GENETIC DIVERSITY OF ETHIOPIAN MUSTARD (B*rassica carinata* A. Braun) USING INTER-SIMPLE SEQUENCE REPEAT MARKERS

MSc THESIS

Andualem Getachew

March 2017

Haramaya University, Haramaya

Genetic Diversity of Ethiopian Mustard (*Brassica carinata* A. Braun) using Inter-Simple Sequence Repeat Markers

A Thesis Submitted to the Department of Biology,

Postgraduate Program Directorate

HARAMAYA UNIVERSITY

In Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE IN BIOTECHNOLOGY

Andualem Getachew

March 2017

Haramaya University, Haramaya

HARAMAYA UNIVERSITY

POSTGRADUATE PROGRAM DIRECTORATE

We hereby certify that we have read and evaluated this Thesis entitled "**Genetic Diversity of Ethiopian Mustard (*Brassica carinata* A. Braun) using Inter Simple Sequence Repeat Markers"** prepared under our guidance by Andualem Getachew. We recommend that it be submitted as fulfilling the thesis requirement.

Tileye Feyissa(PhD) ______________________ __________________

Major advisor Signature Date

Yohannes Petros(PhD) ______________________ __________________

Co-Advisor Signature Date

As members of the Board of Examiners of the MSc Thesis Open Defense, we certified that we have read and evaluated the Thesis prepared by Andualem Getachew and examined the candidate. We recommend that the thesis be accepted as fulfilling the Thesis requirements for the degree of Master of Science in Biotechnology.

____________________ ______________________ ___________________

Chairperson Signature Date

____________________ ______________________ ___________________

Internal Examiner Signature Date

____________________ ______________________ ___________________

External Examiner Signature Date

Final approval and acceptance of the Thesis is contingent upon the submission of its final copy to the Council of Graduate studies (CGS) through the candidate's department or school graduate committee (DGC or SGS

STETMENT OF THE AUTHOR

First, I declare that this thesis is my legitimate work and that all sources of materials used for this thesis have been appropriately acknowledged. This thesis has been submitted in partial fulfillment of the requirements for MSc. Degree at the Haramaya University and is deposited at the University Library to be made available to borrowers under the rules of the Library. I declare that this thesis is not submitted to any other institute anywhere for the award of any academic degree, diploma or certificate.

Brief quotations from this thesis are allowed without special permission provided that accurate acknowledgement of the source is made. Requests for permission for extended quotations from or reproduction of this manuscript in whole or in part may be granted by the Head of the Department of Biology or the Dean of School of graduate studies when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Name: Andualem Getachew Signature: ______________________

Date: 6 April, 2017

School/Department: Biology

BIOGRAPHICAL SKETCH OF THE AUTHOR

The author was born on March 08, 1979 in Addis Ababa of his father Getachew Asefa and his mother Belaynesh Gebremichael. He attended his primary and junior secondary educations at Mekaneyesus Primary School from 1987 to 1997. He then joined Adola Senior Secondary School, Adola Town in 1998 for grades 9 to 12 and completed his high school education in 2002.

He joined Ambo University, the then Ambo College of Agriculture, in 2003 and graduated with Diploma in Biology Laboratory Techniques on July 01, 2004. He also graduated with Bachelor of Arts in Economics and Bachelor of Science in Biology from Haramaya University on November 07, 2009 and July 06, 2013 respectively. He was recruited as Technical Assistant in the Department of Biology of Haramaya University in August 2004 and has been serving the department at positions of technical assistant, graduate assistant and assistant lecturer. In October 2014, he joined the Post Graduate Program of Haramaya University to study MSc in Biotechnology.

ACKNOWLEDGEMENTS

I am highly grateful to my advisers Dr. Tileye Feyissa and Dr. Yohhanes Petros for their critical evaluation of the research work, encouragement, excellent comments, guidance and careful supervision beginning from the research proposal to thesis write-up. The successful accomplishment of this research work would have not been possible without their unreserved guidance and assistance. Thus, I am very much indebted to them for their tireless efforts that enabled me to successfully complete my research.

My special thanks goes to Dr. Tileye Feyissa for allowing me to conduct the laboratory activities in his research laboratory, providing the required equipments and consumables for DNA Extraction, PCR amplification as well as his follow up during the laboratory work.

I thank Mr. Leweye Getie for his kind assistance in reagent preparation, DNA extraction, primer selection and optimization and PCR amplification during the laboratory work. I am also thankful to my friends and all persons who helped me with valuable ideas.

Due thanks is also extended to Ethiopian Biodiversity Institute (EBI) for providing me the Ethiopian mustard seeds used for this study and to Institute of Biotechnology, Addis Ababa University for allowing me green house and laboratories for sowing the seeds and conducting laboratory work of the research.

I would like to extend sincere thanks to Haramaya University and Biology Department for the study leave and provision of fund for my graduate study.

ACRONYMS AND ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
CTAB	Cetyl Trimethyl Ammonium Bromide
ISSR	Inter Simple Sequence Repeat
NTSYS - pc	Numerical Taxonomic SYStem for Personal Computer
OECD	Organization for Economic Co-operation and Development
PCR	Polymerase Chain Reaction
PGR	Plant Genetic Resource
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SSR	Simple Sequence Repeat
TBE	Tris-borate-EDTA
TE	Tris-EDTA
UPGMA	Unweighted Pair Group Method with Arithmetic mean
USDA	United States Department of Agriculture
PCoA	Principal Coordinate Analysis

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF TABLES IN THE APPENDIX

Genetic Diversity of Ethiopian Mustard (*Brassica carinata* A. Braun) using Inter Simple Sequence Repeat Markers

ABSTRACT

Brassica carinata *(2n = 34, BBCC) is an amphidiploid plant arisen as a result of interspecific cross between* Brassica nigra (2n = 16) *and* Brassica oleracea (2n = 18). *It is economically important plant being cultivated for use as vegetable and oilseed. Though* B. carinata *is among important oil crops, study on its genetic variability is lacking and this study is meant to examine the genetic diversity and relationships of* B. carinata *accessions collected from selected* B. carinata *growing areas of Ethiopia. Six ISSR markers were used to analyze the genetic diversity and relationships of 60 accessions of* B. carinata *collected from four geographical areas; Wello, West Shewa, East Wellega and Illu Ababora and assigned to four populations according to the collection areas. A total of 73 bands were generated with an average of 12.17 bands per primer. The percentage polymorphism observed in the populations ranged from 65.75% to 84.93%. Nei's gene diversity (0.34) was highest in accessions collected from Illu Ababora and least (0.26) in accessions collected from Wello. The proportion of the total genetic variation contributed by among population variation (G_{ST}) was 0.227 while 77.3% of the total genetic diversity was attributed to the within population variation. AMOVA analysis revealed higher (75%) genetic variation within the populations and less (25%) genetic variations among the populations. Clustering analysis and dendrogram constructed with UPGMA method resulted in similar clustering pattern. The patterns of clustering in 2D and 3D scatter plots were supportive of the clustering pattern obtained with UPGMA. The result in this study showed that ISSR Markers disclosed high level of genetic diversity. Although ISSR markers enabled detection of high genetic diversity in the 60 accessions used in this study, more analysis should be done with more number of accessions collected from vast geographical areas.*

Key words: AMOVA, *Brassica carinata*, Clustering, Genetic Variation, Polymorphism

1. INTRODUCTION

Brassica species, commonly called as rapeseed-mustard, are diverse plants belonging to family *Brassicaceae* and genus *Brassica*. The genus comprises 41 species among which six species are agronomically and economically important, namely, *B. napus*, *B. rapa*, *B. juncea*, *B. nigra*, *B. oleracea* and *B. carinata* (Tsige *et al*., 2005a). They are cultivated in different areas worldwide with China, India, Canada, Japan and Germany being the major rapeseed-mustard growing countries (Vinu *et al*., 2013). The purposes of cultivation of such plants include oil production, condiments, vegetable and subsistence income generation by small holder farmers. *Brassica* oilseeds provide 14% of the world's edible vegetable oil and are the third most important source of edible oil after soybean and palm (OECD, 2012).

Ethiopia is the center of origin for many cultivated plants including Ethiopian mustard (*Brassica carinata*) (EBI, 2014). *B. carinata* is an amphidiploid plant believed to have arisen about 10 thousand years ago as a result of an interspecific cross between *B. nigra* and *B. oleracea* (OECD, 2012). It is economically important plant in Ethiopia being cultivated for use as vegetable and oilseed next to niger seed (*Guizotia absyynica*) and linseed (*Linum ustatismum L.*) (Tesfaye *et al*., 2014). *B. carinata* has important agronomic features such as drought tolerance and disease resistance that contribute to its adaptation in temperate climate.

Genetic diversity is vital for management and conservation of germplasm collections. Since germplasms are collected from different geographical and ecological locations studying genetic diversity will provide information about the genetic structure of the entire collection which provides ease to properly manage the collections. Genetic diversity study is also crucial to help breeders make selection of individual germplasm based on genetic structure of the collection (Mohammadi and Prasanna, 2003).

Moreover, availability and accessibility of genetic diversity is crucial for successful improvement of crop quality, yield and development of a species. Govindaraj *et al*. (2015) suggested that availability of genetic diversity in crops provides option for breeders to develop new varieties and hybrids and genetic diversity can be assessed through phenotypic and molecular characterization of plant genetic diversity (PGR).

Genetic diversity can be determined through evaluation of morphological, biochemical, and molecular markers. Morphological and biochemical markers are less effective than molecular markers for genetic diversity study. These markers do not provide accurate information on the genetic make-up of the organism due to silent mutation and environmental cues (Nybom, 2015). As a result, the genetic variability at molecular level may escape detection through biochemical and morphological markers. Studying genetic diversity at molecular level reveals the molecular information which can be detected through molecular techniques (Semagn *et al*., 2006; Xu, 2010).

There are several DNA-based markers, each of which has its own advantages and disadvantages. Therefore, in choosing the DNA based marker for a particular purpose, the researcher should consider such factors as time, cost, technical demand, efficiency, reproducibility, abundance and variability (Sarwat, 2012). Inter-simple sequence repeat markers are highly variable and ubiquitously distributed across the genome, achieve higher reproducibility compared to random amplified polymorphic DNA and costs less in terms of time and money compared to amplified fragment length polymorphism (Ng and Tan, 2015; Goswami and Tripathi, 2010).

Though Ethiopia is the origin of *B. carinata* (EBI, 2014), studies on genetic diversity of this plant are limited. Different biochemical markers (Nigussie and Heiko, 2005; Tsige *et al*., 2005b) and morphological markers have been used to study the genetic diversity of *B. carinata* (Zada *et al*., 2013; Adeniji and Aloyce, 2012; Misteru and Yared, 2013). Tsige *et al*. (2005a) employed AFLP to study genetic diversity of *B. carinata* accessions. However, diversity studies of *B. carinata* based on DNA based markers are lacking. Therefore, this study is aimed at employing ISSR markers to analyze the genetic diversity of *B. carinata* accessions collected from different regions of Ethiopia.

Objectives

General Objective

The general objective of this study is to assess the genetic diversity and relationships of Ethiopian mustard (*B. carinata*) accessions collected from different geographical areas of Ethiopia using ISSR markers.

Specific Objectives

The specific objectives of the study are:

- To assess the extent of genetic diversity of Ethiopian mustard accessions collected from different regions of Ethiopia.
- To examine the genetic apportionment among and within populations of the accessions

2. LITERATURE REVIEW

2.1. Origin, Taxonomy and Distribution of *Brassica carinata*

Brassica species are categorized under family Brassicaceae (Cruciferae), genus *Brassica.* According to Rich (1991) and Christopher *et al*. (2005), Brassicaceae family consists of about 350 genera and 3500 species. However, El-Esawi *et al*. (2012) reported that different scientists recorded different numbers regarding the number of species in the genus *Brassica*. *Brassica carinata* belongs to the Tribe Brassiceae (Arias *et al*., 2014).

Brassica carinata is one of the six agronomically and economically important oil seed mustards. Three of these species are monogenetic diploids and the other three are amphidiploids. The former includes *Brassica rapa* (2n = 20, AA), *Brassica nigra* (2n = 16, BB) and *Brassica oleracea* (2n = 18, CC). *Brassica napus* (2n=38, AACC), *Brassica juncea* (2n=36, AABB) and *Brassica carinata* (2n = 34, BBCC) originated as amphidiploids from crosses among pairs of species with lower chromosome numbers like *Brassica nigra* (2n=16, BB), *Brassica oleracea* (2n=18, CC) and *Brassica rapa* (2n = 20, AA) (Arias *et al*., 2014). *Brassica carinata* is one of the "U triangle species" of *Brassica* which reveals the phylogenetic relationship of the diploid and amphidiploid species of genus *Brassica* (Figure 1).

