ADDIS ABABA UNIVERSITY

Faculty of Life Sciences

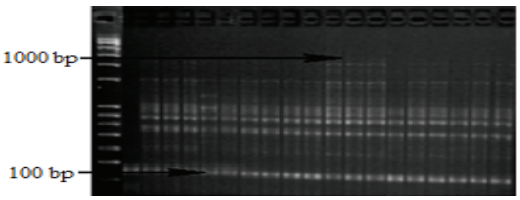
Microbial, Cellular and Molecular Biology Program Unit

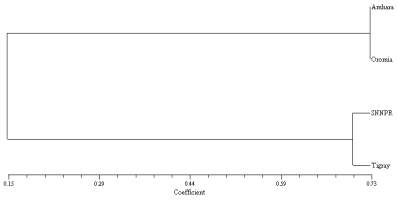
Investigation of Genetic Diversity in Ethiopian Collections of Safflower (*Carthamus tinctorius*) using ISSR Markers

By

Baye Wodajo

A Thesis Submitted to the School of Graduate Studies of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biology (Applied Genetics).





AAU, Ethiopia

June, 2012

Dedication

I dedicate this thesis manuscript to my mother Shesheg Ayalew and my brother Abrham Wodajo, for their magnificent thinking as taught man and for his steady constructive advice which is bases for success of my life. Hence, **Abrham is the secret of behind me to be better man**!!

Declaration

I, the undersigned, declare that this thesis is my own work and has not been presented in other universities, colleges or institutions, seeking for similar degree or other purposes. All sources of materials used for the thesis have been duly acknowledge.

Name: Baye Wodajo

Signature: _____

Date: _____

Advisor: Kassahun Tesfaye (PhD)

Signature: _____

Date: _____

Acknowledgements

I would like to express my deep whole-hearted gratitude and indebtedness to my advisor Dr. Kassahun Tesfaye, not only for accepting me as his student and providing the laboratory resource without hesitation for the resource but also for providing me the initial motivation and encouragement that led to the realization of this work and also thanks for his unreserved assistance, guidance and very constructive supervision from inception until the completion of the research work.

I am also grateful to the Institute of Biodiversity Conservation (IBC) for its goodwill in providing the seed samples of cultivated safflower (*Carthamus tinctorius).*

I would like to acknowledge the Ministry of Education for giving me the opportunity to join postgraduate program and I would like to express my sincere thanks to Microbial Cellular and Molecular Biology Program Unit, Addis Ababa University, for accepting me to attend postgraduate program and covering the research expenses.

I wish to express my appreciation to my fellow classmates Mr. Abush Zenaw, Shemekit Tadele, Sewalem Tehay and Said Mohamed for their collaborative works and assistance in laboratory activities, analysis and in editing the thesis report.

Finally, I would like to give many thanks to my brother, Abrham Wodajo, for his care, understanding, support and encouragement. He was a driving force through my study time.

Above all, I praise GOD for his uncountable compassions and for everything he did for me. Everything is possible with him and I accomplished this thesis with HIS help

TABLE OF CONTENTS

Contents **Pages**

LIST OF TABLES page

LIST OF FIGURES **page**

LIST OF ABBREVIATIONS

2D- Two Dimensions

3D-Three Dimensions

AFLP-Amplified Fragment Length Polymorphism

AMOVA-Analysis of Molecular Variance

DdH2O- double distilled water

h- Nei's gene Diversity

I- Shannon's diversity index

IBC- Institute of Biodiversity Conservation

ISSR-Inter Simple Sequence Repeat

NJ-Neighbor Joining analysis

NPL-number of polymorphic loci

PCO-Principal coordinated Analysis

PGRFA- Plant Genetic Resources for Food and Agriculture

PP (%) -percent polymorphism

RAPD-Random Amplified Polymorphic DNA

RFLP-Restriction Fragment Length Polymorphism

SNNPRs- Southern nation, nationality and people region

UBC- University of British Colombia

UPGMA – Un-weighted pair group method with arithmetic mean.

Abstract

Safflower, Carthamus tinctorius*, L. is an oilseed crop that belongs to the family Asteraceae, a diverse group of flowering plants that grow in many parts of the world. So far, the characterization of safflower using molecular markers has been limited. The objective of this study was to investigate the genetic diversity of safflower accessions collected from different regions of Ethiopia using ISSR molecular markers. For this purpose, seeds of seventy landrace accessions collected from four administrative regions of Ethiopia (Amhara, Oromia, Tigray and SNNPR) were obtained from the IBC and grown in greenhouse at Addis Ababa University, Faculty of Life Science. DNA was extracted from a bulk leaf sample of five randomly selected plants per accession using a triple CTAB extraction technique. Four primers were selected. The four selected ISSR primers produced a total of 43 bands across the 70 safflower accession. The number of amplified fragments with ISSR primers ranged from 6 to 14 per primer with varied in size of 100 to 1000 base pairs. Out of the 43 bands detected, 87.5% were polymorphic. In terms of region, Oromia showed the highest percentage of polymorphism (86.1%) with Tigray having the least polymorphism (20.9%). Safflower population from Oromia has 0.32 and 0.48 gene diversity and Shannon diversity index, respectively whereas Amhara population of safflower has shown 0.27 and 0.39 gene diversity and Shannon diversity index, respectively. AMOVA also showed that 98.9% of the variation is attributed to within population while 1.1% is among populations variation. Both UPGMA and neighbor joining trees based on Jaccard's similarity coefficient showed weak grouping among individuals collected from the same regions. The results reveal the presence of higher genetic diversity that deserves conservation attention and sustainable use strategy to improve the productivity of safflower.*

Key words: - *Carthamus tinctorius* L, genetic diversity, ISSR molecular marker, safflower, polymorphism, Ethiopia.

1. INTRODUCTION

Safflower, *Carthamus tinctorius*, L. is an oilseed crop that belongs to the family Asteraceae, a diverse group of flowering plants that grow in many parts of the world. The genus *Carthamus* has 25 species, of which *C. tinctorius* is the only cultivated one, and has 2n = 24 chromosomes (Helm *et al.,* 1991) and it is dicotyledonous, herbaceous, and annual plant. It is highly self-fertilizing (Knowles, 1969). It has colorful flower heads, a deep taproot, and the production of white, oil-bearing seeds. It has adapted to grow in hot, dry climates, and well-drained soil. Depending on environmental conditions, the typical generation length of safflower varies from about 17-20 weeks (Smith, 1996). Safflower is commonly known with different local name in different countries; as '*kusum*' in India and Pakistan, '*honghua*' (red flower) in China, '*Muswar*' in Afghanistan, '*Suff*' in Ethiopia (Chavan, 1961; Smith, 1996).

