence of two subpopulations. Based on this value, Clumpak result (bar plot) detected high level of genetic admixture and hence there was low degree of geographic origin-based structuring of populations.

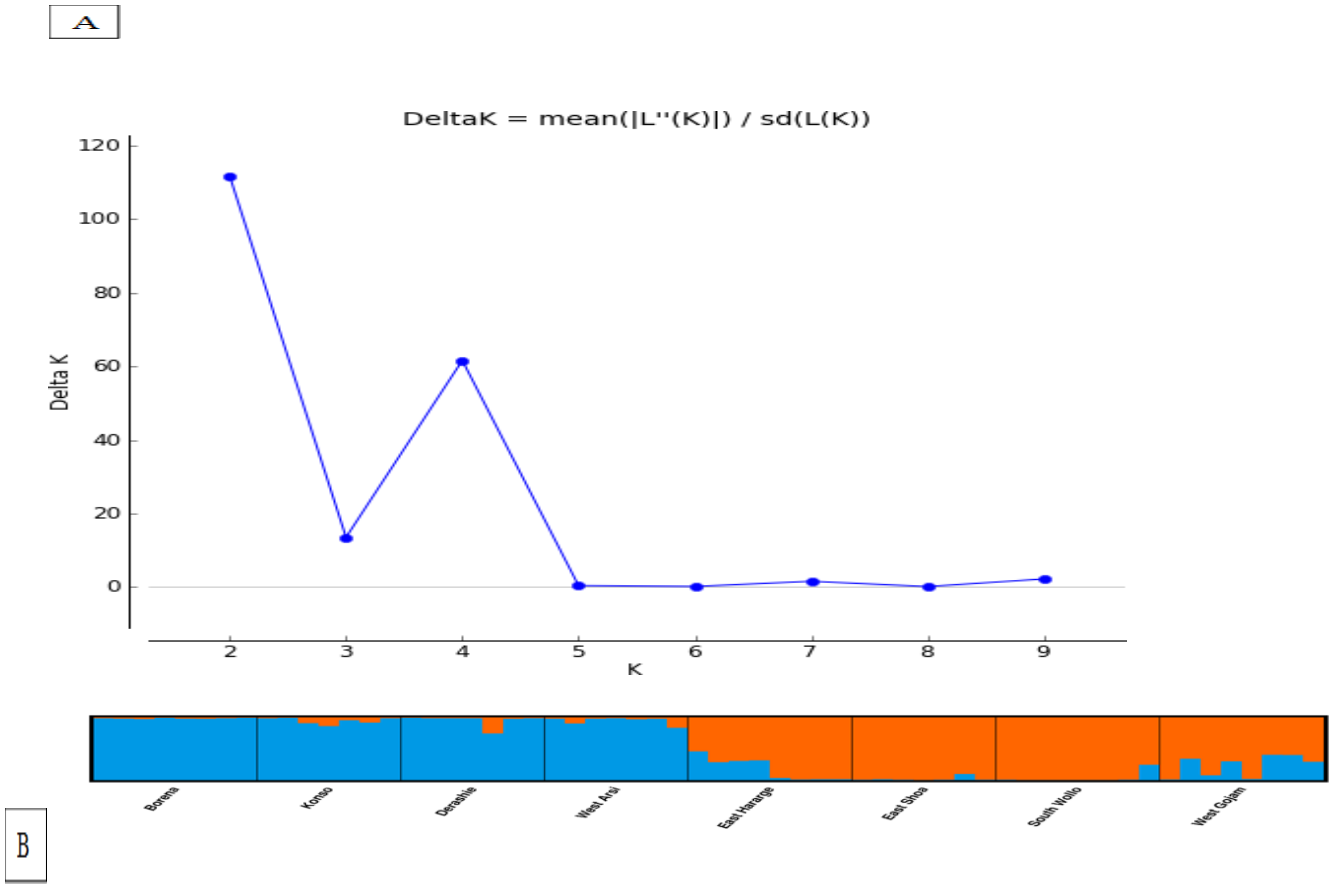


Figure 6. Structure analysis for 8 populations representing 60 V. galamensis germplasm in Ethiopia.

(A) delta K value estimated using Evanno *et al*. (2005); method and (B) Bayesian model-based estimated population structure for K = 2 rendering to geographical locations. The different (blue and orange) colors represent genetic groups or sub-populations designated by Structure Harvester.

5.DISCUSSIONS

This study was the first report to use ISSR markers to assess genetic diversity and population structure in *V. galamensis*. In the present study, eight ISSR markers amplified a total of 83 alleles. The number of alleles per locus ranged from eight for primer 845 to thirteen for primer 881 with a mean of 10.4 alleles per locus. The mean number of alleles observed in the current study is higher than the report of (Alemneh, 2020), who observed total of 79 alleles with a mean of 3.9 alleles per locus using 20 SSR markers in 150 *V. galamensis* germplasm. The result also justly higher than the result of Aikpokpodion *et al.,* (2018) on genus *Vernonia* (*Vernonia amygdalina*) reported that total of 29 alleles with a mean of 5.8 per locus using 5 RAPD markers. The difference could be due to the types of markers, the size of the population, the number of germplasms used and reproducibility of the markers.