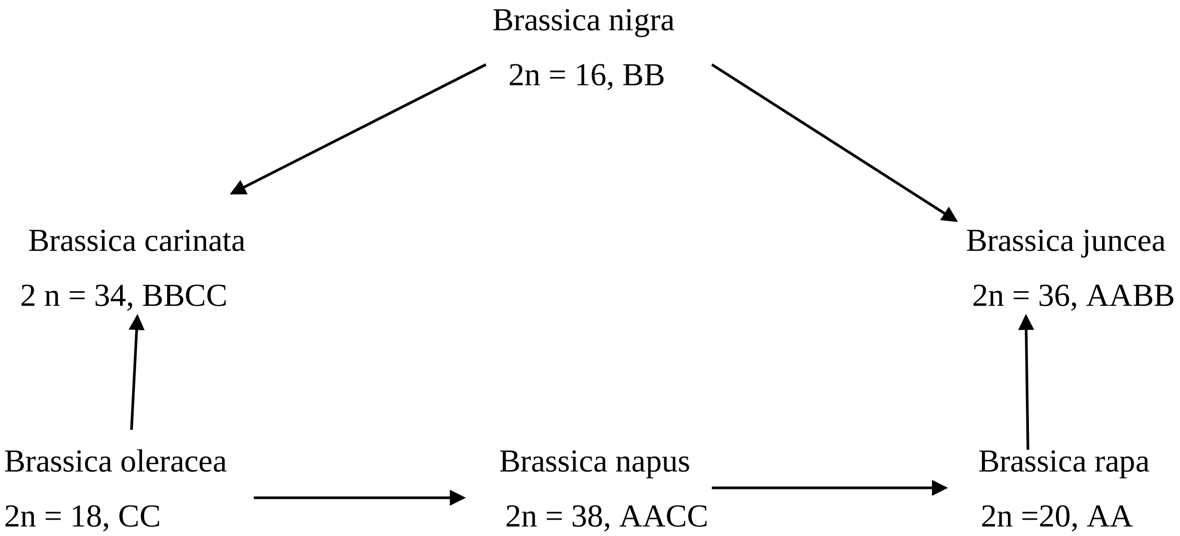

Figure 1: The "triangle of U" showing the genetic relationship of the diploids and amphidiploids species of genus *Brassica*.

Brassica carinata is believed to have arisen about 10 thousand years ago as a result of inter-specific cross between *B. nigra* and *B. oleracea* (OCED, 2012) in highlands of Ethiopian plateau, the adjust countries of east Africa and along the Mediterranean coast (Gomez-Campo and Prakash, 1999). Its distribution has been limited to Ethiopia and neighboring East African countries (OCED, 2012). Recently it has been introduced as an oil crop to India, Spain, Canada, USA, Italy, and Pakistan as a species of commercial interest and field-tested for biofuel production (USDA, 2014).

2.2. Botanical Description of *Brassica carinata*

Brassica carinata is an erect, annual, occasionally biennial plant which can grow up to 2 m tall (Adeniji and Agatha, 2012). It usually grows in branching pattern, glabrous to slightly hairy stem and petiole bases (Mnzava and Schippers, 2007) and has large seed size relative to other *Brassica* species. Leaves are simple, thick, dark green and pinnately lobed. The phyllotaxy is spiral, i.e. leaves are arranged alternately on stems. The crop has strong tap root system. B. *carinata* has appalling harvestability characteristics such as lodging and pod shatter resistance that makes direct cutting possible without shattering seeds.

Brassica carinata bears terminal inflorescence of long raceme, borne on the main stem and branches. Flowers are perfect, with four petals arranged in cross shape. The petals are usually pale-yellow, occasionally cream color and some forms show deep yellow color and are without clasping auricles (Adeniji and Alyoce, 2012). The plant is dioecious, mainly self-pollinating though cross pollination with agents including wind and insects like bees. The seeds are borne in a long narrow pod. The pod, which is the fruit of the plant, consists of two fused carpels separated by pseudo-septum (Figure 2 shows B. carinata plant).

Figure 2: Brassica carinata plant in a field

Source: https://www.linkedin.com/pulse/ethiopian-rape-mustard-abyssinian-brassica-carinata

2.3. Importance of *Brassica carinata*

Brassica carinata is cultivated for a wide range of utilities. It is one of the major oilseeds in Ethiopia though it is mainly cultivated for non-oil use in the country including leaf vegetable, animal feed and subsistent income generation. Leaves, stalks and seeds are edible. The leaves and seeds of the crop are marketed in the local market by farmers to generate subsistence income. Practices like saving and/or obtaining seeds through exchange, barter, gift and local market are major ways to access *B. carinata* seeds in Ethiopia for farmers (Amsalu *et al*., 2014).

Brassica carinata has attracted interest due to its better agronomic performance and commercial purpose in Spain, US, Canada and Italy (Cardone *et al*., 2003; Seepaul *et al*., 2015). The environmental condition of these countries is not favorable for cultivation of *B. napus*, the most common rapeseed in Europe. *B. carinata* is more productive in these areas attributed to its tolerance to adverse conditions and low cropping systems (Cardone *et al*., 2003). The seed of *B. carinata* contains 25-48% oil (Tsige *et al*., 2005) and 25-45 % protein which is comparable to that of pulses (Mnzava and Schippers, 2007).

Brassica carinata is grown as a leafy vegetable in most sub-Sahara African countries though its cultivation is largely limited to production as oilseed in Ethiopia (Mnzava and Schippers, 2007). However, due to the presence of high level of erucic acid and glucosinolate, the oil is not suitable for consumption and considered as non-food oil (Seepaul *et al*., 2015). Therefore, *B. carinata* seed oil is industrially more important for biofuel production than household consumption (Tsige *et al*., 2005). Southeast Farm Press, on August 17, 2013, reported that a jet aircraft was powered by an unblended biofuel produced from 100% *B. carinata* seeds. The biofuel produced from *B. carinata* seed meets the petroleum jet fuel specifications (http://southeastfarmpress.com/markets/carinata-valuable-biofuel-potential-florida-farmers)

2.4. Genetic Diversity Assessment

With increasing human population, climate change and selective breeding and cultivation, natural genetic diversity resources deterioration, erosion and loss seems to be the present and future challenge of human kind. With this in mind, conservation and preservation of genetic diversity are key factors to avoid or minimize loss of plant genetic resources and extinction of valuable plants. Effective conservation, management, and efficient utilization of plant genetic resources (PGR) require adequate knowledge of the existing genetic diversity.

Genetic diversity is the variability in genetic constituent of individuals, groups of individuals or populations. Analysis of the genetic diversity within and among populations provides important information regarding the genetic variability in the population. Improvement of crop genetic resources is dependent on continuous infusions of wild relatives, traditional varieties and the use of modern breeding techniques (Mondini *et al*., 2009). These all processes require assessment of genetic diversity to select plant varieties with desired trait from the range of prevailing individual plants with differing genetic makeup.

Diversity in plant genetic resource provides opportunity for plant breeders to develop new and improved cultivars with desirable characteristics, which include both farmer-preferred traits like yield potential and large seed and breeders preferred traits including pest and disease resistance and photosensitivity (Govindaraj *et al*., 2015).

Study of genetic diversity is the process by which variations among individuals or groups of individuals or population is analyzed with a specific or combinations of methods (Mohammadi and Prasanna, 2003). The extent of genetic variation in a species and its distribution among and within populations is determined by breeding system, historical events, population size, migration between populations and many biotic and abiotic ecological factors (Nybom *et al.*, 2014). Population differentiation (subdivision) is also a fundamental process of evolution and many studies in population genetics including phylogeography and conservation biology benefited from inference of population differentiation (MA *et al*., 2015). Therefore, estimating the level of population genetic diversity and differentiation is of paramount importance.

2.5. Population Genetic Diversity Measurements

Estimation of the amount of genetic diversity among plant genotypes and genetic differentiation in populations have been conducted based on data obtained from morphological characteristics, isoenzymes and DNA finger printing using different DNA-based markers (Rossetto and Rymer , 2013; Nybom *et al.*, 2014). Different parameters including number of polymorphic loci, percentage polymorphic loci, average hetrozygosity (H), also called Nei's gene diversity and population differentiation statistics like coefficient of genetic differentiation F_{ST} (Wright, 1965) and it analogue G_{ST} (Nei, 1973) have been widely used to estimate population genetic diversity and the distribution of genetic variation within and among populations.

The proportion of polymorphic loci and level of hetrozygosity reflects the extent of genetic variability present in the populations while F_{ST} and G_{ST} reflects the apportioning of the total genetic variability among subpopulations (Write, 1965; Nei, 1975). The number and percentage polymorphic loci are crud estimates of polymorphism reside in population and obtained through counting and calculating the proportion of polymorphic loci in the total population. The average hetrozygosity or Nei' gene diversity (H) is calculated by subtracting frequency of homozygosis (gene identity) from 1($H = 1 – J$; J = frequency of homozygosis) at a locus in each population, performing the operation for all loci in subpopulations and then averaging.

F_{ST} is calculated from the within and among population variances with the following formula:

$$F_{ST} = \frac{\sigma_p^2}{\overline{p}(1-\overline{p})}$$

Where σ_p^2 is the variance observed among subpopulations for a given allele and $\overline{p}$ is the average frequency of the allele in the total population (Wright, 1965).

G_{ST} is defined as the proportion of genetic diversity that resides among populations. It is equivalent to F_{ST} when there are only to alleles at a locus, and, in the case of multiple alleles, it is equivalent to the weighted average of F_{ST} for all alleles (Nei, 1973). G_{ST} is calculated from the average total genetic diversity in the pooled populations (H_T) and average diversity within each population (H_S) according to the formula:

$$G_{ST} = \frac{H_T - H_S}{H_T} = \frac{D_{ST}}{H_T}$$

.

Or alternatively as:

$$G_{ST} = 1 - \frac{H_S}{H_T}$$

Where D_{ST} is the average total genetic diversity distribution among population, H_S is calculated as the mean of expected hetrozygosity (He) values over all populations, where He is the expected proportion of heterozygous loci per individual given as:

$H_e = 1 - \Sigma p_i^2$, Where p_i is the frequency of a given allele.

G_{ST} is used to describe the average amount of genetic differentiation observed over multiple loci since the quantities H_S and H_T represent averages over all loci examined. The values of F_{ST} and G_{ST} range from zero to one with low values (close to zero) and high values (close to one) indicating that little and large amount of genetic variation is apportioned among populations respectively. G_{ST} enables estimation of gene flow a characteristic and it can be default measure of population genetic differentiations at present and in the near future (MA *et al*., 2015)

Another approach in genetic diversity studies is cluster analysis. Clustering analysis often performed to evaluate the genetic relationship or genetic distance among individual accessions or germplasms. Unweighted pair-group method based on arithmetic mean (UPGMA) algorism is the most used agglomerative clustering and dendrogram construction methods (Mohammadi and Prasanna, 2003). It is a distance based method and requires calculation of similarity/dissimilarity between genotypes to generate distance data which is used as input data for cluster analysis. Jaccard's similarity coefficient is one of the genetic distance calculation methods for binary data recorded from dominant DNA-based markers like RAPD and ISSR and is often used to calculate genetic similarities between genotypes.

2.6. Markers for Genetic Diversity Assessment

Diversity among organisms arises as a result of variations in DNA sequences and environmental factors. Each individual of a species, with the exception of monozygotic twins, possesses a unique DNA sequence. Variations in DNA nucleotide sequences are caused by mutations resulting from substitution of single nucleotides, insertion or deletion of DNA fragments of various lengths, duplication or inversion of DNA fragments (Russell, 2010).

A marker is the characteristic features by which a particular organism can be distinguished from the range of individuals in the population. Genetic markers are biological features that are determined by allelic forms and can be used as experimental probes or tags to keep track of an individual, a tissue, cell, nucleus, chromosome or gene (Xu, 2010). Markers can be morphological, biochemical or molecular and are routinely used for genetic diversity assessment within and among plant populations (Govindaraj *et al*., 2015; Semagn *et al*., 2006).

2.6.1. Morphological markers

Morphological markers generally represent genetic polymorphisms which are visible as differences in appearance, such as the relative difference in plant height and color, distinct differences in response to abiotic and biotic stresses, and the presence or absence of other specific morphological characteristics (Xu, 2010). These marker traits are often susceptible to phenotypic plasticity. Morphological markers are still having advantage and they are mandatory for distinguishing the adult plants in the field. Morphological markers do not require expensive technologies but large plot of land area for field experiments (Mondini *et al*., 2009).

2.6.2. Biochemical Markers

Biochemical markers are allelic variants of enzymes called isozymes or protein marker that catalyze the same chemical reaction but differ in amino acid sequence and therefore also in the speed taken to travel through an electrophoretic gel (Nybom, 2014).These markers are detected by electrophoresis and specific staining because they differ in molecular weight and mobility in an electric field (Russell, 2010) Since proteins can be separated based on their molecular weight and charge, which varies depending on the amino acids constituents, they can be separated as long as alleles of the same gene produce protein products with different molecular weight and charge (Xu, 2010). Isozymes reflect the products of different alleles rather than different genes.

Protein electrophoresis provides a technique for quick determination of genotypes of many individuals at many loci used to examine genetic variation in hundreds of plant and animal species (Russell, 2010). Isozymes are important to detect diversity at gene level, have simple inheritance pattern and require only small amounts of plant material for detection. However, only a few number of enzyme markers are available and hence confer low resolution power in revealing genetic diversity resided in the genome of organisms (Mondini *et al*., 2009). Isozyme markers are co-dominant in nature.

2.6.3. Molecular Markers

Molecular markers are segments of DNA that serve as identification mark of individuals. Molecular markers or DNA markers reveal neutral segments at the DNA sequence level and, unlike morphological markers, these markers do not show themselves in the phenotype (Jones *et al*., 1997) and are not affected by the environmental cues. DNA-based markers have advantages over other markers because they are neutral and independent of environmental factors, developmental stage and age of the organism (Sarwat, 2012). Molecular markers are usually located in non-coding regions of DNA in a chromosome and generally measure DNA variations which may not affect the phenotype of the plant (Govindaraj *et al*., 2015; Xu, 2010).

DNA markers differ in their efficiency attributed to their different characteristics. In general, ideal DNA marker should reveal high level of genetic polymorphism, show co- dominance, unambiguous to score, distributed on the entire genome; neutral to selection , easy for detection; low cost for marker development and genotyping, reproducible, so that the data can be accumulated and shared between laboratories (Xu, 2010; Mondini *et al*., 2009). Nevertheless, no molecular marker presents all these attributes.