Safflower originated in the eastern Mediterranean, and spread to Egypt, Ethiopia, southern Europe, south Asia and the Far East early in its evolution (Smith, 1996). However, Vavilov (1951) considered Ethiopia, Afghanistan and India as the primary centers of origin of safflower. Currently, it is grown as an oilseed crop in over 60 countries worldwide and India as the largest producer of safflower (Weiss, 200).

Safflower is mainly cultivated for its seed, which is used primarily for edible oil. In the past, the crop is grown for its flowers used for coloring and flavoring foods, making dyes and medicine. However, due to an increasing demand for vegetable oil in human diet, its production as an oil seed crop has received a great deal of attention. It is well known for its quality edible oil in view of higher proportion of linoleic and oleic acid content compared to other vegetable oils. The oil is semidry in nature and is used in paints, textile

and leather industries (Zhang, 1997). In Ethiopia, boiled and finely pounded safflower kernels are mixed with water and the supernatants is used to prepare the so called 'fitfit', which is used as fasting-food. Also the supernatant is used as a drink mixed with sugar during fasting seasons. Roasted seeds, generally mixed with roasted chickpeas, barley or wheat, are eaten as a snack food in Ethiopia and Sudan (Belayneh and Wolde-Mariam, 1991; cited in Dajue and Mündel, 1996). Safflower has some agronomic advantages such as drought resistance and adaptation to arid and semiarid climatic conditions (Weiss, 2000).

Ethiopian safflower is very much neglected and it is cultivated only as a minor oil crop with very limited information available on its genetic resources. In Ethiopia, safflower cultivation is mostly done by small farmers in well fertile and drained field, usually around homesteads. Nevertheless, seed can be harvested from the plant for oil extraction, roasted seed and other traditional uses. Despite all these local uses and industrial applications of safflower plants, it remains as low acreage crop compared to other oilseed crops. Moreover, limited attention has been given to conserve, add value and improve the productivity of safflower in Ethiopia (Edwards, 1991).

2. LITERATURE REVIEW

2.1. Taxonomy and botany of *Carthamus tinctorius*

2.1.1. The Genus *Carthamus* and its chromosome number

Genus *carthamus* is belongs to Asteraceae family, which comprises 25 species. The only cultivated species is C. *tinctorius*. The remaining 24 species are the wild safflower species . *C. oxycantha, C. palaestinus*, *C. alexandrius*, *C. glaucus*, *C. syriacus*, *C.tenuis, C. lanatus, C. baeticus,C. creticus,* and *C. turkestanicus* are some examples of the wild safflower species (Khidir and Knowles, 1970). *Carthamus* classified in to four sections based on chromosome numbers. Section one (2n=20) includes *C. oxycantha* and *C. palaestinus*, section two (2n=24) includes *C. tinctorius*, *C. alexandrius*, *C. glaucus*, *C. syriacus* and *C.tenuis*, section three (2n=44) *C. lanatus* and section four (2n=64) includes *C. baeticus, C. creticus* and *C. turkestanicus*. The first two sections are diploids; the third and the fourth section are polyploidy species (Khidir and Knowles, 1970). However, information with regard to the origin of cultivated safflower from the *C. oxycantha* and *C. palaestinus* while origin of the polyploid species such as *C. lanatus* (2n=44), *C. creticus* and *C. turkestanicus* (2n=64) were pointed to possible progenitors. Studies of Chapman *et al.* (2007) on characterization of *Carthamus* species with 2n = 24 using universal markers specific to Asteraceae indicated *C. palaestinus* as the progenitor species of cultivated safflower. Among the *Carthamus* species cultivated safflower harbored the lowest levels of nucleotide diversity, while *C. oxyacanthus* and *C. palaestinus* exhibited highest and intermediate level of diversity (Chapman *et al.,* 2007). However, the somatic chromosome analysis of the genus *Carthamus* is very difficult due to poor stainability, stickiness and tendency to overlap at metaphase and diffuse

appearance of primary and secondary constrictions of the chromosomes. An efficient squash technique for resolving the somatic chromosomes and permit detailed analysis of safflower karyotype was developed by Malik and Srivastava, (2009).

2.1.2. Botanical description and reproductive biology

Safflower is a dicotyledonous, herbaceous, winter annual, thistle-like plant that is highly self-fertilizing, with out-crossing rates of less than 10% (Knowles, 1969). Its most prominent features are colorful flower heads, a deep taproot, and the production of white, oil-bearing fruits. Safflower is suited to grow in hot, dry climates, where soils are moist in early spring but generally well-drained and can be cultivated in regions where seasonal rainfall is around 375 mm and 1400m altitude. Depending on environmental conditions, the typical generation length of safflower varies from about 17-20 weeks and the growth cycle is divided into the following stages: emergence, rosette, stem elongation, branching, flowering and maturity (Figure 1.Mündel *et al*., 2004).

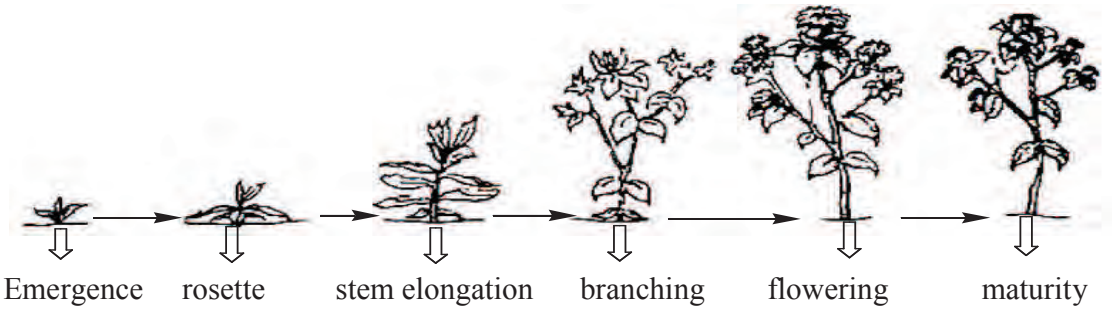


Figure 1: Growth cycle of safflower plants (Mündel *et. al.*, 2004.)