Gene diversity (h) and Shannon's information index (I) of ISSR primers level used in this studies were ranged from 0.27-0.63 and 0.24-0.61 with overall mean of 0.39 and 0.40 respectively. The level of polymorphism observed across all 8 primers used is approving the genetic variation within *V. galamensis* collection. Shannon diversity index value observed in the present study was lower than Nwakanma *et al.,* (2018), 0.76, genetic diversity study for 50 vernonia species in Nigeria using RAPD Markers. This is due to types of primers used. The average percent of polymorphism per population revealed in the present study (64%) was considerably higher than that reported by Ramalema *et al.,* (2010) which was 40%. This implies that Ethiopian *V. galamensis* populations exhibit higher genetic diversity.

Moreover, the PIC of the loci ranged from 0.50 -0.71 with overall mean of 0.51, indicating that all the used markers are highly informative and very suitable to assess genetic diversity and phylogenetic relationship analysis among *V. galamensis* genotypes. According to Botstein *et al.* (1980), marker with PIC values below (0.25) is somewhat informative, a marker with PIC values between (0.25) and (0.5) is considered as moderate informative, and those greater than (0.5) is said to be highly informative. The authors also report that the maximal value of PIC for dominant markers is 0.5. However, for the markers with multiple alleles and equal distribution in the population, the PIC values are much higher than 0.5. Based on our findings, PIC values was greater than 0.50.

Analysis of the genetic variation within populations revealed that relatively *V. galamensis* collected from East Shoa exhibited the highest gene diversity (h= 0.29) and Shannon diversity index (I = 0.43), which was followed by the populations of East Hararghe that showed gene diversity value of 0.28 and Shannon diversity value of 0.41. This implies that the areas representing these populations could be considered as genetic diversity hot spots and a potential in-situ conservation sites for *V. galamensis*.

AMOVA revealed that *V. galamensis* has high genetic differentiation (Fst =0.16) with relatively low among populations variation which accounted 37% of the total genetic variation. In our study high level of gene flow (Nm=1.93) confirmed the presence of high genetic exchange or gene flow between the populations of *V. galamensis.* This result is consistent with earlier studies on similar and related crops. For instance, Alemneh (2020) reported that 37% of the genetic variations was accounted as among populations, using SSR. Similarly, using SSR markers Kassahun *et al.,* (2006) in coffee and Edossa *et al.,* (2010) in Lentil reported an among population genetic variation of 22% and 43.7%, respectively. Conversely, Oumer *et al*., (2011) in white lupine using ISSR primers and reported that 74.6% of the total genetic variation is due to within population, leaving only 25.4% of among populations.

NJ, UPGMA, and PCoA clustering analyses were employed to better visualize the genetic relationships of the 60 samples of *V. galamensis* that originated from three different region of Ethiopia. The Neighbor joining cluster analysis showed a moderate clustering pattern, confirming the presence of genetic differentiation among the populations and suggesting that the genetic background of *V. galamensis* populations does somewhat correlate with their geographical origin. It was observed that some samples from different populations (collection sites) clustered together and formed moderate pattern of clustering on the basis of geographic origins, which may imply the presence of gene flow between populations of different collection sites with different factors like wind insect and birds. UPGMA based clustering of the eight populations resulted in two major clusters, confirming the presence of weak population differentiation as a result of the presence of high gene flow. Moreover, the PCoA showed that the genotypes of the eight populations not tended to form grouping based on their geographic regions.

However, some individuals were intermixed from one to another population, likely due to the presence of gene flow among *V. galamensis* populations.

Population genetic structure analysis also revealed that delta K reached its maximum (a sharp peak at K=2), indicating that the populations of *V. galamensis* studied could be grouped into two main populations with genetic admixture. The analysis revealed that most populations possessed genetic background (alleles) from both sub-populations confirming the presence of admixture due to the nature of the primers used, the origin, distribution and pollination patterns of *V. galamensis* germplasms and gene flow among the populations.

6.CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

V. galamensis is one of the most important oil crops that has wide industrial applications. It is widely distributed in the Eastern, Southern and South-eastern parts of Ethiopia. Knowledge of the genetic diversity of crop plants is important for its improvement through selection and also conservation purposes. All the markers used were highly polymorphic and informative to describe the genetic diversity and population structure.

Most of the studied populations showed high gene diversities and partitioning of genetic diversity by analysis of molecular variance using grouped populations revealed that outof the total genetic diversity, most of the ISSR diversity was distributed between individual, with the remaining diversity being distributed among the populations. Cluster analysis of the present study also showed that the existence of weak genetic differentiation and most of the *V. galamensis* population shared genetic background originated from two sub-populations hinted that the presence of high genetic admixture resulted from gene flow among populations which was confirmed by UPGMA and PCoA analyses. In this study ISSR markers were observed to be an appropriate molecular marker for generating the detail intraspecific genetic diversity data to evaluate the extent and distribution of genetic diversity within and among *V. galamensis* germplasm.