DNA markers are useful tools for genetic diversity and population genetic structure (pattern of genetic similarities and differences of individuals in a population) investigation, phylogenetic studies, gene tagging, genome mapping, breeding, identification and taxonomic classification, evolutionary biology in a wide range of crop species and devise germplasm conservation programs (Sarwat, 2012; Reddy *et al*., 2002). With the development of high throughput technologies, DNA polymorphisms have become the markers of choice for DNA-based surveys of genetic variations in animals and plants. Different DNA marker systems have been used for genetic diversity study in plants.

Those mostly used markers in genetic diversity studies include RFLP (restriction fragment length polymorphism), SSR (simple sequence repeats or microsatellites), RAPD (random amplified polymorphic DNA) or AP-PCR (arbitrarily primed PCR), ISSR (inter-simple sequence repeats) and AFLP (amplified fragment length polymorphism). Based on the techniques employed for their detection, molecular markers are broadly divided into three classes: hybridization based, polymerase chain reaction (PCR) based and DNA chip and sequencing based DNA markers (Gupta *et al*., 1999; Govindaraj *et al*., 2015).

2.6.3.1. Restriction Fragment Length Polymorphism

Restriction fragment length polymorphism (RFLP) is hybridization based DNA marker. It was used in human linkage mapping by Botstein and his coworkers (Botstein *et al*., 1980) for the first time and this pioneered the utilization of DNA polymorphisms as genetic markers. Genetic analysis based on RFLP marker involves digestion of DNA with one or more sequence specific restriction endonuclease, gel electrophoresis, blotting and hybridization with labeled probe and visualization with autoradiography.

Restriction fragment length polymorphism arises due to change in the nucleotide sequences of the restriction sites as a result of mutation events such as indels, inversions and transpositions. A single nucleotide change can create or destroy a restriction site and causes variation in the number of cleavage sites resulting in different restriction fragment length and number. Thus, if there is variation in the restriction site of DNA between individuals, digestion with the same endonuclease results in restriction fragments of different sizes. Restriction fragment length polymorphism markers have been used in plant breeding, gene linkage mapping, molecular epidemiology, diversity study and taxonomy (Gupta *et al*., 1999; Botstein *et al*., 1980; Mikkonen *et* al., 2005; Goswami and Tripathi 2010; Nybom, 2015; Janssen *et al*., 1996)

Restriction fragment length polymorphisms are co-dominant markers that enable detection of both the dominant and recessive alleles of the heterozygous genotype and all the three morphs at a locus (Figure 3). If the base pair mutation is present in one chromosome but not in the other, both fragment bands appear on the gel, and the sample is said to be heterozygous for the marker. Therefore, RFLP markers are more informative than morphological and dominant DNA-based markers because they can make identification of the three genotypes; dominant homozygous, dominant heterozygous and recessive homozygous genotypes possible (Jones *et al*., 1997).

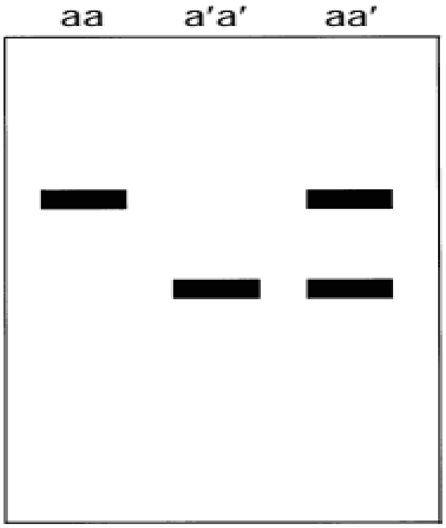

Figure 3: Band pattern of DNA from plants with RFLP genotypes at a locus.
Tracks aa and a’a’ depict bands resulted from homozygous genotype and aa’ from heterozygous genotype. The co-dominance of RFLPs allows for all three genotypes at a locus to be scored. Source: Jones *et al.*, (1997), Markers and mapping: we are all geneticists now

Restriction fragment length polymorphic markers have strong attributes like high reproducibility, co-dominance, no prior sequence information requirement, and relatively easiness to score due to large size difference between fragments. However, there are limitations attached to RFLP markers: requirement of high quantity and quality of DNA, needs development of specific probe libraries for the species, non-amenability for automation, low level of polymorphism and few loci are detected per assay, time consuming, laborious, and expensive and usually requires radioactively labeled probes (Semagn *et al*., 2006).

2.6.3.2. Random Amplified Polymorphic DNA

Random Amplified Polymorphic DNA (RAPD) is the first PCR based method used to amplify DNA fragments from any species without prior knowledge of DNA sequences (William *et al*., 1990). RAPD maker system makes use of a single PCR primer of 10 nucleotides long with arbitrary sequence that binds to its homologous sequences in DNA by chance and enables amplification of several different region of the genome. Typically, in standard RAPD analysis, each primer is a 10-mer oligonucleotide with 60–70% GC content and with no self-complementary ends (Jhang and Shasany, 2012). The difference among the amplified DNA fragments of genomic DNA samples taken from organisms reveals polymorphism in the population of organisms.

RAPD markers do not provide genotypic information if the organism is heterozygous due to their dominant nature and hinder analysis which require knowledge of hetrozygosity. It is not possible to distinguish whether a DNA segment is amplified from a locus that is heterozygous or homozygous with a dominant RAPD marker (Jhang and Shasany, 2012). RAPD marker also lacks reproducibility. However, RAPD marker system is a method of choice for many molecular approaches because it requires no previous sequence knowledge, relatively easy and not expensive techniques. Moreover, multiple loci can be amplified from a single primer and a small amount of DNA is required.

RAPD techniques have useful applications in plant genetic characterization and determine the genetic structure of plant populations (Martín *et al*., 1999), genetic diversity (Bussell, 1999) genotype identification (Goswami and Tripathi, 2010), genotype fingerprinting (Abdelmigid, 2012) and gene mapping (Jhang and Shasany, 2012) to mention some.

2.6.3.3. Amplified Fragment Length Polymorphism

Amplified fragment length polymorphism (AFLP) technique is a PCR based technique and uses restriction endonucleases for DNA fingerprinting in identification of organisms and can be used for DNA of any origin or complexity (Vos *et al*., 1995). It is based on the amplification of restriction fragments from the whole genome, and separation of amplified products by polyacrylamide gel electrophoresis.

Amplified fragment length polymorphism analysis involves the use of multistep processes. The first step in the generation of AFLPs is to double-digest genomic DNA with two restriction enzymes. Next, a specific short DNA sequence, adaptor, is linked to one end of the fragment, and a different adaptor sequence is added to the other end. These sequences, together with the adjacent restriction sites serve as binding sites for PCR primers (Jones *et al*., 1997). AFLP marker system is technically demanding and expensive to set up, but it detects a large number of loci, reveals a great deal of polymorphism and produces high complexity DNA fingerprints (Jones *et al*., 1997)

Like other DNA markers, AFLP has a number of advantages and disadvantages. The advantages include requirement of less amount of DNA, generation of more information associated with high frequency of AFLPs, high reproducibility, and fast and cost effectiveness compared to RFLP (Paun, O. and Schönswetter, P. 2012). Moreover, AFLP generates fingerprints of any DNA regardless of its source, and without prior knowledge of DNA sequence (Mondini *et al*., 2009). However, since it is a dominant marker, AFLP does not allow the differentiation between some heterozygous and homozygous genotypes.

Amplified fragment length polymorphism markers have been applied in gene mapping (Jones *et al*., 1997), DNA fingerprinting (Vos *et al*., 1995), molecular epidemiology (Mikkonen *et al*., 2005), plant breeding (Gupta *et al*., 1999), genetic diversity study (Goswami and Tripathi, 2010), and bacterial taxonomy (Janssen *et al*., 1996). Tsige *et al*., (2005) used AFLP markers to investigate the genetic relationship of 39 *B. carinata* genotypes and suggested that AFLP is a reliable tool and it permits greater insight into the genetic diversity of *B. carinata*.

2.6.3.4. Simple Sequence Repeat

Simple sequence repeats (SSRs) or simple tandem repeats (STRs) or microsatellites are short stretches of DNA sequences usually 1 to 6 base pairs (bps), repeated multiple times in tandem and randomly distributed in eukaryotic genomes both in coding and non-coding regions (Sarwat, 2012). The repeated sequence is often simple, consisting of di, tri or tetra nucleotides. SSRs are the smallest class of simple repetitive DNA sequences and are flanked by unique sequences (Semagn *et al.*, 2006; Xu, 2010). SSRs flanking sequences provide templates for specific primers used to amplify the SSR alleles through PCR amplification using pairs of oligonucleotide primers homologous to those DNA sequences flanking the SSR sequences (Xu, 2010).

Microsatellite islands are unstable regions and experience high rate of mutation than other non- repetitive DNA sequences (Antiqueira, 2013). Their high mutation rate and co-dominant nature render SSRs suitable for estimation of genetic diversity within and among non-distant populations. Hence, closely related species or populations can easily be distinguished by microsatellite-based markers where other molecular marker tools were not proven useful (Sarwat, 2012).

Microsatellites are PCR based DNA markers and are informative in plants because they are ubiquitous throughout the genome allowing a more complete coverage of the genome. The high level of polymorphism of microsatellites relative to RFLPs and RAPDs combined with a high interspersion rate make them an abundant source of genetic markers (Gupta *et al*., 1999). Though SSRs are useful genetic markers due to their genome coverage and high polymorphism, they are costly and require specific primer for each species.

2.6.3.5. Inter-Simple Sequence Repeat

Inter-Simple Sequence Repeat (ISSR) markers are DNA sequences flanked by inverted SSR sequence. ISSR marker technique is a PCR based method that relies on amplification of ISSR sequences flanked by the inverted SSRs, given that ISSRs are located at amplifiable distance. ISSR markers are multilocus and dominant markers. ISSR techniques and RAPD techniques are much similar except that ISSR primers are designed from microsatellite regions and annealing temperatures in the PCR reaction of ISSR are higher in comparison to RAPDs because ISSR primers are longer than primers in RAPDs (Sarwat, 2012). Annealing temperature (Ta) of ISSR primers varies from 45-65°C depending on the GC content of the primer employed (Reddy *et al*., 2002).

ISSR markers permit detection of polymorphisms in inter-microsatellite loci using an ISSR-primer, designed from di-nucleotide, tri-nucleotides, tetra-nucleotide or penta-nucleotide microsatellites to target multiple genomic loci in a single PCR reaction to generate ISSR fragments of different sizes (Abdelmigid, 2012; Reddy *et al*., 2002). The SSRs are used for primer attachment sites to amplify the inter-SSR regions. Mutation in these PCR priming sites and inter-microsatellite sequences are the sources of variation in amplified fragments between individuals as a result of structural variation in the SSRs, loss of primer binding sites and changes in sequences flanked by SSRs. ISSR primers are usually 16 to 25bps long (Abdelmigid, 2012; Ng and Tan, 2015), and amplify inter-SSR sequences in the range 100-3000 bps long (Sarwat, 2012).

Since their first development by Zietkiewicz *et al*. (1994), ISSR markers have been employed in diversity study of a variety of plants such as *Cornus spp.* accessions (Shi *et al*., 2010), *Brassica napus* genotypes (Abdelmigid, 2012; Safari *et al*., 2013), *Brassica juncea* (Yadav and Rana, 2012) genetic diversity in *Barassica* taxa and *Arabidopsis thaliana* (Bornet and Branchard, 2004), *Vigna radiate* (Singh *et al*., 2012), *Poa angustifolia* and *Poa trivialis* accessions (Arslan and Tamkoc, 2011),*Trichosanthes dioica* Roxb. cultivars (Goswami and Tripathi, 2010), species of *Cerrado* (Antiqueira, 2013), *Dioscorea opposite* Thunb cultivars (Zhou *et al*., 2008), *Lupines albus* L. (Oumer *et al*., 2015), *Trifolium steudneri* (Tadesse and Kassahun, 2017).

ISSR markers are technically simple, require no radioactive materials, have good throughput with relatively less cost (Goswami and Tripathi, 2010), fast, informative, detect high level of polymorphism and do not require prior knowledge of DNA sequences (Sarwat, 2012). Although ISSR markers are advantageous with these respects, they are not free from drawbacks. They suffer from low reproducibility and informativeness in revealing genotypes per locus due to their dominant nature compared to co-dominant markers. These problems can be reduced by using high number of loci and leaving out non-reproducible loci (Ng and Tan, 2015).

3. MATERIALS AND METHODS

3.1. Plant Material

Sixty accessions (Appendix 1) of *B. carinata* collected from four geographical areas (Wello, West Shewa, East Wellega and Illu Ababora) were obtained from Ethiopian Biodiversity Institute (EBI) and grouped in to four populations of fifteen individuals each according to the collection areas. The accessions were planted in pots in the green house of Addis Ababa University. The seedlings were transferred to Plant Molecular Laboratory, Institute of Biotechnology, Addis Ababa University (AAU) prior to collection of leaves in order to keep the leaves fresh for DNA extraction. Healthy, young and fresh leaves were collected from 3-weeks-old bulked 5 young plants of each accession.