Depending on moisture availability and temperature safflower seeding emergence occurs between 3 days to 3 weeks. Safflower emerges once soil temperatures rise to 4.4°C, with emergence rates far higher when temperatures are 15.6°C or above (Kaftka, 1965). If

planted when soil temperatures are the 4.4°C required for germination, safflower seedlings can take up to 3 weeks to emerge. If planted when temperatures are above 15.6°C, seedlings emerge in 3–4 days. Once the cotyledons of emerged seedlings have spread, leaves begin to grow and the plant enters a rosette stage, where the stem does not elongate but the plant develops a long taproot that can grow 16-36 inches in depth (Dajue and Mündel, 1996).The nature of the rooting system in safflower, being that it has one deep taproot and several successive lateral roots, enables safflower to obtain water and nutrients from well below the surface soil level which gives safflower a degree of drought resistance. The length of the taproot and the configuration of the laterals will depend on the type of soil and amount and depth of available moisture. The deep rooting character helps safflower do well in areas of relatively low rainfall (Knowles *et al.,* 1965).

After spending 2-3 weeks in a rosette stage, stem elongation proceeds very quickly, and much branching of the main axis occurs (Mündel *et al*., 2004). The safflower stem is stiff and cylindrical, thick at its base and becomes thinner as branching progresses, relatively smooth and hairless, white to light gray or green in color, and marked with very fine longitudinal grooves (Weiss, 1971). The circumference of the plant stem at its base varies from 3 to 12 cm (Chavan, 1961). Stem begins to branch after it reaches a height of 8–15 inches. The first branches of the main axis are called the primary branches and the branches stemming from the primary branches are called the secondary branches and so on. Safflower can grow to a height of 0.5-1.8m (Mündel *et al*., 2004).

Safflower leaves are arranged in a special manner on the stems, set opposite each other. The leaves are divided longitudinally by a noticeable midrib that is more pronounced next to the stem. There are lightly defined lateral veins leading off the midrib. The leaf

size varies greatly among cultivars, and leaf shape can be linear, lanceloate, ovate or oblong which is the leaf are between 2.5 to 5 cm wide and from 10 to 15 cm long with acuminate (pointed) tips with increasing spininess towards the top of the plant. Normally, spines are present on the leaves and bracts of safflower (Bradley *et al*., 1999). Leaves become shorter and stiffer on the upper reaches of the plant until reaching the terminal bud branches where the leaves become still shorter, ovate to obovate shapes, getting closer and closer together until they crowd on each other in an involucral whorl around the flower head (Li *et al.,* 1993).

Flowering begins once stems are elongated and branching is extensive. Groups of flowers are contained in a flower head is a capitulum, a characteristic trait of all plants in the Compositae family. A safflower plant has many capitula, each of which is found at the end of a stem and surrounded by involucral bracts. A given capitulum can have 20 to 180 flowers or florets. Within a capitulum, flowering starts at the outer florets and proceeds centripetally towards the centre. Flowering begins at the primary capitulum of a safflower plant which is located at the top of the main stem and then continues toward the outer stems. Florets are tubular and contain five corolla lobes which spread during anthesis. On each floret are five stamens which are united in the anther tube. As the style and stigma grow up through the surrounding anther tube, self-pollination occurs. An ovary is located at the base of each floret, which develops into a single-seeded achene or fruit (Knowles, 1980). Flower color varies from white (rarely seen) to deep red with shades of yellow and orange being most common (Dajue and Mündel, 1996).

The achenes, which simply are called the seed, are four-sided with white color. The achene mature within 30 to 35 days of flowering and then after which safflower requires

another two weeks to dry, prior to harvest. Safflower seeds consist of fibrous hull that protects a kernel made up of two cotyledons and an embryo. The oil-bearing fruits contain 33-60% hull and 40-67% kernel with oil content typically ranging from 20-45% (Dajue and Mündel, 1996). The hull makes up 18–59% of the seed weight. Once the leaves are brown and the capitula have dried such that very little green, if any, remains on the bracts of the last developing capitula, plants are ready for harvest (Mündel *et al*., 2004).

During the growth cycle, safflower plants are subject to numerous diseases with Alternaria leaf blight caused by *Alternaria carthami* and Sclerotinia head rot caused by *Sclerotinia sclerotiorum* being the most common (Kaftka, 1965). Both can significantly affect yield particularly in unseasonably wet conditions. Safflower can also suffer damage by pests and herbivores, including grasshoppers, thrips, deer and weeds. Safflower has some agronomic advantages such as drought, salt resistance, and adaptation to arid and semiarid climatic conditions (Weiss, 2000).

2.2. Center of origin and its geographical distribution

Population genetic analyses of crop gene pools can provide insight into the origin and subsequent evolution of crop plants and can also result in the identification of novel sources of genetic variation for the continued improvement of crop plants (Zohary and Hopf, 2000). Early researchers proposed the existence of a number of centers of safflower gene pool. Accessions within each of the proposed centers are known to be quite similar to one another with respect to certain attributes such as height, branching, head size and flower color whereas differences are evident between the centers. For instance, early in

its evolution, *C. tinctorius* spread to Egypt, Ethiopia, southern Europe, south Asia, and the Far East (Ashri, 1973). But Vavilov (1951) proposed three centers of origin for cultivated safflower (*Carthamus tinctorius L.*). India as first center was proposed based on variability and ancient culture of safflower production. Afghanistan as the second center was proposed based on safflower diversity and proximity to wild species. A third center of origin in Ethiopia was primarily based upon the presence of the wild safflower species (*C. lanatus*) However, contrary to the above, Ashri and Knowles (1960) and Hanelt (1961) indicated the center of origin to be in the Near East. This assumption was based on the similarity of cultivated safflower to two closely related wild species: *C. flavescens* reported from Turkey, Syria, and Lebanon and *C. palaestinus* found in desert areas of western Iraq and southern Israel. Currently, over 60 countries grow safflower with over half of the world's production occurring in India, mainly for the domestic vegetable oil market. Production in Argentina, Australia, Ethiopia, Kazakhstan, Mexico, and the United States constitutes most of the remainder (Ekin, 2005).

2.3. Importance of safflower

Safflower was originally grown for the flowers that were used in making red and yellow dyes for clothing and food preparation. Today, this crop supplies oil, meal, birdseed, and medicinal purpose (Weiss, 1983).