In conclusion, among the studied populations, based on all analysed parameters those from East Shoa and East Hararghe showed high gene diversity suggesting that these locations could be potential sites for *V. galamensis* genetic diversity and should be targeted for traits improvement through breeding and conservation. Conversely, it was observed that South Wollo population had lowest genetic diversity which should need special attention in order to balance genetic resources of such beneficial genes across the various agro ecology and genetic origin. Therefore, the present study generated valuable information that is important for *V. galamensis* improvement through selection, conservation and provides baseline information for future studies.

6.2. RECOMMENDATIONS

Based on the study, the following recommendations were made:

- Because of various reason, this study is not exhaustive in terms of geographic area coverage and number of primers used. Thus, similar further studies should be carried out on other populations from the remaining parts of the country using a greater number of primers to generate a nationwide picture of the diversity of the crop.
- The study involved ISSR based genetic analysis which is relatively less powerful as compared to advanced marker systems. Thus, further genetic diversity analysis based on high density markers such as SNPs can provide better information.

7.REFERENCES

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.

Adebayo, M.A., Menkir, A., Gedil, M. (2015). Diversity assessment of drought tolerant exotic and adapted maize (Zea mays L.) inbred lines with microsatellite markers. J. Crop Sci. Biotechnol. 18: 147–154.

Aikpokpodion, P. O., Abebe, J. and IGWE, D. (2018). Genetic diversity in *Vernonia amygdalina* Delile accessions revealed by random amplified polymorphic DNAs (RAPDs). Journal of Biotechnology, 99 (2): 147–152.

Alemneh Mideksa, (2020). Genetic diversity, oil content and fatty acid composition of Vernonia [*Vernonia galamensis (Cass.) Less]* germplasm from Ethiopia, Ph.D. Dissertation Addis Ababa University.

Althoff, D. M., Gitzendanner, M. A., Seagraves, K. A. (2007). The utility of amplified fragment length polymorphisms in phylogenetics. A comparison of homology within and between genomes. *Syst. Boil.* 56: 477- 484.

American Journal of Animal and Veterinary Sciences, 9(1), pp.6-13.

Anne, C. (2006). Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects', *Genetic*, vol. 127, pp. 101–120.

Appl. Genet. 128:2131–2142.

Arya, V., Yadav, S. and Yadav, J. P. (2011). *Intra Specific Genetic Diversity of Different Accessions of Cassia Occidentals by RAPD markers*. Department of Genetics, M.D. University, Haryana, India. 4: 97 -123.

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.

Bali S, Robinson BR, Sathuvalli V, Bamberg J, Goyer A. (2018). Single Nucleotide Polymorphism (SNP) markers associated with high folate content in wild potato species.

2: 13 -17.

Becker, H. (2005). Genetic variability and interrelationship of traits in the industrial oil crop *Vernonia galamensis*. Euphytica 142, 119–129.

Bhandari, H. R, Bhanu, A. N, Srivastava, K, Singh, M. N. and Shreya, A.H. (2017). Assessment of genetic diversity in crop plants - An overview. *Adv. Plants. Agric. Res*. 7(3): 110-154

Bornet, B. and M. Branchard. (2002). No anchored inter-simple sequence repeat (ISSR) markers: Reproducible and specific tools for genome fingerprinting. Plant Mol Biol Reporter 19: 209–215.

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.

Burow, G., Franks, C. D., Xin, Z. and Burke, J. J. (2012). Genetic diversity in a collection of Chinese sorghum landraces assessed by microsatellites. *American Journal of Plant Sciences*, 3: 1722-1729.

Chittaranjan, K. and Albert, G. (2008). *Principles and Practices of Plant Genomics Volume 1: Genome Mapping*. Science Publishers. South Carolina, USA. pp 393.

Dalirsefat, S., Meyer, A. and Mirhoseini, S. (2009). Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silk worm, bombyxmori. *J. Insect Sci.* 9: 71-79.

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.

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.

Esayas Aga, Endashew 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.

Esquinas-Alcazar, J. (2005). Protecting crop genetic diversity for food security: Political, ethical

and technical challenges. Nat. Rev. Genet 6: 946–953

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUC TURE: a simulation study. *Mol. Ecol.* 14, 2611–2620.

Fedlu Hassen, Endashew Bekele, Workneh Ayalew and Tadele Dessie, (2007). Genetic variability of five indigenous Ethiopian cattle breeds using RAPD markers. *Afro. J. BiothecnnolBiotechnol* 6:2274-2279.

Fu, B.Y. (2015). Understanding crop genetic diversity under modern plant breeding. *Theor. Appl. Genet. 128:2131–2142.*

Getachew Ali (2008). Genetic diversity of agro-morphological traits, self-compatibility and chromosomal cytology in Ethiopian populations of *Vernonia galamensis* (L). MSc. Thesis. Addis Ababa University, Addis Ababa, 61 pp.