3.2. DNA Extraction

Genomic DNA is extracted using CetylTrimethyl Ammonium Bromide (CTAB) (2% Cetyltrimethyl ammonium Bromide; 1% polyvinylpyrolidone; 100mMTris, pH= 8; 20mM EDTA; 1.4M NaCl; 0.2 % β-Mercapto-ethanol) protocol based on Borsch *et al*. (2003) from 0.3g of fresh young leaves sampled for each accession. The leaf samples were pulverized in liquid nitrogen using clean, sterilized and chilled mortar and pestle. The powder was immediately transferred into 2ml eppendorf tube, dissolved with 700µl of warm CTAB solution and incubated at 65°C for 45 minutes. The incubated solution was centrifuged at 13000 rpm for 7 minutes and the supernatant was transferred in to 1.5 ml eppendorf tube using cut blue pipette tips. Six hundred microliters of chloroform was added in to the eppendorf tube with supernatant and shaked thoroughly by turning-inversing for 5 minutes and centrifuged at 13000 rpm for 7 minutes (this step was repeated). The supernatant was transferred into new 1.5ml eppendorf tube and 300 µl cooled isopropanol (4°C) was added, shaked carefully by inverting the eppendorf tube, freezed at -20°C for 2 hours and centrifuged at 13000rpm for 13 minutes. The centrifuged solution was aspirated using yellow tips without touching the pellets, 200µl of 70% Ethanol was added and centrifuged at 1300 rpm for 13 minutes in a cool centrifuge. The Ethanol was aspirated using yellow tips and the DNA-pellet was dried at room temperature for 15 minutes. The dried DNA pellets were dissolved in 100µl TE buffer and stored at 4°C over night. Fifty microliters of cooled (4°C) of 7.5M NH_4Ac solution was added and mixed carefully. Three hundred microliters of cooled (4°C) 100% Ethanol was added, mixed carefully and freezed at -20°C for 2 hours. Then centrifuged at 13000rpm for 35 minutes and the liquid was aspirated. Two hundred microliters of 70% Ethanol was added, centrifuged at 13000 rpm for 13 minutes and the liquid was aspirated. The DNA pellet was dried at room temperature for 15 minutes and dissolved in 100µl TE buffer. Fifty microliters of cooled (4°C) of 3M NaAc solution was added and mixed carefully. Three hundred microliters of cooled (4°C) 100% Ethanol was added, mixed carefully and freezed at -20°C for 2 hours. Then centrifuged at 13000rpm for 35 minutes and the liquid was aspirated. Two hundred microliters of 70% Ethanol was added, centrifuged at 13000 rpm for 13 minutes and the liquid was aspirated. The DNA pellet was dried at room temperature for 15 minutes, dissolved in 100µl TE buffer

The DNA pellets dissolved in 100µl TE buffer were gel electrophoresed in 1x TBE buffer on 1 % agarose gel to check the DNA quality. Electrophoresis was run for 80 minutes at constant voltage of 100V. The quality and quantity of the genomic DNA was also verified with NanoDrop (NanoDrop$^{\text{TM}}$2000/2000c) spectrophotometer (Appendix II). The DNA was stored at -20°C for ISSR amplification.

3.3. PCR Amplification and Gel Electrophoresis

Thirteen ISSR-primers (Sigma Aldrich) were screened for the presence of PCR product, polymorphism, reproducibility and clarity of bands and six primers comprising four di-nucleotide, one tri-nucleotide and one penta-nucleotide primers were selected for PCR amplification and further analysis (Table 1).

Table 1: ISSR primers screened and selected for amplification

Primer	Primer Sequence(5’-3’)	Annealing Temperature(ºC)	Repeat motif	Band Formation Pattern
UCB-809	$(AG)_8G$	47	di-nucleotide	Poor
UCB-810	$(GA)_8T$	45	di-nucleotide	poor
UCB-812*	$(GA)_8A$	45	di-nucleotide	Good, polymorphic
UCB-826*	$(AC)_8C$	47	di-nucleotide	Good, polymorphic, reproducible
UCB-828*	$(TG)_8A$	45	di-nucleotide	Good, polymorphic, reproducible
UCB-835	(AG)8TC	49	di-nucleotide	Poor
UCB-836*	$(AG)_8AT$	47	di-nucleotide	Good, polymorphic, reproducible
UCB-844	(CT)8AC	49	di-nucleotide	Poor
UCB-854	(TC)8AG	49	di-nucleotide	Poor
UCB-857	(AC)8TG	49	di-nucleotide	poor
UCB-866*	(CTC)6	55	tri-nucleotide	Good, polymorphic, reproducible
UCB-880	$(GGAGA)_3$	43	penta-nucleotide	Poor
UCB-881*	(GGGTG)3	49	penta-nucleotide	Good, polymorphic, reproducible

*primers selected for further amplification

Three solutions with concentration of 10ng/µl, 25ng/µl and 50ng/µl were prepared based on the NanoDrop measurements by diluting genomic DNA of 8 samples, two from each population, to test the optimum concentration that result in amplification of good and reproducible bands. The DNA solution of concentration 25ng/µl produced relatively better bands than the two concentrations and working solutions of 25ng/µl were prepared from each sample through dilution.

PCR reaction was carried out in 26µl of reaction volume. The master mix (Table 2) was prepared by adding all the PCR reaction components in to a 2ml eppendorf tube except the genomic DNA. Reaction mixture of 24µl was transferred into 0.5ml eppendorf tube and 2µl genomic DNA was added, mixed gently and PCR reaction was carried out in PCR machine.

Table 2: PCR reaction components, concentration and volume

Component	Concentration	Volume(µl)
Sterile H_2O	-	17.6
$MgCl_2$	25mM	3
Reaction buffer(BD)	10x	2.5
Primer	20 $pmol^{-1}$	0.4
dNTP mix	100mM	0.2
Taq polymerase	$5U\mu l^{-1}$	0.3
Genomic DNA	25ng/ µl	2
Total	-	26

PCR amplification was performed with cycling parameters of: initial denaturation at 94°C for 4 minutes, followed by 40 cycles with denaturation at 94°C for 30 seconds, varying annealing temperature (45°C/47°C/49°C/55°C) according to the primer used, extension at 72^0C for 1 minute and 30 seconds and final extension at 72°C for 7 minutes using TC- 412, 96 x 0.2 ml block Thermocycler.

Agarose gel was prepared by boiling 1.67g of agarose powder in 100ml of 1x TBE buffer. The gel was stained by adding 3µl Ethidium bromide (10mg/mL) and poured onto gel casting tray with comb on the leveled bench top after cooling for 2 min. The agarose gel was removed from the tray after 20 minutes when it solidified and placed into electrophoresis tank containing 1x TBE buffer. The amplification products of each sample were loaded onto wells with the DNA ladder and control sample loaded onto the most left and most right wells respectively and electrophoresed for 3 hours at constant voltage of 100V.

The amplification products were visualized and photographed under UV light using Gel documentation system (Biosens SC750) and pictures were saved.

3.4. Band Scoring

The gel electrophoresis results were scored as 1, 0 and "?" for presence, absence and ambiguousness of bands respectively from the gel photograph (Figure 4). Clear and distinct bands were considered for scoring in order to minimize stochastic and personal errors. The data were recorded in MS-excel spread sheet as binary matrix for further statistical analysis.

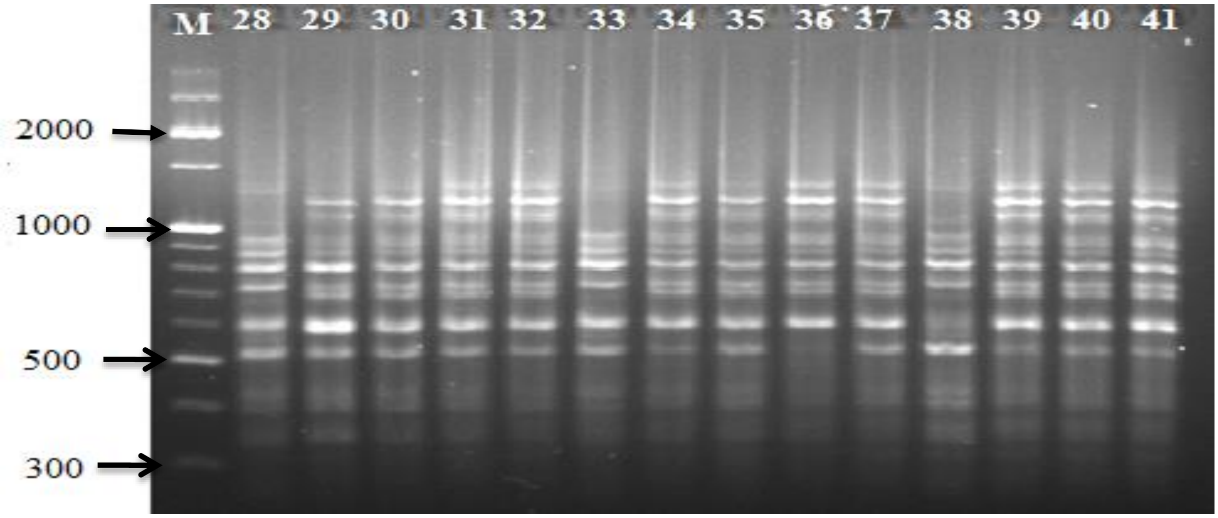

Figure 4: ISSR bands produced with primer 836
Numbers 28-41 represent samples IL_2-IL_{14}; IL= Illu Ababora

3.5. Data Analysis

3.5.1. Genetic Diversity Analysis

POPEGENE version 1.31 (Yeh *et al*., 1997) was used to calculate population genetic diversity and differenciation measures like number of polymorphic loci (NPL), percentage polymorphic loci (PPL), hetrozygosity in subpopulation (H_S), total hetrozygosity in pooled populations (H_T), Nei's gene diversity ($H = 1-\sum Pi^2$, P_i = probability of i^{th} locus), coefficient of genetic differentiation ($G_{ST} = H_T - H_S/H_T$),gene flow rate ($Nm = 0.5(1 - G_{ST})/G_{ST}$) and Shannon Diversity Index ($I = -\Sigma p_i \log_2 p_i$, P_i = probability of i^{th} locus).

3.5.2 Genetic Differentiation Analysis

Among and within populations genetic variations and the coefficient of genetic differentiation (F_{ST}) were calculated using Analysis of Molecular Variance (AMOVA) with GenAlex 6.502 software (Peakall and Smouse, 2012). F_{ST} is calculated from estimated within and among population variances as:

$$F_{ST} = AP / (WP + AP) = AP / TOT$$

Where,
AP = Est. variance among populations
WP = Est. Variance within populations and
TOT = Estimated total variance

3.5.3. Cluster Analysis

NTSYS- pc version 2.02 (Rohlf, 2000) software was used to generate the Unweighted Pair Group Method with Arithmetic mean (UPGMA) dendrogram to analyze and compare the individual accessions based on Jaccard's similarity coefficient(J) calculated with the formula:

$$J_{ij} = a_{ij}/(a_{ij} + b_{i0} + c_{0j})$$

Where, a_{ij} = the number of loci shared by both "i" and "j" individuals
b_{i0} = the number of loci where individual "i" has a band, but individual "j" does not
c_{0j} = the number of loci where individual "j" has a band, but "i" does not

To further examine the patterns of variation among individual samples, principal coordinate analysis (PCoA) was performed based on Jaccard's coefficient (Jaccard, 1908). The two dimensional and three dimensional scatter plots were plotted with GenAlex 6.502 software (Peakall and Smouse, 2012) and STATISTICA version 5.5 software (Hammer *et al*., 2001; Statistica Soft, Inc. 2001) respectively.

4. RESULTS AND DISCUSSION

4.1. Polymorphism Detected with each Primer

The results of the number of bands scored, number of polymorphic bands, percent of polymorphic bans, Nei's Gene diversity and Shannon's diversity index are presented in table 3. A total of 73 bands were generated by all the six primers of which 69 (94.52%) bands were found to be polymorphic. The average number of bands per primer was 12.17 with the size of the bands ranging from 200 to 900bp. High percentage (100%) of polymorphic bands was observed in primers UCB-812, UCB-826, UCB-828 and UCB-866. The least percentage (76.92%) of polymorphic bands was observed in primer UCB-836. The highest gene diversity (0.4376) and the least gene diversity (0.2705) were revealed with primer UCB-826 and primer UCB-836, respectively. Shannon's index was also highest in primer UCB-826 while it was the least in primer UCB-836. The level of polymorphism, Nei's gene diversity and Shannon's diversity index indicate the presence of high level of diversity in the 60 *B. carinata* accessions examined.

Table 3: Primers, Number of Bands, Number of Polymorphic Bands, Percent of Polymorphic Bands, Nei's Diversity and Shannon's Diversity Index

Primer	NB	NPL	PPL	H	I
UCB-812	12	12	100	0.3724	0.5473
UCB-826	13	13	100	0.4376	0.6264
UCB-828	11	11	100	0.3429	0.5108
UCB-836	13	10	76.92	0.2705	0.4057
UCB-866	11	11	100	0.4223	0.6069
UCB-881	13	12	92.31	0.4016	0.5733
Total	73	69	94.52		

NB = number of bands, NPL = number of polymorphic loci, PPL= percent of polymorphic loci, H = Nei's gene diversity, I = Shannon's diversity index,

4.2. Genetic Diversity and Differentiation

The observed percentage polymorphism ranged from 65.75% to 84.93% among the four populations (Table 4). The highest number of polymorphic loci (62, 84.93%) was revealed in

Illu Ababora population while the least number of polymorphic loci (48, 65.75 %) was detected in population of the accessions collected from Wello. Nei's gene diversity (H) value was highest in Illu Ababora population (0.34) and least in Wello (0.26). The same level (0.28) of Nei's gene diversity is revealed in both West Shewa and East Wellega populations. The observed total polymorphism and Shannon's diversity index (I), 94.52% and 0.5443 respectively, indicate the presence of high genetic diversity in *B. carinata* accessions considered in this study (Table 5).