Mostly safflower is cultivated for its oil which is widely used for cooking throughout the world (Dajue and Mundel, 1996). Seeds consist of 30-45% oil, 15-20% protein and 35-45% hull (Dajue and Mundel, 1996). It produces two types of oils, one rich in monounsaturated fatty acid (oleic acid, MUFA) and the other rich in polyunsaturated

fatty acid (linoleic acid, PUFA). Standard safflower oil contains 6-8% palmitic acid, 2-3% stearic acid, 16-20% oleic acid, and 71-75% linoleic acid (Nagaraj, 1993). Safflower oil has the highest source of polyunsaturated fats (PUFA) than that found in any other type of vegetable oil that is almost 79 percent polyunsaturated fatty acids (linoleic acid) and 13 percent of monounsaturated fatty acids (oleic acid) but only 8 percent of saturated fatty acids. The standard polyunsaturated type is used in soft margarines and for cooking, as well as in surface coatings because of its high linoleic acid content. The monounsaturated fatty acid (oleic acid) is known to reduce low-density lipoproteins (LDLs; bad cholesterol) without affecting high-density lipoproteins (HDLs; good cholesterol) in blood (Smith, 1996). High content of linoleic acid in safflower oil can control arteriosclerosis, reduce serum cholesterol, and thus cure hypertension, high blood fat, coronary heart disease, and obesity and control blood pressure (Dajue and Mündel, 1996).The other essential nutrients present in safflower oil are omega-6 fatty acids, cis-linoleic acid, vitamin E and vitamin K(Dajue and Mundel, 1996). Safflower oil can be used in manufacturing of surfactant, alcohol and varnishes due to its high linoleic acid content (Dajue and Mundel, 1996). Safflower oil is similar to olive oil in nutrition value due to high levels of linoleic acid for being very stable on heating and not giving any smell during frying (Gyulai, 1996).

The safflower oil is also extremely important for a bio-fuel or fuel extender. Biodiesel fuel is attractive because of its potential to reduce greenhouse gas emissions and ability to be used directly in diesel engines with only minor alterations. Safflower oil has the ideal low melting point for biodiesel production. However, high oleic acid producing cultivars of safflower have greater oxidative stability. High oleic acid producing cultivars of

safflower do have a possible pollutant-reducing effect because the oil is biodegradable, largely free of sulphur, and lacks the fossil fuel carbon dioxide (Flynn and Berman, 2001).

Different reports are available in literature which shows that safflower is an important medicinal plant. In China, safflower is cultivated for its flowers which are medicinally used for blood circulation. Some therapeutic medicines have been prepared such as safflower injection, safflower tea, safflower capsule etc. which are used to combat different diseases (Stobart and Stymne, 1985; Zhaomu and Lijie, 2001). Besides China, safflower is also used as herbal medicines in Korea, India and Japan. Safflower florets are used as a cure of human diseases such as diabetes and high blood pressure (Zhaomu and Lijie, 2001).

In Ethiopia, boiled and finely pounded safflower kernels are mixed with water to prepare the so called 'safflower's fitfit', which is used as fasting-food. Roasted safflower seeds, are mixed with roasted seeds of chickpeas, barley or wheat, are eaten as a snack food in Ethiopia and Sudan (Belayneh and Wolde-Mariam, 1991; cited in Dajue and Mündel, 1996). Safflower also supplies meal used in animal forage and poultry. Safflower forage is of comparable quality to cereal or alfalfa and is desired by cattle, sheep and goat and birds. (Landaua *et al*., 2004).

2.4. **Marker systems and their applications in genetic diversity analysis**

There are different types of markers: those based on visually assessable traits (morphological and agronomic traits), those based on gene product (biochemical markers) and those relying on a DNA assay (molecular markers).

2.4.1. Morphological markers

Morphological markers are a marker system based on phenotypic appearance. It is the earliest genetic markers used for assessment of variation and still has great importance. Moreover, morphological characters are simple to score and inexpensive too (Vithanage *et al.*, 1995).

The main disadvantages of this approach are the influence of the environment and resources needed for evaluation in the field and greenhouse. Moreover, some morphological traits are observed to be very plastic in nature which could be easily affected by environmental change and affect the exact relationships of plant species and/or populations (Vithanage *et al.*, 1995).

2.4.2. Biochemical Markers

Biochemical markers are markers based on protein polymorphisms through electrophoresis separation of protein molecules. A tissue extract is prepared and electrophoresed on a non denaturing starch or polyacrylamide gel. The proteins of this extract are separated by their net charge and size. After electrophoresis, the position of a particular enzyme in the gel is detected by adding a colorless substrate that is converted in to a dye under appropriate reaction conditions. Depending on the number of loci their state of homo or heterozygosity and the enzyme configuration (i.e., the number of separable units), from one to several bands are visualized. The positions of these bands can be polymorphic and can be considered as informative loci (Weising *et al*., 2005).

The main advantages of protein markers are their co-dominant inheritance (not all proteins co-dominant) and the technical simplicity and low cost of the assay. Its

limitation includes the restricted number of suitable allozyme loci in the genome, the requirement of fresh tissue and the sometimes limited variation (Weising *et al*., 2005).

2.4.3. **Molecular markers and its significance**

In current scenario, the DNA markers become the marker of choice for the study of crop genetic diversity and have become revolutionized plant biotechnology (Karp *et al*., 1997). Increasingly, techniques are being developed to more precisely, quickly and cheaply assess genetic variation. DNA-based molecular markers have acted as versatile tools and have found their own position in various fields like taxonomy, physiology, embryology, genetic engineering. The discovery of PCR (polymerase chain reaction) was a landmark in this effort and proved to be a unique process that brought about a new class of DNA profiling markers. Provide valuable data on diversity through their ability to detect variation at the DNA level (Karp *et al*., 1997).

Properties of the molecular markers

An ideal molecular marker must have some desirable properties. Below are the ideal properties of molecular markers as described by Weising *et al.* (1995).

1. Highly polymorphic nature: It must be polymorphic as it is polymorphism that is measured for genetic diversity studies. 2. Co-dominant inheritance: determination of homozygous and heterozygous states of diploid organisms. 3. Frequent occurrence in genome: A marker should be evenly and frequently distributed throughout the genome. 4. Selective neutral behaviors. 5. Easy access (availability): It should be easy, fast and cheap to detect. 6. Easy and fast assay 7. High reproducibility 8. Easy exchange of data between laboratories.

It is extremely difficult to find a molecular marker which would meet all the above criteria. A wide range of molecular techniques is available that detects polymorphism at the DNA level. Depending on the type of study to be undertaken a marker system can be identified that would fulfill at least a few of the above characteristics (Weising *et al.,* 1995).

In recent years different marker systems such as Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNAs (RAPDs), Amplified Fragment Length Polymorphisms (AFLPs), Simple Sequence Repeats (SSRs) or micro-satellites and others have been developed and applied to a range of crops genetic variation (Reddy *et al.,* 2002).