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

Gilbert, M. G. (1986). East African *Vernonia* (Compositae), a revision of the *Vernonia galamensis* complex. Kew Bulletin, 41 (1): 19-35.

Govers, (2005). DNA Extraction Protocol for Plants with High Levels of Secondary Metabolites and Polysaccharides without Using Liquid Nitrogen and Phenol", International Scholarly Research Notices, 6:62-71.

Govindaraj, M., Vetriventhan, M. and Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances. *Gene. R. Inter.*

Hadebe, T. S., Modi, A. T. and Shimelis, H. A. (2017). Determination of optimum ethyl methane sulfonate conditions for chemical mutagenesis of selected *Vernonia* (*Centrapalus pauciflorus*) accessions. *South African Journal of Plant and Soil*, 1-7.

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

Hamrick JL and Godt MJW. (1989). Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and germplasm resources. Sinauer Associates, Sunderland, pp 43–63.

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

Wallingford, UK, pp. 512.

Hussein Shimelis and A. Hugo. (2011). Determination of Selection Criteria for Seed Yield and Seed Oil Content in Vernonia (*Vernonia galamensis* variety *ethiopica*). Industrial Crops and Products 33, 436-439.

Hussein Shimelis and Gwata, E. T. (2013). Evaluation of the Agronomic performance of Vernonia (*Vernonia galamensis*) germplasm. *African Center for Crop Improvement, Scottsville, South, DOI: 10.5772/55016.*

Hussein Shimelis, A., Mashela, P. W. and Hugo, A. (2013). Principal agronomic and seed oil traits in the industrial oil crop vernonia (Centrapalus pauciflorus var. ethiopica). South African Journal of Plant and Soil, 30: 131–137.

Hussein Shimelis, Mashela, P.W. and Hugo, A. (2008). Performance of Vernonia as an alternative industrial oil crop in Limpopo Province of South Africa. *Crop Science*, 48: 236.

Idrees, M. and Irshad, M. (2014). Molecular markers in plants for analysis of genetic diversity*: European Academic Research*, 1 (1).

Isshiki, (2008).Genetic variability of Indonesian eggplant (Solanum melongena) based on ISSR markers, 20:3049-3055.

Jaworski, J. and Cahoon, E.B. (2003). Industrial oils from transgenic plants. Curr.Opin. Plant Biol. 6, 178–184.

Jones N, Ougham H and Thomas H. (1997). Markers and mapping: we are all geneticists now. New Philologist; 137(1):165-77.

Kassahun Tesfaye, Govers, K., Endashew 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.

Kedir Wulchafo, (2019). Genetic Diversity of Vernonia (*Vernonia Galamensis* L.) in Ethiopia International Journal of Engineering and Information Systems *Vol*. 3:31-38.

Kopelman, N.M., Mayzel, J., Jacobson, M., Rosenberg, N.A. and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources, 15*(5), pp.1179-1191.

Kumari, N. and Thakur, S.K., (2014). Randomly amplified polymorphic DNA-a brief review. American Journal of Animal and Veterinary Sciences, 9(1), pp.6-13.

Liu, K. and Muse, S.V., (2005). Power Marker: an integrated analysis environment for genetic marker analysis. Bioinformatics, 21(9), pp.2128-2129.

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

Misra, A., Shasany, A. K., Shukla, A. K., Darokar, M. P., Singh, S. C., Sundaresan, V., Singh, J., Bagchi, G. D., Jain, S. P., Salkia, D. and Khanuja, S. P. S. (2010). AFLP markers for identification of swertia species (Gentianaceae). Genet. Mol.Res. 9: 1535-1544.

Mondini, L., Noorani, A. and Pagnotta, M. A. (2009). Assessing plant genetic diversity by molecular tools. Diversity 1: 19-35.

Montes, L. R., Azurdia, C., Jongschaap, R. E. E., Van Loo, B. E., Visser, R. and Mejia, L. (2009). Global Evaluation of Genetic Variability in Jatrophacurcas. Wageningen University Plant Breeding Research Day, Wageningen.

Moose, S. P. and Mumm, R. H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiol. 147: 969- 977.

Nagaraju, J., Reddy, K. D., Nagaraja, G. M. and Sethuraman, B. N. (2001). Comparison of Multi Locus RFLPs and PCR-based Marker Systems for Genetic Analysis of the Silk Worm, Reviews, vol. 3, no. 2, pp. 32–45.

Naliaka S. (1990). Study of the wood anatomy of the Vernonia leopoldii complex with the implication of its taxonomy. M.Sc Thesis, Alemaya University of Agriculture, Addis Ababa, Ethiopia.

Nwakanma, N. M., Oboh, B. O., Adekoya, K. O. and Ogunkanmi, L. A. (2018). Genetic diversity of Vernonia as revealed by random amplified polymorphic DNA (RAPD) markers. Egypt. Acad. J. Biolog. Sci. 9 (1): 27-38.