Table 4: Polymorphism and genetic diversity of populations

S.No.	Population	NPL	PP	H + SD	I + SD
1.	West Shewa	52	71.23%	0.28 + 0.21	0.40 + 0.29
2.	East Wellega	49	67.12%	0.28 + 0.22	0.40 + 0.30
3.	Illu Ababora	62	84.93	0.34 + 0.17	0.50 + 0.24
4.	Wello	48	65.75%	0.26+0.21	0.38+0.29

NPL= number of polymorphic loci, PP = percentage of polymorphic loci, H = Nei's genetic diversity, I = Shannon's diversity index, SD = Standard deviation

Table 5: Genetic diversity and differentiation among populations

	West Shewa	East Wellega	Illu Ababora	Wello	Mean	Total
Na	1.7127	1.6712	1.8493	1.6575	1.7226	1.9452
Ne	1.4945	1.5099	1.6049	1.4568	1.5165	1.6719
P	71.23	67.12	84.93	65.75	72.26	94.52
H	0.2774	0.2783	0.3420	0.2598	0.2894	0.3741
I	0.4047	0.4016	0.4981	0.3806	0.4213	0.5443
H_T						0.3744
H_S						0.2893
G_{ST}						0.2271
Nm						1.7013

Na = number of observed alleles; Ne = number of expected alleles; p = percentage polymorphism; H = Nei's gene diversity; I = Shannon's information index; H_T = average total genetic diversity; H_S = average genetic diversity within populations; G_{ST} = genetic differentiation among populations; Nm = gene flow among populations

The total genetic diversity (H_T) and within populations genetic diversity (H_S) were 0.3744 and 0.2893, respectively (Table 5). The proportion of genetic variation of the total genetic variation contributed by interpopulation genetic variations (G_{ST}) was 0.227 and 77.3% of genetic variation was attributed to the within population genetic variations. The gene flow (Nm) value among populations is 1.7013.

Genetic diversity Studies of *B. carinata* accessions based on AFLP markers reported high level of genetic diversity. Jiang *et al.* (2007) evaluated the genetic diversity of 110 *B. carinata* accessions with 36 pairs of AFLP markers and reported wide range of genetic variability in most accessions with genetic similarity varying from 0.46 to 0.88.Tsige *et al. (*2005) used six AFLP primers pairs to detect the genetic diversity of 39 *B. carinata* accessions and reported high level of genetic variation. Among 278 AFLP bands produced, 189 were polymorphic revealing 68% of polymorphism in the AFLP fragments. Though the level of polymorphism detected in these studies is not identical to polymorphism detected in this study, results showed high genetic diversity in *B. carinata* accessions.

Different levels of genetic diversity were observed among and between the populations of the 60 *B. carinata* accessions. Factors such as mating types and migration may be the causes that might have led the in genetic similarity, divergence and differentiation. High genetic diversity detected in *B. carinata* is likely to be attributable to its mating system and pollen and seed dispersal. The gene flow value (Nm= 1.7013) showed the occurrence of moderate gene flow among populations. Gene flow occurs due to the dispersal of germplasm by human activity and other dispersal agents like insects and wind. In Ethiopia, it is customary to cultivate *B. carinata* at small scale level around resident areas and its distribution for cultivation is through seed exchange and marketing. *B. carinata* seed is marketed for household uses like clay plate smoothening (Nigussie and Heiko, 2001; Amsalu *et al*., 2014) which may contribute to its gene flow.

4.3. Analysis of Molecular Variance

Analysis of molecular variance (AMOVA) of the four populations without grouping showed that most of the variation (75%) is due to genetic variation within the populations, and the remaining 25% variation is as a result of variation among the populations. The coefficient of genetic differentiation (F_{ST}) value was 0.253 which implied that there is moderate differentiation between the four populations (Table 6).

Table 6: Analysis of molecular variance of the four populations

Source of variation	df	SS	MS	Est.Var.	PVC	F_{ST}	p
Among populations	3	175.300	58.433	3.254	25%	0.253	p >0.001
Within populations	56	539.267	9.630	9. 630	75%		p >0.001
Total	59	714.567		12.883	100%		

df = degree of freedom, SS = sum of squares, MS = mean sum of squares, Est.Var= estimated variance, PVC = percentage of variance, Fst= coefficient of genetic differentiation

Among and within species genetic variation distribution is strongly linked to life-history traits, particularly dispersal and reproductive mode (Hamrick *et al*., 1991). Analysis of molecular variance showed less genetic diversity among population and higher genetic diversity within population of the 60 *B. carinata* accessions. The G_{ST} (F_{ST}) values imply moderate gene flow among the populations and it can be the possible reason for such pattern of genetic diversity apportion.

The rate of gene flow among populations is affected by geographical distance between populations and mating type which causes flow gene packed in pollen. In this study, some of the accessions collected from neighboring areas (West Shewa and East Wellega) showed to be intermingled while forming clusters which indicates that gene flow may have been occurred between these areas. Jiang *et al.* (2007) evaluated genetic variability of 104 Ethiopian origin *B. carinata* using AFLP markers and detected no significant relationships between accessions originated from the same region and that suggested could indicate gene flow between accessions from Ethiopia. High genetic variation observed in accessions within populations examined in this study implies that there is less genetic relationship between the accessions which may also show the existence of gene flow.

Within populations genetic variation tends to be higher where there is gene flow due to the influx of genotypes and mixing of alleles though crossing with individuals of the receiving population (Templeton, 2006). Mixing of accessions in some groups in cluster formation shows migration of genotypes, crossing and exchange of alleles between mating plants.

Genetic diversity within and among populations can be influenced by the mating type the plant experiences (Wu *et al*., 2015). *B. carinata* is self-pollinating plant; however, pollination involves such pollinating agents as insect and wind which make cross pollination to occur. Hence, mixed mating type can be the reason for distribution of high proportion of genetic diversity within the four populations of the 60 accessions used in this study.

4.4 . Clustering Analysis

Population and individual based cluster analysis was performed with UPGMA method based on Jaccard's similarity coefficient. The dendrogram generated for the populations grouped the four populations separately with some admixture of accessions collected from different areas. Accessions population of Wello, West Shewa and East Shewa are genetically closely related while Ill Aubabora population tends to be distant from the other populations (Figure 5). Pair wise Jaccard similarity coefficient calculated for the four populations ranges from 0.5337 to 0.6500. The highest similarity coefficient (0.6500) was observed between West Shewa and Wello populations while the least similarity coefficient (0.5337) was recorded between West Shewa and Illu Ababora populations (Table 7).

Table 7: Jaccard's similarly coefficient of the four populations

	WS	EW	IL	W
WS	1.000			
EW	0.6454	1		
IL	0.5337	0.5740	1	
W	0.6500	0.6252	0.5575	1

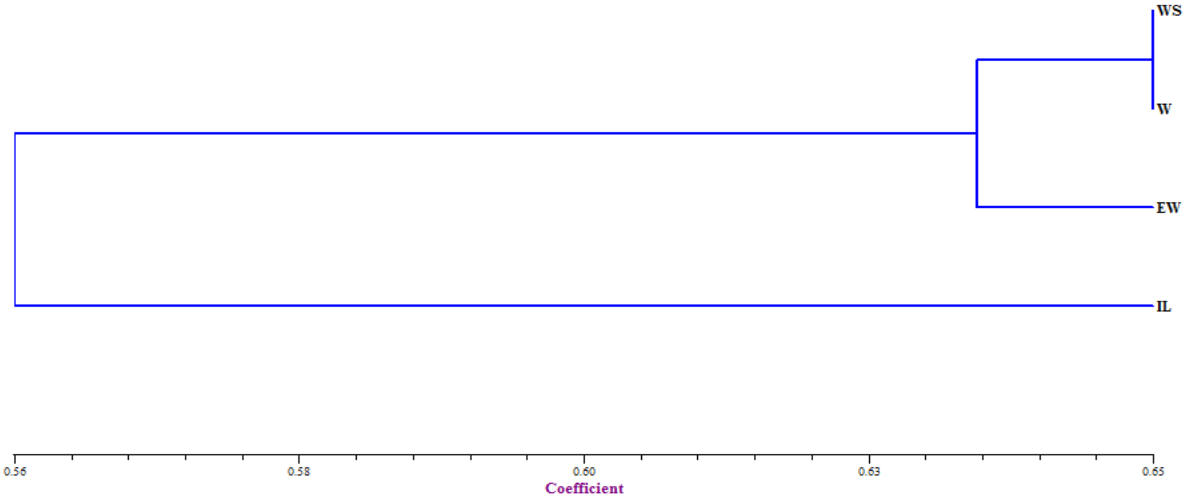

Figure 5: Dendrogram of the four populations constructed with UPGMA method based on Jaccard's similarly coefficient.
WS = West Shewa,W = Wello, EW = East Wellega, IL = Ill Ababora

The dendrogram constructed for the 60 accessions using UPGMA method clustered the accessions in to five major clusters (I, II, III, IV and V) with one out group accession (IL3) at similarity coefficient of 0.58 (Figure 6). Cluster I comprises eight accessions all collected from Illu Ababora. Cluster II includes six accessions, three accessions collected from Illu Ababora and the other three are collected from East Wellega. Clusters III and IV include three accessions from Illu Ababora and two accessions from West Shewa respectively. Cluster V is further divided in to two sub clusters (V1 and V2) in which cluster V2 is again divided in to two sub clusters (V2A and V2B).Cluster V1 consists of four accessions from Wello. Cluster V2A consists of 16 accessions with one singleton accession (WS3) of which two of them are from East Wellega (EW1 and EW2).Among accessions grouped in V2A cluster 98% similarity was observed between EW8 and EW9 accessions which may indicate that these accessions are duplicates or are progenies produced with cross between closely related parents. Cluster V2B consists of 19 accessions (nine and ten accession from Wello and East Wellega respectively).

The clustering pattern of the accessions showed that all accessions, except five accessions from East Wellega, were clustered together and may exhibit some level of similarity. The five accessions from East Wellega (EW_1, EW_2, EW_{13}, EW_{14}, and EW15) intermingled with West Shewa and Illu Ababora accessions. The two accessions (EW1 and EW2) which are intermingled with accessions from West Shewa may imply there is short geographic range genetic flow since East Wellega and West Shewa are geographically close to each other and admixture of accessions from East Wellega with Illu Ababora indicates long geographic range gene flow.

The pattern of clustering formed in both clustering method showed clustering of accessions collected from the same locality in to the same group though some accessions seen to be intermingled. Accessions which have genetic similarity cluster together and the clustering pattern seen in cluster analysis imply that there is considerable amount of genetic similarity within the accessions collected from the same geographical area. As revealed by the dendrogram, Illu Ababora populations were distant from other populations which may be due to the geographical distance between the collection areas.

Clustering of accessions of *B. carinata* genotypes in to the same cluster according to their geographical origin is also reported by Yared (2011) in his study on genetic variability, correlation and path analysis of 39 Ethiopian mustard genotypes. The same clustering pattern was also reported by Fekadu *et al*. (2016) based on agronomic trait of 49 *B. carinata* genotypes. AFLP based study of genetic diversity of 39 accessions of Ethiopian mustard conducted by Tsege *et al*.(200) depicted that the accessions form 7 distinct clusters at a cut of 0.5 with dissimilarity ranging from 0.00 to 0.60 though the geographical correlation of clustering is not clear.

The clustering pattern observed in this study and previous studies indicate that accessions collected from the same area are relatively genetically related and accessions with different clusters exhibit less genetic relationship. Therefore, selection of parents from different clusters for breeding can help to develop hybrids with desired traits and avoid inbreeding depression that may arise from inbreeding of genetically similar or related parents.

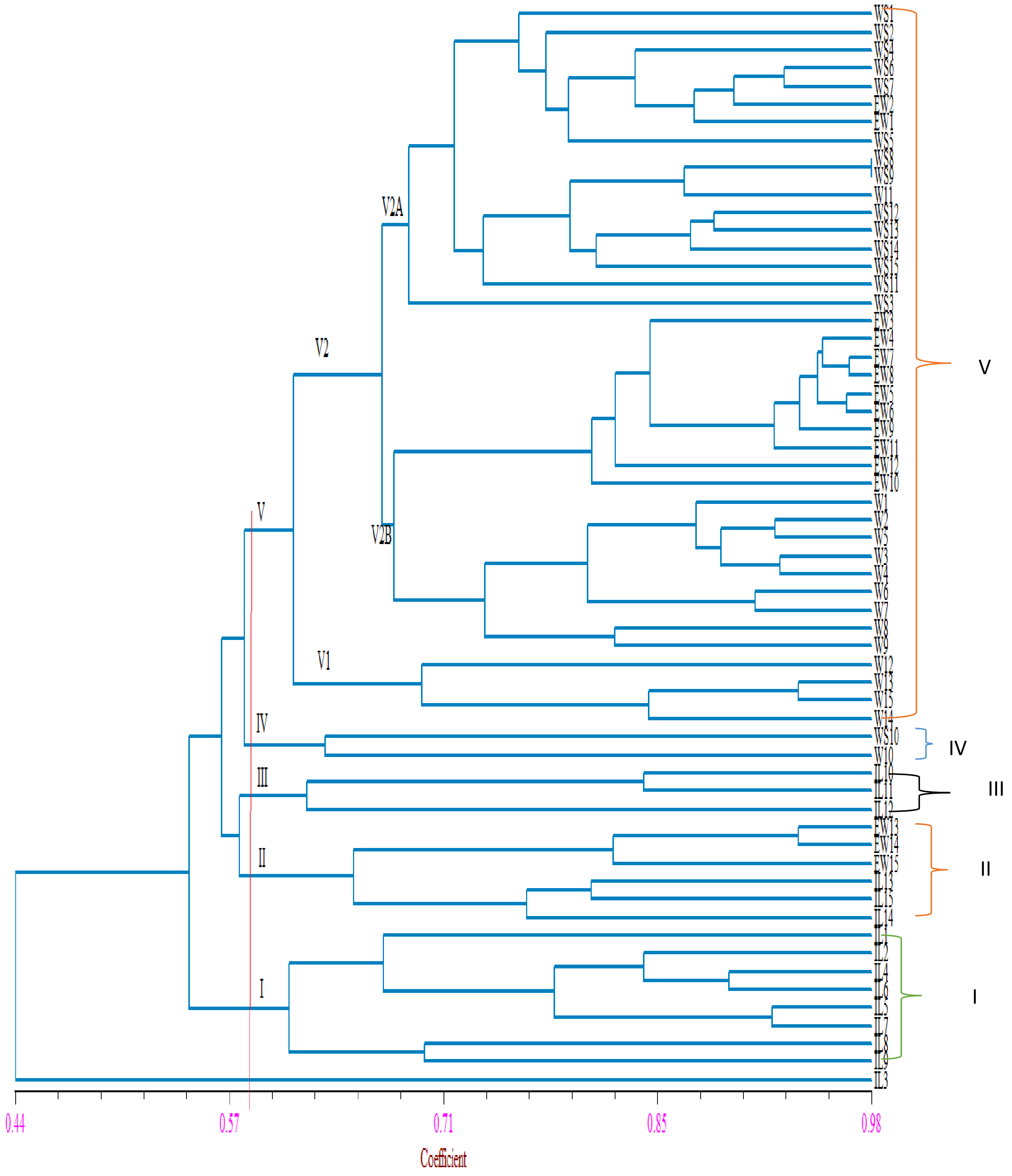

Figure 6: Dendrogram of 60 accessions of *B. carinata* constructed using UPGMA method based on Jaccar's similarity coefficient.
WS = West Shewa,W = Wello EW = East Wellega, IL = Illu Ababora

4.5 . Principal Coordinate Analyses

Principal coordinate analysis (PCoA) is performed using GenAlex 6.5 and Statistica version 5.5 for 2D and 3D scatter plot respectively. The Eigen values of the first three principal components used for principal coordinate analysis were 13.557, 12.056 and 9.289 respectively. Of the total variation, 37.52% was contributed by these principal components. The 2D scatter plot was congruent with the clustering pattern of accessions observed in UPGMA method. Accessions EW13 and EW14 are grouped with accessions from IL and accessions EW1 and EW2 grouped with accessions from WS in both clustering methods and in 2D scatter plot (Figure 7).