2.5.3.1. Random Amplified Polymorphic DNA (RAPD)

In 1991 Welsh and McClelland developed a new PCR-based genetic assay namely randomly amplified polymorphic DNA (RAPD). RAPDs are DNA fragments amplified by the PCR using short synthetic primers (10 bp) of random sequence. These oligo-nucleotides serve as both forward and reverse primer and are usually able to amplify fragments from 1–10 genomic sites simultaneously. Amplified products are separated on agarose gels in the presence of ethidium bromide and view under ultraviolet light (Jones *et al.,* 1997).

The main advantage of RAPDs is that they are quick and easy to assay because PCR is involved only low quantities of template DNA are required, usually 5–50 ng per reaction. Since random primers are commercially available, no sequence data for primer construction are needed. Moreover, RAPDs have a very high genomic abundance and are

randomly distributed throughout the genome (Williams *et al*., 1993). While the main drawback of RAPDs is their low reproducibility and hence highly standardized experimental procedures are needed because of their sensitivity to the reaction conditions. RAPD analyses generally require purified, high molecular weight DNA, and precautions are needed to avoid contamination of DNA samples (Schierwater and Ender, 1993).

2.5.3.2. Amplified Fragment Length Polymorphism (AFLP)

AFLP is based on a selectively amplifying a subset of restriction fragments from a complex mixture of DNA fragments (80–500 bp) obtained after digestion of genomic DNA with restriction endonucleases. Polymorphisms are detected from differences in the length of the amplified fragments by polyacrylamide gel electrophoresis (PAGE) (Matthes *et al.,* 1998) or by capillary electrophoresis. The technique involves four steps: (1) restriction of DNA and ligation of oligonucletide adapters (2) preselected amplification (3) selective amplification (4) gel analysis of amplified fragments. These fragments are viewed on denaturing polyacrylamide gels either through autoradiography or fluorescence methodologies (Vos *et al*., 1995, Jones *et al*., 1997).

AFLP have high genomic abundance, considerable reproducibility and the generation of many informative bands per reaction, their wide range of applications and the fact that no sequence data for primer construction are required. AFLPs may not be totally randomly distributed around the genome as clustering in certain genomic regions, such as centromeres has been reported for some crops (Alonso-Blanco *et al.,* 1998; Saal and Wricke, 2002). Its limitation includes the need for purified, high molecular weight DNA, the dominance of alleles. In addition, due to the high number and different intensity of

bands per primer combination, there is the need to adopt certain strict but subjectively determined criteria for acceptance of bands in the analysis (Vos *et al.,* 1995).

2.5.3.3. Microsatellites or Simple sequence Repeat (SSR)

The term microsatellites were coined by Litt and Lutty (1989) and it also known as Simple Sequence Repeats (SSRs), it consists of tandomly repeated 1-7 base pair units, and distributed widely throughout the genome. These are heritable, useful to monitor gene flow, excellent for parentage determination and ideally suitable for analysis via multiplexing with highly reproducible profiles (Powell *et al.,* 1996). They are easy to detect with PCR and a typical microsatellite marker has more variants than those from other marker systems. Initial identification and screening of microsatellite loci are time-consuming. Microsatellite sequences are especially suited to distinguish closely related genotypes because of their high degree of variability (Smith and Devey, 1994). Microsatellites are ideal genetic markers for detecting differences between and within species. The variation in the number of tandomly repeated units results in highly polymorphic banding patterns. Microsatellites are not limited to the nuclear genome but they occur in chloroplast as well as in mitochondrial genome (Soranzo *et al*., 1999). Nuclear microsatellites are bi-parental inheritance and possess few loci with many alleles per locus. They are co-dominant, can detect large variation within population confer stability to the genome (Farooq and Azam, 2002).

The strengths of microsatellites include the co-dominance of alleles, their high genomic abundance in eukaryotes and their random distribution throughout the genome, with preferential association in low-copy regions (Morgante et al., 2002). Because the

technique is PCR-based only low quantities of template DNA (10–100 ng per reaction) are required. Due to the use of long PCR primers, the reproducibility of microsatellites is high and analyses do not require high quality DNA while main drawbacks of microsatellites is that high development costs are involved if adequate primer sequences for the species of interest are unavailable, making them difficult to apply to unstudied groups (Ghislain *et al*., 2004).

2.5.3.4. Inter Simple Sequence Repeats (ISSR)

ISSRs are DNA fragments of about 100–3000 bp located between adjacent, oppositely oriented microsatellite regions. These techniques, first reported by Zietkiewicz et al. (1994) are primers based on microsatellites are utilized to amplify inter-SSR DNA sequences. ISSRs are amplified by PCR using microsatellite core sequences as primers with a few selective nucleotides as anchors into the non-repeat adjacent regions (16–18 bp). About 10–60 fragments from multiple loci are generated simultaneously separated by gel electrophoresis and scored as the presence or absence of fragments of particular size.

The main advantage of ISSRs is that no sequence data for primer construction are needed. Because the analytical procedures include PCR only low quantities of template DNA are required (5–50 ng per reaction). Furthermore, ISSRs are randomly distributed throughout the genome and characterize as multilocus dominant marker. Its disadvantages include the possible non homology of similar sized fragments (Godwin *et al.,* 1997, Zietkiewicz *et al.,* 1994, Gupta *et al.,* 1994).

Generally, Inter simple sequence repeats (ISSR) as a dominant marker with a large number of polymorphic fragments per primer has been successfully employed to reveal genetic variation and characterize genotypes in some cultivated crops such as genetic diversity in rice (Gezahagn Girma, 2007), Coffee (Solomon Balami, 2007), Sesame (Dagmawi Teshome, 2011) and lentil (Edossa Fikiru *et al.*, 2011). Recently, it has effectively provided information on genetic diversity among cultivated (Yang *et al.,* 2007) and wild (Ash *et al.,* 2003) safflower germplasm.

The major limitations of methods which are low reproducibility of RAPD, high cost of amplified fragment length polymorphism (AFLP) and the need to know the flanking sequences to develop species specific primers for SSR polymorphism so that ISSR-PCR is a technique that overcomes most of these limitations (Wu *et al.,* 1994). It is rapidly being used by the research community in various fields of plant improvement (Reddy *et al*., 2002). ISSR-PCR gives multilocus patterns which are very reproducible, abundant and polymorphic in plant genomes (Zietkiewicz *et al.,* 1994). As a result of these advantages and their universality and easiness of development ISSR markers are employed to study and characterize germplasm.