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

Perrier, X. and Jacquemoud-Collet, J.P. (2006). DARwin–dissimilarity analysis and representation for windows. Version, 5, p.158.

Preetha, S. and Raveendren T. S., (2008). Molecular marker technology in cotton," Biotechnology and Molecular Biology Reviews, vol. 3, no. 2, pp. 32–45.

Pritchard, J.K., Stephens, M. and Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics, 155(2), pp.945-959.

Prude, R. E. JR., Kenneth, D.C., and Michael, G. G. (1986). *Vernonia galamensis* potential new crop source of epoxy acid. Econ. Bot. 40, 54–68.

Rahman A.H.M.M., M.S. Alam, S.K. Khan, A. Ferdous, A.K.M. Rafiul Islam and M.M. Rahman. (2008). Taxonomic Studies on the Family Asteraceae (Compositae) of the ajshahi Division. *Research Journal of Agricultural and Biological Sciences* 4, 134-140.

Ramalema, S. P., Shimelis, H., Ncube, I., Kunert, K. K. and Mashela, P.W. (2010). Genetic analysis among selected vernonia lines through seed oil content, fatty acids and RAPD DNA markers. *African Journal of Biotechnology*, 8 (4): pp 635-640.

Reinisch A. J., Dong J.-M. Brubaker C. L., Stelly D. M., Wendel J. F., and A. H. Paterson., (1994). A detailed RFLP map of cotton, Gossypium hirsutum x Gossypium barbadense: chromosome organization and evolution in a disomic polyploid genome," *Genetics,* vol. 138, no. 3, pp. 829–847

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

Said Mohammed and Kassahun Tesfaye (2012). Genetic diversity of Lepidium *sativum* population from Ethiopia using morphological characters and ISSR markers, Addis ababa university, Addis AbabaEthiopia,14(17): pp1461-1470.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: pp406-425.

Saleh Alansi, Mohamed Tarroum, Fahad Al-Qurainy, Salim Khan and Mohammad Nadeem. (2016). Use of ISSR markers to assess the genetic diversity in wild medicinal *Ziziphus spina-christi* (L.) Willd. Collected from different regions of Saudi Arabia, Biotechnology and Biotechnological Equipment, 30:5, pp942-947.

Schlotterer, C. (2004). The evaluation of molecular markers: Just a matter of fashion? *Nat. Rev Genet.* 5:63-69.

Semagn, Kasa, Bjornstad, A. and Ndjiondjop, M. N (2006). An overview of molecular marker Methods for plants. *Afr. J. Biotech*. 5: pp2540-2568.

Shafiei-Astani, B, Ong, AHK, Valdiani, A, Tan, SG, Yong, CSY, Ahmady, F, Alitheen, NB, Ng, WL and Kaur, T. (2015). Molecular genetic variation and structure of Southeast Asian crocodile (*Tomistoma schlegelii*): comparative potentials of SSRs versus ISSRs'. *Gene*, vol. 571, pp. 107–116.

Shen J., Jia X., Ni H., Sun P., Niu S. and Chen X. (2010). AFLP analysis of genetic diversity of Jatropha curcas grown in Hainan, China Trees.

Singh, P., Singh, S., Mishra, S. P. and Bhatia, S. K. (2010). *Molecular Characterization of Genetic Diversity in Jatropha curcas Mo*lecular biology and tree physiology laboratory, Institute of Forest Productivity, India.

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

Song, J. (2005). *Genetic Diversity and Flowering in Clianthus and New Zealand Sophora (Fabaceae). Ph.D. Dissertation, Massey University, Palmerston North, New Zealand.*

Srivastava, S. and Gupta, O. S. (2009). Inter Simple Sequence Repeat profile as a genetic marker system in sugarcane. Sugar Technol. 10: 48-52.

Tadesse Mebrahtu, Tesfaselassie Gebremariam, Asmerom Kidane and Woldeamlak Araia., (2009). Performance of *Vernonia galamensis* as a Potential and Viable Industrial Oil Plant in Eritrea: Yield and Oil Content. African Journal of Biotechnology, February, 8, 4, 635-640.

Tafese Geremew, Mekonnen, Yalemtsehay Mekonnen, E., Majinda, R.T., Bojase-Moleta, G., & Yeboah, S.O. (2018). Antibacterial activity of crude extracts and pure compounds isolated from *Vernonia galamensis* leaves. African Journal of Pharmacy and Pharmacology, 12, 136- 141.

Thomson, M.J., Septiningsih, E.M., Suwardjo, F., Santoso, T.J., Silitonga, T.S., and McCouch, S.R. (2007). Genetic diversity analysis of traditional and improved Indonesian rice (Oryza sativa L.) germplasm using microsatellite markers. *Theor. Appl. Genet*., 114, 559–568.

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.

Van der Vossen H.A.M. and Mkamilo G.S. (Eds). (2007). Plant Resources of Tropical Africa 14. Vegetable Oils. PROTA Foundation, Netherlands. 178-181.

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

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). *DNA Fingerprinting in Plants: Principles, Methods and Applications*. Taylor and Francis Group, USA. 444pp.