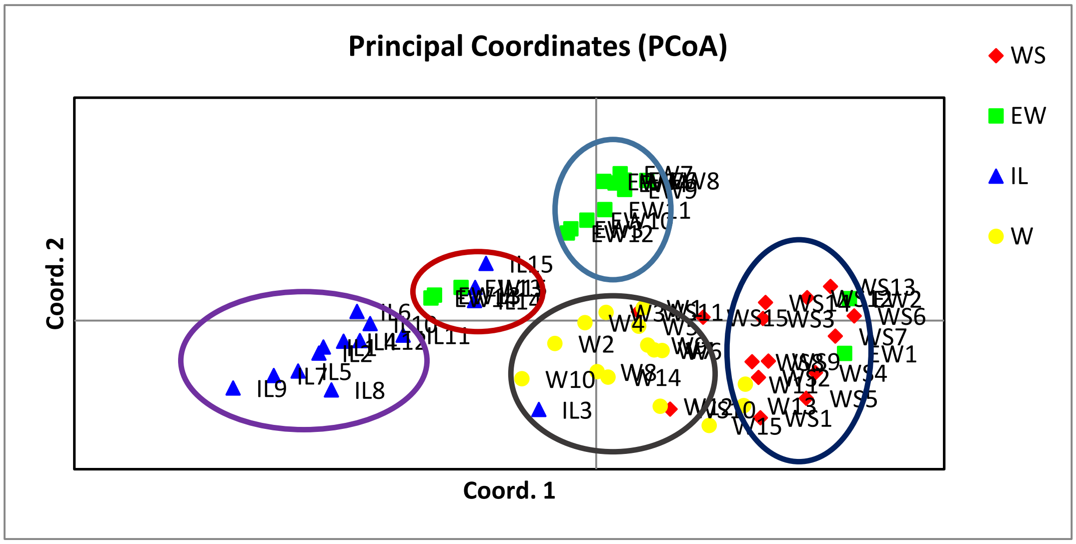

Figure 7: Two dimensional Scatter plot of 60 *B. carinata* accessions
WS = West Shewa, EW = East WEllega, IL= Illu Ababora, W = Wello

The 3D scatter plot (Figure 8) also showed that accessions from same areas plot together like the clustering pattern observed using the UPGMA clustering method. The accessions appear to cluster according to their collection area in 2D and 3D scatter plots as it has been seen grouped in the UPGMA clustering method.

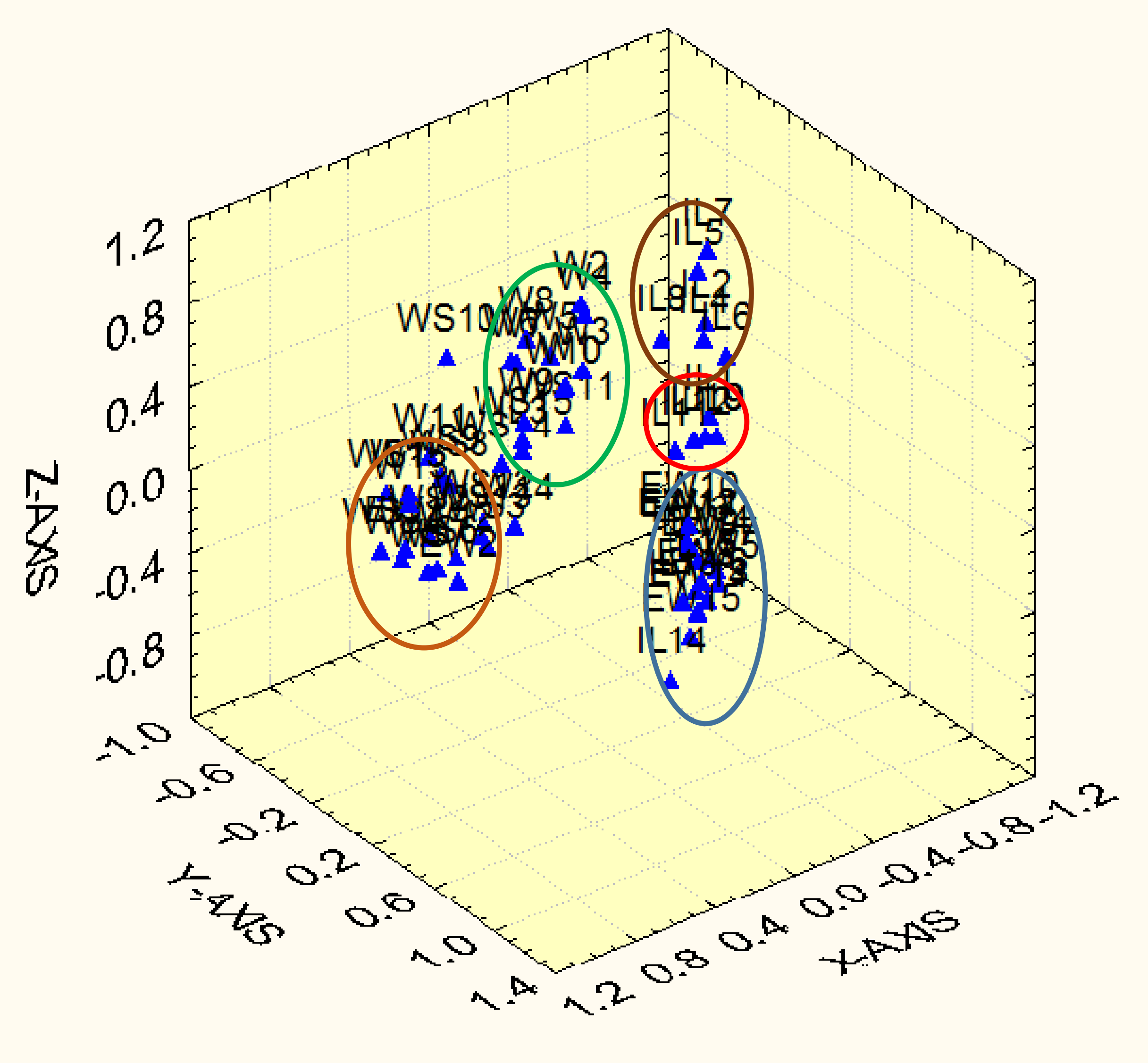

Figure 8: Three dimensional Scatter plot of 60 *B. carinata* accessions
WS = West Shewa, EW = East WEllega, IL= Illu Ababora, W = Wello

PCoA is usually done in order to compliment the clustering methods in order to evaluate the reliability of the methods. Clustering patterns in the 2D and 3D scatter plots depicted congruence with clusters generated using the UPGMA and NJ methods though the 3D plot showed admixture of accessions and some distortion. In the PCoA scatter plots, the distance between the accessions is related to their genetic distance. More distant accessions appear apart from each other and less distant accessions lie close to each other in the scatter plots.

Most of the accessions collected from East Wellega formed separate cluster with individual accession lie close to each other in both 2D and 3D scatter plots. Accessions collected from

West Shewa, Illu Ababora, and Wello were formed groups of accessions more dispersed from each other relative to that of accessions collected from East Wellega. The clustering pattern observed in both 2D and 3D scatter plots seems to show less relatedness of individual accessions collected from West Shewa, Illu Ababora, and Wello and more relatedness of accessions collected from East Wellega. Generally, the 2D and 3D principal coordinate analysis revealed the existence of more genetic relationship between accessions collected from East Wellega than between accessions collected from West Shewa, Illu Ababora, and Wello.

5. SUMMARY

Sixty accessions of *Brassica carinata* were examined for genetic diversity using ISSR markers to analyze the genetic relationship of the accessions. The polymorphic amplicons produced by each primer ranged from 76.92% to 100% and the ISSR primers were informative with respect to revealing polymorphism. A total of 73 loci were observed among which 69(94.52%) were polymorphic. The genetic diversity detected with each primer ranged from 0.2705 and 0.4376.

The percentage polymorphism observed in the populations range from 65.75% to 84.93%. Nei's gene diversity was highest in Illubabor population (0.34) and least in Wello (0.26). The proportion of total genetic variation contributed by among population differences (G_{ST}) was 0.2271 and 77.3% of genetic variation was attributed to the within population differences. The gene flow value (Nm = 1.7013) showed moderate gene flow among populations.

AMOVA analysis revealed higher genetic variation within the populations. The distribution of genetic variation detected within and among population was 75% and 25% respectively. This value is consistent with Gst value which implies that less proportion of the total genetic variation reside among populations.

Jaccared's similarity coefficient and dendrogram showed that Illubabor population was genetically distant from other populations. Clustering analysis and dendrogram constructed with both UPGMA and NJ method resulted in similar clustering pattern. Accessions collected from the same area tend to cluster together with some accessions mingled with accessions collected from other areas. The pattern of clustering in 2D and 3D scatter plots was supportive of the clustering pattern obtained with UPGMA and NJ methods. The first three principal components used for PCoA analysis accounts of 37.52% of the total variation and deemed informative. PCoA analysis revealed more genetic relatedness between East Wellega accessions than between accessions collected from the other three areas considered in this study.

6. CONCLUSIONS

ISSR markers used in this study enabled detection of high level of genetic diversity in the 60 *B. carinata* accessions examined. High genetic diversity resided among the 60 accessions examined with Gst value showing that relatively high proportion of the total genetic variation is distributed within the populations and the gene flow value indicated moderate gene flow among populations. The proportion of genetic variation distributed among population was less and it can be due to moderate gene flow that resulted in mixing of genetic elements from migrating individuals with individuals of receiving population which contributes to the less variation among population.

The existence of genetic variability in the accessions is important to select genetically distant accessions for the purpose of breeding and management of gene banks. As the result in this study indicated, there seems to be more genetic distance between accessions collected from Illu Ababora and west Shewa than between accessions collected from other areas considered in this study. Therefore, hybridization between accessions selected from these areas may result in progeny with hybrid vigor. In other hands, accessions WS8 and WS9 exhibited 98% similarity which may indicate that these accessions might be duplicates and one of the accessions can be kept in the gene bank to avoid redundancy and maintain genetic diversity of accessions in the collection.

The result obtained in this study can give an insight regarding the genetic variability of *B. carinata* accessions maintained at EBI gene bank and can be used as input for breeders and researchers interested to conduct further investigation on *B. carinata* germplasm collections originated from Ethiopia.

7. RECOMMENDATIONS

Knowledge of genetic structure is crucial in germplasm conservation, restoration, development of cultivars etc. ISSR markers can be used to manage germplasm collection, identification and genetic analysis to avoid redundancy in the accessions collection and help breeders select germplasms based on genetic diversity. To have more comprehensive information on genetic diversity and structure it is important to include adequate sample size from more inclusive geographical array and also conduct further investigation using other type of markers specially co-dominant markers like SSR. Therefore, studies that include large number of accessions collected from vast geographical areas of Ethiopia should be conducted in order to have more comprehensive information on genetic structure of the whole collection.

8. REFERENCES

Abdelmigid, M. H. 2012. Efficiency of random amplified polymorphic DNA (RAPD) and inter-simple sequence repeats (ISSR) markers for genotype fingerprinting and genetic diversity studies in canola (*Brassica napus*). *African Journal of Biotechnology,* 11(24): 6409-6419.

Adeniji, O.T. and Aloyce, A. A. 2012. Floral and seed variability patterns among Ethiopian Mustard (*B. carinata* A. Braun) of East Africa. *Tropicultura,* 30(3):133-140.

Amsalu Ayana, Victor Afari-Sefa, Bezabih Emana, Fekadu F. Dinssa, Tesfaye Balemi and Milkessa Temesgen. 2014. Analysis of vegetable seed systems and implications for vegetable development in the humid tropics of Ethiopia. *International Journal of Agriculture and Forestry*, 4(4): 325-337.