Safflower possesses considerable diversity across different regions of the world. In order to design an appropriate breeding program, it is important to know how much the phenotypic variation of a trait is heritable (Kearsey and Pooni, 1996), since the efficiency of a selection program is mainly dependent on the magnitude of genetic variation and heritability of a trait (Falconer and Mackay, 1996). Genetic diversity of some safflower germplasm has been previously investigated based on agro-morphological markers were the first to be used for genetic diversity study (Ashri, 1975; Jaradat and Shahid, 2006).

Zhang *et al*. (2000) and Han and Li (1992) used isozymes techniques to study and classify *C.tinctirius* germplasm assembled from India. RAPD markers were also applied to study the genetic diversity of *C.tinctirius* landraces germplasm collected from Iran landraces (Amiri *et al*., 2001; Vilatersana *et al.*, 2005) while ISSR marker has been used by Ash *et al.* (2003) and Yang *et al*. (2007) to investigate relationship among *C.tinctirius* germplasm collected from different part of the world. AFLP marker also was used by Sehgal and Raina (2005) and Johnson *et al.* (2005) studied the diversity of safflower populations and accessions from diverse geographic and genetic sources to characterize safflower germplasm using amplified fragment length polymorphism (AFLP) markers. None of the above studies either focused on Ethiopian germplasm or include samples from Ethiopia. Hence, this study was designed with the objective of investigating the genetic diversity of safflower (*C. tinctorius*) collected from Ethiopia using ISSR molecular markers.

3. OBJECTIVES OF THE STUDY

3.1. General objective:

- The main objective of the present research is to investigate the genetic diversity and population structure of the safflower *(C. tinctorius*) germplasm collected from Ethiopia using ISSR marker

3.2. Specific objectives:

- To assess the extent of genetic variation within and between populations of cultivated safflower.
- To identify sites with high genetic diversity for germplasm collection and conservation purpose.
- To determine population differentiation among population of safflower with respect to regions.

4. MATERIALS AND METHODS

4.1. Plant materials and growth conditions

Seeds of 70 safflower accessions collected from four different administrative regions of Ethiopia were obtained from Institute of Biodiversity Conservation (IBC), Addis Ababa, Ethiopia (Table 1). All the 70 safflower accessions seeds samples were grown in a greenhouse and fresh leaves from five week old plants were used for genomic DNA extraction.

Table 1: List and origin of the safflower (*C. tinctorius*) populations.

No.	Accession	Region codes	Region	Zone	Latitude (degree)	Longitude (degree)	Altitude (meter)
1	207454	AMH1	Amhara	Debub Gondar	-	-	-
2	200486	AMH2	Amhara	Debub Gondar	-	-	-
3	207458	AMH3	Amhara	Debub Gondar	-	-	-
4	207475	AMH4	Amhara	Debub Gondar	-	-	-
5	207476	AMH5	Amhara	Debub Gondar	-	-	-
6	207477	AMH6	Amhara	Debub Wello	-	-	-
7	212588	AMH7	Amhara	Debub Wello	11-35-00-N	39-38-00-E	1440
8	51526	AMH8	Amhara	Debub Wello	-	-	-
9	51523	AMH9	Amhara	Misrak Gojam	10-07-00-N	38-09-00-E	1880
10	51524	AMH10	Amhara	Misrak Gojam	10-06-00-N	38-10-00-E	1610

11	51517	AMH11	Amhara	Misrak Gojam	-	-	-
12	51522	AMH12	Amhara	Misrak Gojam	10-27-00-N	38-12-00-E	2570
13	239094	AMH13	Amhara	Misrak Gojam	-	-	-
14	212683	AMH14	Amhara	Misrak Gojam	-	-	1850
15	51518	AMH15	Amhara	Misrak Gojam	-	-	
16	241791	AMH16	Amhara	Semen Gondar	12-19-00-N	37-42-00-E	1915
17	208017	AMH17	Amhara	Semen Gondar	12-40-00-N	37-23-00-E	
18	207457	AMH18	Amhara	Semen Gondar	-	-	-
19	208019	AMH19	Amhara	Semen Gondar	12-27-00-N	36-53-00-E	-
20	207456	AMH20	Amhara	Semen Gondar	-	-	-
21	207469	AMH21	Amhara	Semen Gondar	-	-	-
22	51510	AMH22	Amhara	Semen Gondar	-	-	-
23	207459	AMH23	Amhara	Semen Gondar	-	-	-
24	207466	AMH24	Amhara	Semen Gondar	-	-	-
25	235026	AMH25	Amhara	Semen Shewa	-	-	-
26	212472	AMH26	Amhara	Semen Shewa	09-56-00-N	38-53-00-E	1580
27	212471	AMH27	Amhara	Semen Shewa	09-43-00-N	38-53-00-E	1720
28	212473	AMH28	Amhara	Semen Shewa	10-14-00-N	39-00-00-E	1570
29	229085	AMH29	Amhara	Semen Shewa	-	-	-
30	241793	AMH30	Amhara	Semen Wello	12-02-00-N	39-02-00-E	2070
31	241794	AMH31	Amhara	Semen Wello	12-02-00-N	39-01-00-E	1990
32	241792	AMH32	Amhara	Semen Wello	12-04-00-N	39-04-00-E	2045

33	214915	AMH33	Amhara	Semen Wello	-	-	-
34	200487	ORO1	Oromia	Arsi	-	-	-
35	200488	ORO2	Oromia	Arsi	-	-	-
36	208827	ORO3	Oromia	Arsi	08-17-00-N	39-20-00-E	1740
37	208826	ORO4	Oromia	Arsi	08-18-00-N	39-15-00-E	2140
38	231323	ORO5	Oromia	Arsi	08-35-00-N	39-52-00-E	1740
39	231324	ORO6	Oromia	Arsi	08-37-00-N	39-50-00-E	1600
40	232187	ORO7	Oromia	Arsi	08-38-00-N	39-49-00-E	1620
41	232188	ORO8	Oromia	Arsi	-	-	-
42	200489	ORO9	Oromia	Arsi	-	-	-
43	216851	ORO10	Oromia	Arsi	-	-	2340
44	229954	ORO11	Oromia	Bale	07-08-00-N	40-41-00-E	1950
45	229953	ORO12	Oromia	Bale	-	-	1970
46	229952	ORO13	Oromia	Bale	06-64-00-N	-	1940
47	235718	ORO14	Oromia	Borena	04-56-00-N	37-50-00-E	1480
48	207983	ORO15	Oromia	Illubabor	08-17-00-N	35-00-00-E	-
49	200485	ORO16	Oromia	Mirab Harerge	-	-	-
50	223285	ORO17	Oromia	Mirab Harerge	-	-	-
51	203243	ORO18	Oromia	Mirab Shewa	-	-	2050
52	229086	ORO19	Oromia	Mirab Shewa	-	-	-
53	51506	ORO20	Oromia	Mirab Shewa	-	-	-
54	203242	ORO21	Oromia	Mirab Shewa	-	-	2160
55	228657	ORO22	Oromia	Mirab Shewa	-	-	-