Werner F.A., Durstewitz G., Habermann F.A., Thaller G., Kramer W., Kollers S., Buitkamp J., Georges M., Brem G., Mosner J., Fries R. (2004). Detection and characterization of SNPs useful for ID control and parentage testing in major European dairy breeds. Animal Genetics, 35, 44–49.

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 marker technologies for plant improvement. World Journalof Microbiology and Biotechnology; 11(4):438-48.

Wright S. (1943). Isolation by distance. Genetics. Return to ref 80 in article CAS PubMed Pub Med Central Google Scholar. 28:114–38.

Wright, S. (1965). The interpretation of population structure by F- statistics with special regard to systems of mating. *Evolution*, pp.395-420.

Yeh, F.C., Yang, R.C. and Boyle, T. (1999). POPGENE version 1.32: Microsoft Windows–based freeware for population genetic analysis, quick user guide. *Center for International Forestry Research, University of Alberta, Edmonton, Alberta, Canada*, pp.1-29.

8.APPENDICES

Appendix 1. Representative pictures to give further illustration of all practical works beginningfrom sample collection to the end of laboratory session


A


B


C


D


E

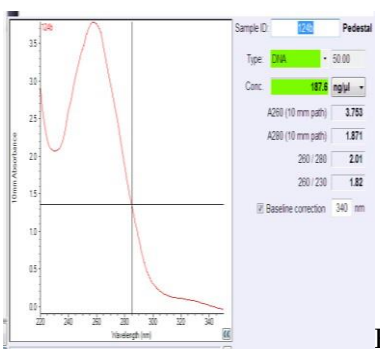

F

A= *V. galamensis* growing, B= sample collection, C=PCR master mix preparation, D= master mix with DNA entering to PCR, E=PCR product loading on agarose gel, F= taking picture from gel Doc

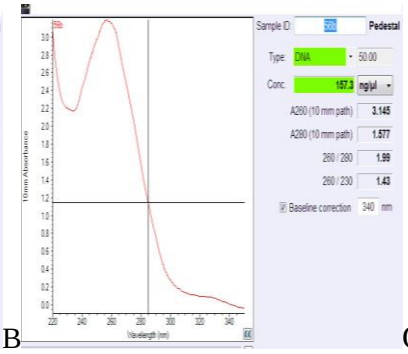
Appendix 2.Quantification and qualification of genomic DNA using Nano drop

The following graphs shows the absorbance patterns of some genomic DNA selected from each populations of *V. galamensis* considered in this study