Antiqueira, L. M. O. R. 2013. Application of microsatellite molecular markers in studies of genetic diversity and conservation of plant species of *Cerrado*. *Journal of Plant Sciences.*1 (1):1-5. doi: 10.11648/j.jps.20130101.11

Arias, T., Beilsteein, A. M., Tang, M., Mckain R. M. and Pires C.J. 2014. Diversification times among Brassica (Brassicaceae) crops suggest hybrid formation after 20 million years of divergence. *American journal of Botany*.101 (1):86-91.

Arslan, E. and Tamkoc, A. 2011. The application of ISSR-PCR to determine the genetic relationship and genetic diversity between narrow leaved bluegrass (*Poa angustifolia*) and rough bluegrass (*Poa trivialis*) accessions. *Turk J Biol,* 35: 415-423.

Bornet, B. and Branchard, M. 2004. Use of ISSR fingerprints to detect microsatellites and genetic diversity in several related Brassica taxa and Arabidopsis thaliana. *Hereditas,* 140: 245-248.

Botstein, D., White, L. R., Skolnick, M. and Davis, W. R. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32:314-331.

Bussell, J.D. (1999). The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). *Molecular Ecology,* 8: 775–789. doi:10.1046/j.1365-294X.1999.00627.x

Cardone, M., Mazzoncini, M., Menini S., Rocco, V., Senatore, A., Seggiani, M. and Vitolo,S. 2003. *Brassica carinata* as alternative oil crop for production of biodiesel in Italy: agronomic evaluation,fuel production by trans-esterification and characterization. *Biomass and Energy*, 25: 623-636.

Christopher, G.L., Andrew, J.R., Geraldine, A.C.L., Clare, J.H., Jacqueline, B., Gary, B., German, C.S. and David E. 2005. Brassica ASTRA: An integrated database for Brassica genomic research. Nucleic Acids Res., 1(33): 656-659.

EBI (Ethiopian Biodiversity Institute), 2014. Ethiopia's Fifth National Report to the Convention on Biological Diversity. EBI, Addis Ababa, Ethiopia.

El Esawi, M., Bourke, P., Germaine, K. and Malone, R. 2012. Assessment of morphological variation in Irish *Brassica oleracea* species. *Journal of Agricultural Science,* 4 (10): 20 - 34 doi: 10.5539/jas.

Excoffier, L., Laval, G. and Schneider, S. 2005.Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1: 47-50.

Fekadu Amsalu Worekeneh, Sentayehu Alamerew Kebede and Bulecha Woyessa 2016. Multivariate analysis of genetic divergent of Ethiopian Mustard (*Brassica carinata* A. Braun) Land races in Seed Yield and Its Related Component Traits. Acad. Res. J. Agri. Sci. Res. 4(5): 197-205.

Gomez-Campo and Prakash, S.1999. Origin and Domestication. Pp. 33-58. In: Omez-Campo C.(ed.), Biology of brassica cenospecies, Elsevier Science B.V., Amsterdam, the Netherlands.

Goswami, S. and Tripathi, V. 2010. The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among *Trichosanthes dioica* Roxb. Cultivars. *International Journal of Biodiversity and Conservation,* 2(12): 405-413.

Govindaraj, M., Vetriventhan, M. and Srinivasan, M. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. Genetics Research International, Hindawi Publishing Corporation.http://dx.doi.org/10.1155/2015/431487.

Gupta, P.K., Varshney, R. K., Sharma, P. C. and Ramesh, B. 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding*, 118:369-390, Blackwell Wissenschafts-Verlag, Berlin.

Hammer O., Harper, D., and Ryan P. 2001.PAST: Paleontological statistics software package for education and data analysis. Paleontol. Electronic, 4:1-9.

Hamrick J. L., Godt, M. J. W., Murawski, D. A. and Loveless, M. D. 1991. Correlations between species traits and allozyme diversity: implications for conservation biology. In Genetics and conservation of rare plants, D. A. Falk and K. E. Holsinger (eds). New York: Oxford University Press.

Jaccard, P. 1908. Nouvellesrescherches Sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44:223-270.

Janssen, P., Coopman, R., Huys, G., and Swings, J.1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology***,** *142:1881*-1893.

Jhang, T. and Shasany, A.K. 2012. Random amplified marker technique for plants rich in polyphenols. pp 61-74. *In*: Nikolaus J. Sucher et al. (eds.), *Plant DNA Fingerprinting and Barcoding: Methods and Protocols*, Methods in Molecular Biology, vol. 862, DOI 10.1007/978-1-61779-609-8_9, SpringerScience+Business Media.

Jiang, Y., Tian, E., Li, R., Chen, L. and Meng, J. 2007. Genetic diversity of Brassica carinata with emphasis on the interspecific crossability with *B. rapa. Plant Breeding*, 126: 487- 491.

Jones N., Ougham, H. and Thomas, H. 1997. Markers and mapping: we are all geneticists now, *New Phytologist*. , 137: 165-177.

MA, L., JI, Y-J. and ZHANG, D-X. 2015. Statistical measures of genetic differentiation of populations: Rationales, history and current states. *Current Zoology*, 61 (5): 886–897.

Martín, C., González-benito, M. E. and Iriondo, J. M. 1999.The use of genetic markers in the identification and characterization of three recently discovered populations of a threatened plant species. *Molecular Ecology*, 8: S31–S40.

Mikkonen, T., Koort, J.M.K., Bjo¨rkroth, K.J., and Sukura, A.2005. Testing of amplified fragment length polymorphism (AFLP) technique as a tool for molecular epidemiology of Trichinella native. *Veterinary Parasitology*.132, 19–22.

Misteru Tesfaye Woldeyohannes and Yared Semahegn Belete. 2013. Genetic variability of Ethiopian Mustard (*Brassica carinata* A. Braun) accessions based on some morphological characters. *International Journal of Plant Breeding and Genetics,* 7(3):169-175.

Mnzava, N.A. and Schippers, R.R. 2007. *Brassica carinata* A. Braun. [Internet] Record from PROTA4U. van der Vossen, H.A.M. & Mkamilo, G.S. (eds). PROTA (Plant Resources of Tropical Africa / Ressourcesvégétales de l'Afriquetropicale), Wageningen, Netherlands. <http://www.prota4u.org/search.asp>.

Mohammadi, S.A. and Prasanna B.M. 2003. Genetic Diversity Analysis-Statistical Tools and Considerations. *Crop Science*, 43:1235-1248.

Mondini L., Noorani A. and Pagnotta, A. M. 2009. Assessing plant genetic diversity by molecular tools. www.mdpi.com/journal/diversity.

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

Nei, M. 1975. Molecular Population Genetics and Evolution. North-Holland publishing Company, Amsterdam, Netherlands.

Ng, W.L. and Tan, S.G 2015. Inter simple sequence repeat (ISSR) markers: are we do it right? *ASM science journal*, 9(1):30-39.

Nigussie Alemayehu and Heiko Becker. 2005. Quantitative genetic analysis of total glucosinolate, oil and protien contents in Ethiopian Musterd (*Brassica carinata A. Braun*). *SINET: Ethiopian Journal of Science*, 28(2):141–150.

Negussie Alemayehu and Heiko Becker (2001). Variation and inheritance of erucic acid content in *Brassica carinata* germplasm collections from Ethiopia. *Plant Breeding*, 120:331-335.

Nybom, H., Weising, K. and Rotter, B. 2014. DNA fingerprinting in botany: past, present, future Investigative Genetics, 5:1-35. http://www.investigativegenetics.com/content/5/1/1

OECD (Organization for Economic Co-operation and Development). 2012. Consensus Document on the Biology of the Brassica Crops (Brassica spp.). OECD, Paris, France.

Oumer Abdie, yohannes petros,Kassahun Tesfaye,Abel Teshome and Endeshaw Bekele.2017. Inter simple sequence repeat (ISSR) Analyisis of Ethiopian White Lupine (*Lupinus albus* L.). *African journal of Biotechnology*, 14(18):1552–1563.DOI:10.5897/AJB2014.1479.

Paun, O. and Schönswetter, P. 2012. Amplified fragment length polymorphism: An invaluable fingerprinting technique for genomic, transcriptomic and epigenetic studies. PP 75- 87. In: Nikolaus J. Sucher *et al (eds.), Plant DNA Fingerprinting and Barcoding: Methods and Protocols*, Methods in Molecular Biology, Vol.862, DOI 10.1007/978-1-61779-609-8-7, Springer+Business Media.

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

Peakall, R. and Smouse P.E. (2012) GenAlEx 6.5: Genetic Analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics*, 28: 2537-2539.

Rakow, G. 2004. Rakow, G. (2004), "Species Origin and Economic Importance of *Brassica*", *Biotechnology in Agric. and Forestry*, 54, pp. 6-7, Springer-Verlag Berlin Heidelberg.

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

Rich, T.C.G. (1991). *Crucifers of Great Britain and Ireland*, B.S.B.I. Handbook, No.6., Botanical Society of the British Isles, London, pp. 336.

Rohlf,F,J. 2000. NTSYSpc: Numerical Taxonomy and Multivariate Analysis System. Version 2.02, Exerter Software, Setauket, New York.

Rossetto M. and Rymer P. D., 2013. Applications of molecular markers in plant conservation. pp. 81-96. In: Henry, J.R. (ed), Molecular Markers in plants.Jhon wiley & sons,Inc., Chichester,West Sussex, England.

Russell, J. P. 2010. *Genetics: A Molecular Approach, 3rd edition*. Pearson Education Inc., San Francisco, USA.

Safari, S., Mehrabi, A-A., and Safari, Z. 2013. Efficiency of RAPD and ISSR markers in assessment of genetic diversity in *Brassica napus* genotypes. *Intl J Agri Crop Sci.*, 5 (3), 273-279.

Sarwat, M. 2012. ISSR: A Reliable and cost-effective technique for detection of DNA polymorphism. pp. 103-122. In: Nikolaus J. Sucher et al. (eds.), *Plant DNA fingerprinting and barcoding: Methods and Protocols,* Methods in Molecular Biology, vol. 862, and DOI 10.1007/978-1-61779-609-8_9, SpringerScience + Business Media.

Seepaul, R., Bliss, C. M., Wright, D. L., Marois, J. J., Leon, R., Dufault, N., George, S. and Olson S. M. 2015. Carinata, the Jet fuel cover crop: 2015 Production manual for the Southern United States.

Semagn, K., Bjørnstad, Å. And Ndjiondjop, M. N. 2006.An overview of molecular marker methods for plants, *African Journal of Biotechnology*, 5 (25):2540-2568.

Shi, A., Kantartzi, S., Mmbaga, M. and Chen, P. 2010. Development of ISS-PCR markers for diversity study in dogwood (*Cornus spp*.). *Agriculture and Biology Journal of North America*, 1(3): 189-194.

Singh, S., Reddy K. S. and Jawali, N. 2012. Genetic diversity analyses of Mung bean (*Vigna radiata* [L]. Wilczek) by ISSR. *International Journal of Plant Breeding,* 6 (2), 73-83.

Southeast Farm Press, August 17, 2013. http://southeastfarmpress.com/markets/carinata-valuable-biofuel-potential-florida-farmers.

Statistica Soft, 2001. Statistical Data Analysis System. Statistica Soft, Inc., USA.

Studier, J.A., Keppler, K.J. 1988. A note on the neighbor joining algorithm of Saitou and Nei. Mol. Biol. Evol. 5:729-731.

Tadesse Abate and Kassahun Tesfaye. 2017. Genetic Diversity Study of Steudneri Clover (*Trifolium steudneri*) Accessions of Ethiopia Using Inter Simple Sequence Repeat (ISSR) Markers. *Advances in Life Science and Technology*, 54:72-81.

Templeton, A. R. 2006.Population Genetics and Microevolutionary Theory. John Wiley & Sons, Inc., New Jersey, USA.

Tesfaye Walle, Adugna Wakjira and Tsige Genet. 2014. Genotypic variability, Heritability and Genetic advance in Ethiopian mustard (B*rassica carinata* A. Braun) Genotypes at Northwestern Ethiopia. *Journal of Plant Breeding and Genetics*, 02 (03): 109-114.

Tsige Genet, Viljoen, C.D. and Labuschagne, M.T. 2005a. Genetic analysis of Ethiopian mustard genotypes using amplified fragment length polymorphism (AFLP) markers. *African Journal of Biotechnology*, 4 (9):891 - 897.

Tsige Genet, Labuschagne, M.T. and Hugo, A. 2005b.Genetic relationships among Ethiopian mustard genotypes based on oil content and fatty acid composition. *African Journal of Biotechnology*, 4 (11):1256-1268.

USDA (United States Department of Agriculture). 2014. Weed Risk Assessment of *Brassica carinata* A. Braun (Brassicaceae) – Ethiopian mustard. Animal and Plant Health Inspection Service.

Vinu, V., Naveen, S., Sujata, V., Devendra, K. Y. and Sushil, K.,Sugandh, N, Sripad, R. B. and Kumble, V. P. 2013. Assessment of genetic diversity in *Brassica juncea* (*Brassicaceae*) genotypes using phenotypic differences and SSR markers. Int. J. Trop. Biol, 61 (4): 1919-1934

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23(21): 4407-4414.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18(65):31–35.

Write, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, 19:395-420.

Wu, Z-h., Shi, J., Xi, M-l., Jiang, F-x., Deng, M-w., and Dayanandan, S. (2015) Inter-Simple Sequence Repeat Data Reveals High Genetic Diversity in Wild Populations of the Narrowly Distributed Endemic Lilium regale in the Minjiang River Valley of China. *PLoS ONE* 10(3): e0118831. doi:10.1371/journal. pone.0118831

Xu, Y. 2010. Molecular Plant Breeding. International Maize and Wheat improvement Center, Mexico.