56	236840	ORO23	Oromia	Mirab Shewa	-	-	-
57	205052	ORO24	Oromia	Misrak Shewa	-	-	-
58	202944	ORO25	Oromia	Misrak Shewa	-	-	1600
59	208895	ORO26	Oromia	Misrak Shewa	08-20-00-N	38-55-00-E	-
60	51505	ORO27	Oromia	Misrak Wellega	09-28-00-N	36-31-00-E	1440
61	236839	ORO28	Oromia	Semen Shewa	-	-	-
62	235715	SNNPR1	SNNPR	Bench Maji	05-58-00-N	37-31-00-E	1200
63	235716	SNNPR2	SNNPR	Bench Maji	05-16-00-N	37-23-00-E	1670
64	235717	SNNPR3	SNNPR	Bench Maji	05-20-00-N	37-26-00-E	1570
65	208894	SNNPR4	SNNPR	Semen Omo	06-00-00-N	37-32-00-E	1440
66	235485	TIG1	Tigray	Debub Awi	-	-	-
67	215576	TIG2	Tigray	Debub Awi	12-13-00-N	38-43-00-E	-
68	238275	TIG3	Tigray	Mehakelegnaw	13-09-00-N	39-05-00-E	2110
69	235484	TIG4	Tigray	Mehakelegnaw	13-43-00-N	39-06-00-E	1500
70	219794	TIG5	Tigray	Mehakelegnaw	-	-	-

4.2. Genomic DNA extraction

Fresh leaf samples were prepared for DNA extraction. Five plant bulks were analyzed in order to represent the genotypic variability present within each cultivated safflower accession. Young leaves were collected separately from five randomly selected individual plants of each accession just before flowering and equal proportions of leaves were grinded for genomic DNA extractions. Genomic DNA was extracted from leaf samples according to Borsch *et al*. (2003), 100mg of fresh leaf tissues were grind to

powder in liquid nitrogen and transferred to 1.5ml capacity eppendorf tube and then 700 µl hot (65ºC) CTAB was added and vortexed and the rack with eppendorf insert to the water bath at (65^{o}c) for 40 minutes and centrifuged at 1300rpm for 7 minutes. Upper layer was carefully removed to clean the tube and to ensure the complete separation of the genomic DNA to remove secondary metabolites, 600 µl chloroform was added and centrifuged at 1300rpm for 7 minutes and this step was repeated for better cleanup of the genomic DNA. Iso-propanol (2/3 of volume) was added to solution and left overnight at -20ºC. The solution was centrifuged at 13000 rpm for 30 minutes to pour off supernatant and pellet was washed with 70% ethanol and then dried in vacuum. Finally the pellet was re- suspended in 100 µl Tris EDTA and stored at -20°C**.** (See the Appendix 1). Purity and concentration of DNA were monitored using agarose gel (50ml 1xTBE and 1% agarose) and DNA pellets stored at 4°C until further analysis.

4.3. Test gel and electrophoresis

An agarose gel (50ml 1xTBE and 0.5g agarose) was prepared and 2 µl each genomic DNA samples, 3µlddH_2O and 1µl loading dye was loaded on to the gel and electrophoresed at constant voltage of 80V for 40 minutes. Gel picture was taken under UV dual-intensity transilluminator by Zenith with digital canon camera. From the second extractions following the Protocol given by (Borsch *et al.*, 2003) those with high band intensity and less smear were selected for further dilution PCR.

4.4. Primer selection and optimization

A total of 10 ISSR primers then were available at Genetics Laboratory were used for the initial testing of variability and reproducibility (UBC 900). For optimization and

screening of primers, one individual was selected from each population with 1:5 dilutions. All pre-selected ten primers were used for reproducibility and polymorphism. Finally, two di-nucleotide primers (primer 812 and 818), one tetra- nucleotide primer (primer 873) and one penta-nucleotide primer (primer 880) (Table 2) were selected based on polymorphism and reproducibility.

Table 2: List of primers, annealing temperature, sequence and amplification pattern

Primers	Annealing temperature	Sequence	Amplification pattern
810	45	5'-GAGAGAGAGAGAGAGAT-3'	Reproducible but not polymorphic
812	45	5'-GAGAGAGAGAGAGAGAA-3'	Good
818	48	5'-CACACACACACACACAG-3'	Good
824	48	5'-TCTCTCTCTCTCTCTCG-3'	Polymorphic but not reproducible
834	45	5'-AGAGAGAGAGAGAGAGYT-3'	Reproducible but not polymorphic
844*	48	5'-CTCTCTCTCTCTCTCTRC-3'	Polymorphic but not reproducible
872*	38	5'-GATAGATAGATAGATA-3'	Polymorphic but not reproducible
873*	45	5'-GACAGACAGACAGACA-3'	Good
874*	48	5'-CCCTCCCTCCCTCCCT-3'	Polymorphic but not reproducible
880**	45	5'-GGAGAGGAGAGGAGA-3'	Good

Source: Primer kit 900 (UBC 900); Single-letter abbreviations for mixed base positions: R = (A, G) Y = (C, T)

*Tetra-nucleotide primers, **Penta-nucleotide primer and the rest are di-nucleotides.

4.5. DNA amplification

ISSR PCR amplification based on four primers were carried out following the reaction mixture below. PCR amplification was carried out in a 25 µl reaction mixture containing 1µl template DNA, 13.45µl ddH20, 5.6µl dNTP (1.25mM), 2.6µl Taq buffer (10xThermopol reaction buffer), 1.25µl Mgcl2 (50mM), 0.6µl primer (20pmol/µl) and 0.5µl Taq Polymerase (5u/µl). The amplification program was performed in a Biometra 2003 version 3.10 TPersonal using 48 well plates under 2 minutes preheating and initial denaturation at 94^0C, followed by a regular cycling event of 40 cycles of 20 seconds denaturation at 94^oC, 1 minute primer annealing at (45^0C/ 48^0C) based on primers used, 1.30 minutes extension at 72^0C. The final extension for 7 minutes at 72^0C followed. The PCR reactions were also stored at 4^0C until loading on gel for electrophoresis.