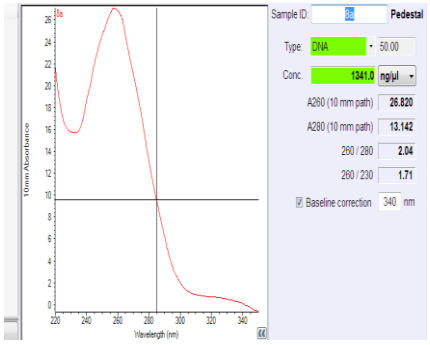
A



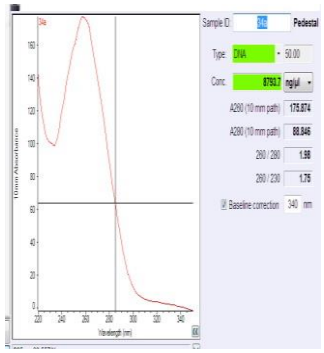
B



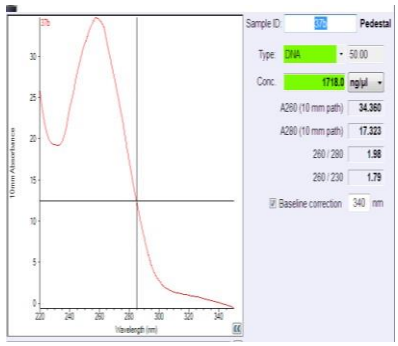
C



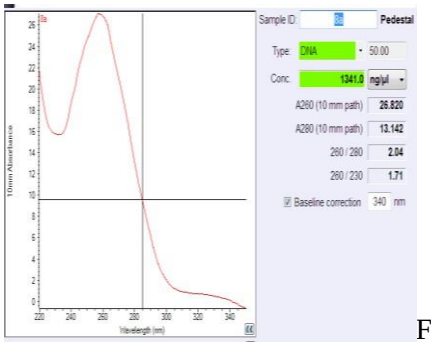
D



E



F



Appendix 3.List of *V. galamensis* samples used in this study

No	Acc.	Region	Zone	Altitude	Longitude	Latitude
1	VgEt-01	Oromia	Borena	2450	E38°48.00	N10°37.00
2	VgEt-02	Oromia	Borena	1380	E37°52.00	N13°20.00
3	VgEt-03	Oromia	Borena	1580	E39°46.40	N10°58.17
4	VgEt-04	Oromia	Borena	1488	E39°37.19	N12°6.42
5	VgEt-05	Oromia	Borena	1580	E39°46.40	N10°58.17
6	VgEt-06	Oromia	Borena	1486	E39°51.00	N09°57.00
7	VgEt-07	Oromia	Borena	1560	E39°46.20	N10°58.33
8	VgEt-08	Oromia	Borena	1490	E38°54.00	N09°54.00
9	VgEt-09	Oromia	East Shoa	2275	E39°44.00	N08°23.00
10	VgEt-010	Oromia	East Shoa	1860	E38°34.00	N07°13.00
11	VgEt-011	Oromia	East Shoa	1860	E38°34.00	N07°13.00
12	VgEt-012	Oromia	East Shoa	1860	E38°34.00	N07°13.00
13	VgEt-013	Oromia	East Shoa	2275	E39°44.00	N08°23.00
14	VgEt-014	Oromia	East Shoa	1930	E38°45.00	N07°21.00
15	VgEt-015	Oromia	East Shoa	1930	E29°42.00	N11°53.32
16	VgEt-16	Oromia	East Hararghe	1900	E36°21.00	N08°21.00
17	VgEt-17	Oromia	East Hararghe	1560	E41°47.30	N09°26.35

18	VgEt-18	Oromia	East Hararghe	2260	E36º39.00	N08º47.00
19	VgEt-19	Oromia	East Hararghe	2020	E41º54.77	N09º27.24
20	VgEt-20	Oromia	East Hararghe	1626	E41º16.55	N09º23.58
21	VgEt-21	Oromia	East Hararghe	1560	E41º34.62	N09º54.85
22	VgEt-22	Oromia	East Hararghe	1560	E41º47.30	N09º26.35
23	VgEt-23	Oromia	East Hararghe	1560	E41º34.62	N09º54.85
24	VgEt-046	Oromia	West Arsi	1899	E39º29.42	N11º53.32
25	VgEt-047	Oromia	West Arsi	1960	E37º22.00	N12º27.00
26	VgEt-048	Oromia	West Arsi	2603	E39º27.14	N09º56.34
27	VgEt-049	Oromia	West Arsi	1560	E39º46.20	N10º58.33
28	VgEt-050	Oromia	West Arsi	1870	E39º15.00	N11º31.00
29	VgEt-051	Oromia	West Arsi	1520	E42º10.10	N09º16.01
30	VgEt-052	Oromia	West Arsi	1640	E41º44.79	N09º29.64
31	VgEt-24	SNNP	Konso	1220	E41º46.38	N09º35.74
32	VgEt-25	SNNP	Konso	1430	E42º06.46	N09º35.83
33	VgEt-26	SNNP	Konso	1900	E42º37.75	N09º29.92
34	VgEt-27	SNNP	Konso	2275	E39º44.00	N08º23.00
35	VgEt-28	SNNP	Konso	1765	E40º58.34	N09º06.78
36	VgEt-29	SNNP	Konso	2275	E39º44.00	N08º23.00

37	VgEt-30	SNNP	Konso	2100	E37°28.00	N08°59.00
38	VgEt-31	SNNP	Derashie	1480	E35°04.00	N08°52.00
39	VgEt-32	SNNP	Derashie	2020	E41°54.77	N09°27.24
40	VgEt-33	SNNP	Derashie	1263	E35°18.08	N07°13.15
41	VgEt-34	SNNP	Derashie	1860	E38°34.00	N07°13.00
42	VgEt-35	SNNP	Derashie	1750	E40°51.00	N09°05.00
43	VgEt-36	SNNP	Derashie	1830	E36°38.00	N09°02.00
44	VgEt-37	SNNP	Derashie	2000	E40°54.31	N09°03.16
45	VgEt-38	Amhara	South Wollo	2450	E38°48.00	N10°37.00
46	VgEt-39	Amhara	South Wollo	1860	E38°34.00	N07°13.00
47	VgEt-40	Amhara	South Wollo	1263	E35°18.08	N07°13.15
48	VgEt-41	Amhara	South Wollo	2275	E39°44.00	N08°23.00
49	VgEt-42	Amhara	South Wollo	1580	E39°46.40	N10°58.17
50	VgEt-43	Amhara	South Wollo	1580	E41°16.55	N09°23.58
51	VgEt-44	Amhara	South Wollo	1560	E39°46.20	N10°58.33
52	VgEt-45	Amhara	South Wollo	1580	E39°46.40	N10°58.17
53	VgEt-53	Amhara	West Gojjam	2270	E37°54.00	N08°48.00
54	VgEt-54	Amhara	West Gojjam	1900	E36°21.00	N08°21.00
55	VgEt-55	Amhara	West Gojjam	1750	E40°5100	N09°05.00

56	VgEt-56	Amhara	West Gojjam	1930	E38°45.00	N07°21.00
57	VgEt-57	Amhara	West Gojjam	1600	E40°50.45	N09°06.30
58	VgEt-58	Amhara	West Gojjam	1430	E40°39.34	N09°11.36
59	VgEt-59	Amhara	West Gojjam	1560	E41°47.30	N09°26.35
60	VgEt-60	Amhara	West Gojjam	2100	E37°54.00	N08°48.00

Appendix 4. A Modified CTAB protocol for plant DNA extraction

1. Introduction

This protocol describes a modified CTAB method for the isolation of DNA from plant material.