Yared Semahegn. 2011. Genetic Variability, Correlation and Path Analysis studies in Ethiopian Mustard (*Brassica carinata* A. Brun). *International Journal of Plant Breeding and Genetics*, (5) 4:328-338.

Yadav, M. and Rana, J. S. 2012. ISSR Markers Assisted Studies on Genetic Diversity in *Brassica juncea*. International Journal of Pharma and Bio Sciences, 3(2): ISSN 0975-6299.

Yeh, Francis C., Yang, R-C., Boyle, T.B.J., YE, Z-H., and Mao, J.X.1997.POPGENE, The user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

Zada, M., Zakir, N., Rabbani A. M. and Shinwari Z. K. 2013.Assessment of genetic variation in Ethiopian mmustard (*Brassica carinata* A. Braun) germplasm using multivariate techniques. Pakistan Jornal of Biotechnology,45 (S1): 583-593.

Zhou, Y., Zhou, C., Yao, H., Liu, Y. and Rongtao, T. 2008. Application of ISSR markers in detection of genetic variation among Chinese yam (*Dioscorea opposite* Thunb) cultivars. *Life Science Journal,* 5(4):6-12.

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

APPENDIX I: Table 1: List of *Brassica carinata* accessions, collection areas, and altitude of collection areas

S.NO	Acc. Number	Region	Zone	Woreda/District	Latitude	Longtude	Altitude	Coll.Date
1.	21283	Amara	Debub Wello	Debresina	10°45'00"N	38°35'00"E		28/11/1981
2.	21284	Amara	Debub Wello	Debresina	10°47'00"N	38°37'00"E		28/11/1981
3.	21285	Amara	Debub Wello	Debresina	10°47'00"N	38°37'00"E		28/11/1981
4.	21310	Amara	Debub Wello	Debresina	-	-		28/11/1981
5.	21311	Oromiya	Mirab Shewa	Ambo	09°00'00"N	37°45'00"E	1880.00	15/11/1982
6.	21312	Oromiya	Mirab Shewa	Ambo	08°57'00"N	37°45'00"E	2140.00	15/01/1982
7.	21313	Oromiya	Mirab Shewa	Ambo	08°57'00"N	37°45'00"E	2140.00	15/01/1982
8.	21314	Oromiya	Mirab Shewa	Tikur	08°48'00"N	37°40'00"E	2410.00	15/01/1982
9.	21315	Oromiya	Mirab Shewa	Tikur	08°48'00"N	37°40'00"E	2410.00	15/01/1982
10.	21316	Oromiya	Mirab Shewa	Tikur	08°48'00"N	37°39'00"E	2430.00	15/01/1982
11.	21317	Oromiya	Mirab Shewa	Cheliya	09°04'00"N	36°36'00"E	2620.00	16/01/1982
12.	21318	Oromiya	Misrak Wellega	Nekemet	09°06'00"N	37°00'00"E	1780.00	17/01/1982
13.	21319	Oromiya	Misrak Wellega	Sibu Sire	09°03'00"N	36°52'00"E	1740.00	17/01/1982
14.	21320	Oromiya	Misrak Wellega	Sibu Sire	09°06'00"N	36°52'00"E	1740.00	17/01/1982

S.NO	Acc. Number	Region	Zone	Woreda/District	Latitude	Longtude	Altitude	Coll.Date
15.	21321	Oromiya	Misrak Wellega	Sibu Sire	09º03'00"N	36º44'00"E	1770.00	17/01/1982
16.	21322	Oromiya	Misrak Wellega	Guto Wayu	09º03'00"N	36º42'00"E	1800.00	17/01/1982
17.	21323	Oromiya	Misrak Wellega	Guto Wayu	09º02'00-N	36º41'00"E	1990.00	17/01/1982
18.	21324	Oromiya	Misrak Wellega	Guto Wayu	09º02'00"N	36º29'00"E	2070.00	18/01/1982
19.	21325	Oromiya	Misrak Wellega	Diga Leka	08º55'00"N	36º28'00"E	2160.00	18/01/1982
20.	21326	Oromiya	Misrak Wellega	Nunu Kumba	08º47'00"N	36º39'00"E	2260.00	18/01/1982
21.	21327	Oromiya	Misrak Wellega	Jimma Arjo	08º45'00"N	36º38'00"E	2080.00	18/01/1982
22.	21328	Oromiya	Misrak Wellega	Jimma Arjo	08º46'00"N	36º35'00"E	2280.00	18/01/1982
23.	21330	Oromiya	Misrak Wellega	Guto Wayu	09º01'00"N	36º25'00"E	2110.00	19/01/1982
24	200403	Oromiya	Misrak Wellega	Jimma Horo	-	-	2330.00	12/09/1981
25	200404	Oromiya	Misrak Wellega	Jimma Horo	-	-	2330.00	12/09/1981
26	200405	Oromiya	Misrak Wellega	Jimma Horo	-	-	2330.00	10/09/1981
27.	21341	Oromiya	Illubabor	Bure	08º16'00"N	35º07'00"E	1680.00	23/01/1982
28.	21343	Oromiya	Illubabor	Bure	-	-		23/01/1982
29.	21344	Oromiya	Illubabor	Bure	-	-		23/01/1982

S.NO	Acc. Number	Region	Zone	Woreda/District	Latitude	Longtude	Altitude	Coll.Date
30.	21345	Oromiya	Illubabor	Ale	08°10'00"N	35°33'00"E	1760.00	24/01/1982
31.	21346	Oromiya	Illubabor	Ale	08°12'00"N	35°34'00"E	1670.00	24/01/1982
32.	21347	Oromiya	Illubabor	Ale	-	-		24/01/1982
33.	21348	Oromiya	Illubabor	Darimu	08°32'00"N	35°40'00"E	1740.00	24/01/1982
34.	21350	Oromiya	Illubabor	Yayu	08°20'00"N	35°36'00"E	1590.00	24/01/1982
35.	21351	Oromiya	Illubabor	Yayu	08°19'00"N	35°37'00"E	1650.00	25/01/1982
36.	21352	Oromiya	Illubabor	Yayu	-	-	1760.00	25/01/1982
37.	21353	Oromiya	Illubabor	Yayu	08°26'00"N	36°20'00"E	1880.00	25/01/1982
38.	21354	Oromiya	Illubabor	Bedele	08°21'00"N	36°21'00"E	1900.00	25/01/1982
39.	21355	Oromiya	Illubabor	Gechi	08°19'00"N	36°26'00"E	1970.00	26/01/1982
40.	21356	Oromiya	Illubabor	Gechi	08°12'00"N	36°27'00"E	2100.00	26/01/1982
41.	21357	Oromiya	Illubabor	Dedesa	07°58'00"N	36°29'00"E	1820.00	26/01/1982
42.	202478	Amara	Semen Wello	Habru	-	-	2000.00	23/11/1982
43.	202479	Amara	Semen Wello	Habru	-	-	1700.00	23/11/1982

S.NO	Acc. Number	Region	Zone	Woreda/District	Latitude	Longtude	Altitude	Coll.Date
44.	202480	Amara	Semen Wello	Habru	-	-	1580.00	23/11/1982
45.	202481	Amara	Semen Wello	Habru	-	-	1570.00	23/11/1982
46.	202482	Amara	Semen Wello	Guba Lafto	-	-	1690.00	23/11/1982
47.	202483	Amara	Semen Wello	Guba Lafto	-	-	1980.00	24/11/1982
48.	202484	Amara	Semen Wello	Guba Lafto	-	-	1850.00	24/11/1982
49.	202485	Amara	Semen Wello	Guba Lafto	-	-	1860.00	24/11/1982
50.	202486	Amara	Semen Wello	Guba Lafto	-	-	1675.00	24/11/1982
51.	202487	Amara	Semen Wello	Guba Lafto	-	-	1640.00	24/11/1982
52.	202489	Amara	Semen Wello	Guba Lafto	-	-	1600	25/11/1982
53.	203224	Oromiya	Mirab Shewa	Adda Berga	-	-	2620.00	23/12/1982
54.	203225	Oromiya	Mirab Shewa	Adda Berga	-	-	2200.00	23/12/1982
55.	203226	Oromiya	Mirab Shewa	Adda Berga	-	-	2560.00	23/12/1982
56.	203227	Oromiya	Mirab Shewa	Alem Gena	-	-	2080.00	24/12/1982
57.	203228	Oromiya	Mirab Shewa	Alem Gena	-	-	2200.00	24/12/1982
58.	203229	Oromiya	Mirab Shewa	Alem Gena	-	-	2400.00	24/12/1982

S.NO	Acc. Number	Region	Zone	Woreda/District	Latitude	Longtude	Altitude	Coll.Date
59.	203230	Oromiya	Mirab Shewa	Alem Gena	-	-	2280.00	24/12/1981
60.	203231	Oromiya	Mirab Shewa	Alem Gena	-	-	2200.00	24/12/1982

APPENDIX II: Table 2: DNA concentrations as measured by NanoDrop

S.No.	Sample ID	Nucleic Acid Conc.(ng/µl)	A260	A280	260/280	260/230	Factor
1.	W1	1726	34.52	16.213	2.13	2.24	50
2.	W2	1726.6	34.532	16.168	2.14	2.18	50
3.	W3	1725.5	34.51	16.14	2.14	2.19	50
4.	W4	2063.2	41.263	19.422	2.12	2.27	50
5.	W5	2443	48.86	23.043	2.12	2.17	50
6.	W6	1614.9	32.298	14.993	2.15	2.19	50
7.	W7	4794.9	95.897	49.911	1.92	2.02	50
8.	W8	1139.6	22.791	10.801	2.11	2.1	50
9.	W9	2186.5	43.73	20.757	2.11	2.1	50
10.	W10	2464.4	49.288	23.535	2.09	2.13	50
11.	W11	2327.1	46.542	22.178	2.1	2.18	50
12.	W12	2103.3	42.066	22.281	1.89	1.71	50
13.	W13	1252.2	25.043	11.937	2.1	1.86	50
14.	W14	617.4	12.348	5.776	2.14	1.92	50
15.	W15	822.5	16.45	7.591	2.17	1.96	50
16.	EW1	750.9	15.018	7.371	2.04	1.94	50
17.	EW2	1296.4	25.927	12.161	2.13	1.93	50
18.	EW3	1976.8	39.537	18.639	2.12	2.02	50
19.	EW4	2006.3	40.126	19.628	2.04	2.02	50

S.No.	**Sample ID**	**Nucleic Acid Conc.(** ng/µl)	**A260**	**A280**	**260/280**	**260/230**	**Factor**
20.	EW5	1976.8	39.535	18.584	2.13	2.01	50
21.	EW6	1817.6	36.352	17.155	2.12	2.06	50
22.	EW7	816.1	16.322	7.64	2.14	1.99	50
23.	EW8	1958.1	39.161	18.621	2.1	2.04	50
24.	EW9	1392.4	27.848	13.143	2.12	2.02	50
25.	EW10	2063.5	41.27	19.585	2.11	2.13	50
26.	EW11	2083.3	41.666	19.603	2.13	2.16	50
27.	EW12	1242.2	24.843	11.644	2.13	2.13	50
28.	IL1	485	9.7	4.577	2.12	1.77	50
29.	IL2	701.7	14.034	6.651	2.11	2.31	50
30.	IL3	228.2	4.564	2.17	2.1	1.79	50
31.	IL4	648.2	12.964	6.2	2.09	1.87	50
32.	IL5	1007.2	20.145	9.354	2.15	1.98	50
33.	IL6	1384	27.679	12.834	2.16	1.89	50
34.	IL7	333.3	6.667	3.096	2.15	2.47	50
35.	IL8	1274.5	25.49	11.652	2.19	1.99	50
36.	IL9	642.1	12.842	6.127	2.1	1.89	50
37.	IL10	744.3	14.885	6.909	2.15	2.02	50
38.	IL11	486.1	9.723	4.541	2.14	1.86	50
37.	IL12	576.9	11.538	5.332	2.16	2.03	50

S.No.	**Sample ID**	**Nucleic Acid Conc.(** ng/µl)	**A260**	**A280**	**260/280**	**260/230**	**Factor**
40.	IL13	881.5	17.63	8.174	2.16	2.43	50
41.	IL14	552.4	11.047	5.223	2.12	2.06	50
42.	IL15	510.9	10.217	4.715	2.17	2.19	50
43.	EW13	1583.8	31.677	14.878	2.13	2.18	50
44.	EW14	1426.5	28.53	13.402	2.13	2.2	50
45.	EW15	1861.4	37.228	17.363	2.14	2.09	50
46.	WS1	1163.1	23.262	10.869	2.14	1.82	50
47.	WS2	1228.9	24.579	11.415	2.15	1.96	50
48.	WS3	740.8	14.816	6.958	2.13	1.88	50
49.	WS4	951.1	19.022	9.036	2.11	2.01	50
50.	WS5	1038.5	20.771	9.768	2.13	1.99	50
51.	WS6	1324.7	26.494	12.311	2.15	1.91	50
52.	WS7	1312.4	26.248	12.283	2.14	2.06	50
53.	WS8	2102.4	42.048	19.817	2.12	2.16	50
54.	WS9	3262.2	65.243	30.398	2.15	2.25	50
55.	WS10	2828.7	56.573	26.739	2.12	2.19	50
56.	WS11	1889	37.78	17.809	2.12	2.26	50
57.	WS12	1415.7	28.314	13.143	2.15	2.23	50
58.	WS13	699.8	13.996	6.528	2.14	2.33	50
59.	WS14	1585.9	31.719	15.017	2.11	2.19	50

S.No.	Sample ID	Nucleic Acid Conc.(ng/µl)	A260	A280	260/280	260/230	Factor
60.	WS15	2101.7	42.034	20.188	2.08	2.12	50