A maximum of 24 samples can be isolated at once using this protocol.

2.Materials and Chemicals used in this study

Materials	Chemicals
Geno grinder	Sodium Acetate
Metal beads	Agarose
65°C water bath	Water (sterile)
Microfuge tubes	Chloroform: Iso Amyl Alcohol (24:1)
tips	75 % Ethanol (ice cold)
Microfuge	Absolute Ethanol (ice cold)
	ß-Mercapto-Ethanol
	NaAc

3. Protocol

1. Pour CTAB solution (700 µl per sample) in a 1.5ml-tube and add 0.2 vol % Mercaptoethanol (Use fume hood!). Mercapto-ethanol is stored at 4°C.

2. Aliquot CTAB in 1.5 ml Eppendorf-caps and warm in water bath up to 65°C.

3. Weigh in 100 mg fresh leave material (50mg dry material) per sample. Pulverise thoroughly using a Geno grinder machine.

4. Add 700 µl of warm CTAB solution to the powdered sample (open the caps carefully), dissolve the powder and incubate the sample for 30 minutes at 65°C.

6. Centrifuge for 5 minutes at 15000 rpm.

7. Transfer the supernatant (only clear liquid) in a new Eppendorf-cap. Use blue pipette tips which are cut.

8. Add new CTAB solution (700 µl) to the tissue pellet and stir slightly with a new 1000 µl pipette tip, incubate 30 min at 65°C.Step 6 and 7 are repeated. The same is carried out for a third extraction. Each fraction proceeds with step 9 and is treated separately.

9. Add 600 µl Chloroform to the cap with supernatant and shake carefully a few times upside down. This chloroform step should be carried out immediately.

10. Shake the samples thoroughly by turning inversing the Eppendorf caps for approximately 5 minutes. (Longer incubation is possible)

11. Centrifuge for 5 min at 15000 rpm.

12. Transfer the supernatant (only clear liquid) in a new Eppendorf-cap. Use blue pipette tips which are cut. Work carefully; do not transfer suspended matter (normally the Chloroform Is covered by a thin layer of fine sediment material). Chloroform has to be disposed of in a special waste bottle.

13. Repeat the chloroform extraction (step 9-12) to make sure that all impurities are removed, and then proceed with step 14.

14. Add cooled Iso-propanol (4°C), approximately 2/3 of the solution volume. Shake carefully by inversing the Eppendorf cap. In most cases DNA becomes visible as white threads. Freeze for more than 2h at -20°C.

15. Centrifuge 10 min at 15000 rpm.

16. Aspirate liquid using yellow tips (without touching pellet!). If pellet is solid enough the larger part of the liquid may be poured out. (Alternatively add TE).

17. Add 200 µl Ethanol 70 % to the pellet. Rinse the inner cap surface by turning the cap.

18. Centrifuge for 10 min at 15000 rpm in a cooled centrifuge.

19. Aspirate Ethanol using yellow tips. Dry the DNA-pellet at room temperature. (Usually 15 min are sufficient; after drying no liquid drops are to be seen)

20. Dissolve pellet in 100 µl TE (1x, p.a. grade) and store at 4°C.

21. Add cooled 7.5 M NH4Ac-solution (4°C, half of the solution volume). Mix carefully.

22. Add cool Ethanol 100 % (double of the solution volume). Mix carefully. Freeze for more than 2 hr at -20°C.

23. Centrifuge 30 min at 15000 rpm. Aspirate fluid carefully.

24. Add 200 µl Ethanol 70%. Rinse the inner cap surface by turning the cap.

25. Centrifuge 10 min. at 15000 rpm. Aspirate liquid and dry pellet at room temperature.

Dissolve the pellet in 100 µl TE (1x, p.a. grade)

26. Repeat steps 21 to 24 with 3 M Na-AC-solution (4°C, half the volume) then proceed with Step 27.

27. Centrifuge 10 min. at 15000 rpm. Aspirate liquid and dry pellet at room temperature. Dissolve the pellet in 100 µl TE (1x, p.a. grade) and store at -20^{o}c for further use.

Appendix 5. Output of the Evanno method results.

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	20	-5908.185	0.386992	—	—	—
2	20	-5267.82	1.495115	640.365	166.925	111.64694
3	20	-4794.38	7.859195	473.44	106.475	13.547825
4	20	-4427.415	2.555134	366.965	157.445	61.619074
5	20	-4217.895	101.18086	209.52	42.95357	0.424523
6	28	-4051.32857	58.088673	166.56643	10.82714	0.18639
7	20	-3873.935	35.580055	177.39357	56.53357	1.588912
8	20	-3753.075	53.798189	120.86	8.565	0.159206