4.6. Preparation of electrophoresis buffer and gel

The quantity and quality of genomic DNA, was tested using agarose gel electrophoresis. Stock solution of 10X Tris Borate EDTA (TBE), commonly used electrophoresis buffer: 108 gm Tris base; 55 gm Boric acid; 40ml EDTA, pH 8.57 components per liter was prepared and stored at room temperature. From the stock, working solution of 10X TBE prepared and then further diluted to 1x TBE and used to prepare the gel as well as fill the electrophoresis tank. The PCR products were checked first on test gel (1% agarose gel, with 1xTBE) for the presence of PCR products and later the ISSR products were run on gel with 1.67%. The ISSR gel with 1.67% was prepared by boiling 1xTBE in 500 ml Erlenmeyer flask in micro oven for 2:30 minutes. After the agarose solution cooled down at room temperature, 12µl of ethidium bromide was added for better visualization of the

gel. Then it was poured on to the gel tray and the comb was inserted immediately after the agarose was poured and the gel was left for three hours to set properly. The comb was carefully removed and put the gel tray in to the electrophoresis tank properly filled with electrophoresis buffer, 1XTBE.

4.7. Loading ISSR PCR product

The amplified ISSR product of 8µl was loaded onto ISSR gel with 2µl loading dye. DNA ladder 100 bp was also used to estimate molecular weight. The electrode was connected and the power supply was turned on; the voltage was adjusted at 100 V and left for three hours.

4.8. Visualization of the gel and staining

The electrophoresed gel was photographed using UV dual-intensity transluminator, Zenith, canon camera connected with computer. But gel picture was taken after stained with 25µl ethidium bromide (10mg/ml) which was mixed with 250 ml double distilled water for 30 minutes and washed for 30 minutes using double distilled water. Different photographs using different lens aperture, were taken and saved for later scoring.

4.9. Data scoring and statistical analysis of diversity

Although a large number of fragments were generated from each primer, only clearly distinguishable and reproducible bands were selected and data were entered in a computer file as a binary matrix “0” coded for absence and “1” for presence of a band. Data from the ISSR studies was analyzed using various statistical programs. POPGENE version1.32 software (Yeh *et al.*, 1999) was used to calculate genetic diversity for each

population as number of polymorphic loci, percent polymorphism, mean of Nei's genetic diversity and Shannon index. Analysis of molecular variance (AMOVA) was used to calculate variation among and within population using Areliquin version 3.01 (Excoffier *et al.,* 2006). NTSYS- pc version 2.02 (Rohlf, 2000) and Free Tree 0.9.1.50 (Pavlicek *et al.,* 1999) software's were used to calculate Jaccard's similarity coefficient which is calculated with the formula:-

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

Where,

'a 'is the total number of bands shared between individuals i and j,

'b' is the total number of bands present in individual i but not in individual j and

'c' is the total number of bands presents in individual j but not in individual i.

The un-weighted pair group method with arithmetic mean (UPGMA) (Sneath and Sokal, 1973) was used to analyze and compare the individual genotypes and generates phenogram using NTSYS- pc version 2.02 (Rohlf, 2000). The neighbor joining (NJ) method (Saitou and Nei1987; Studier and Keppler, 1988) was used to compare individual genotypes and evaluate patterns of genotype clustering using Free Tree 0.9.1.50 Software (Pavlicek *et al.,* 1999). To further examine the patterns of variation among individual samples, a principal coordinated analysis (PCO) was performed based on Jaccard's coefficient (Jaccard, 1908). The calculation of Jaccard's coefficient was made with PAST soft ware version 1.18 (Hammer *et al.,* 2001). The first three axes were later used to plot with STATISTICA version 6.0 software (Statistica soft, Inc.2001).

5. RESULTS

5.1. Banding patterns and ISSR primers

Out of the ten primers tested initially, four of them gave relatively clear banding pattern and they were selected and used in this investigation (Table 3). The molecular weight of the bands amplified using the four primers were in the range of 100 bp to 1000bp (Figure 2). A total of 43 bands were scored from four primers in which the least (six bands) were scored from 818 and the highest number of bands were scored from primer 873 which was 15 bands, while primer 812 and 880 generated 12 and 10 bands respectively.

Table 3: Fingerprint patterns generated using four ISSR primers; selected for this study.

Selected primer	**Repeated motif**	**Number of clear scorable bands**	**Amplification pattern**	**ISSR gel observation**
812	$(GA)_8A$	12	Good	Good with small smear
818	$(CA)_8G$	6	Good with smear	Good with smear
873	$(GACA)_4$	15	Best	Good
880	$(GGAGA)_3$	10	Good	Good
		43		

Source: primer kit 900 (UBC 900); Single-letter abbreviations for mixed base positions.

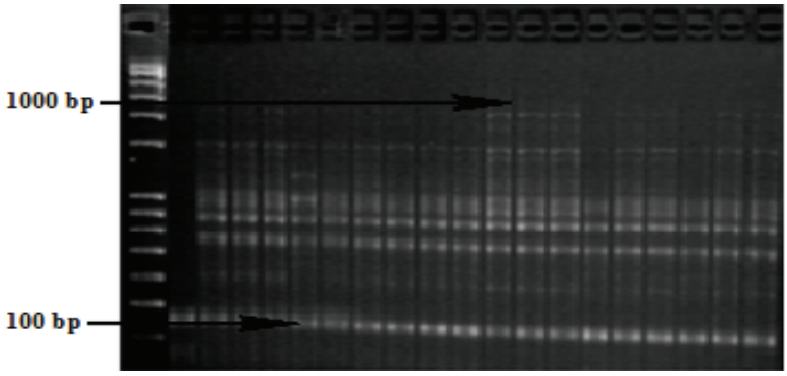


Figure 2: ISSR fingerprinting of primer 873. *Key: The first lane is DNA ladder and lane 1-18 represented accession from Amhara region (AMH 1-18, Table1).*

5.2. **Polymorphic loci and percent polymorphism**

The number of polymorphic loci ranges from six for 818 to 14 for 873. The di-nucleotide primers, namely 812 and 818 were observed to have eight and six polymorphic loci respectively. The tetra-nucleotide (873) and penta-nucleotides (880) were observed to have 14 and 9 polymorphic loci, respectively. The percent of polymorphic bands for 818, 873, 880 and 812, primers were 100%, 93.33%, 90% and 66.67%, respectively (Table 4). The highest gene diversity (0.37) and Shannon index (0.55) was shown by primer 873 and followed by primer 880 with gene diversity and Shannon index value of 0.36, 0.52, respectively.

In addition, among all populations studied, the Oromia population were found to have higher percentage polymorphism, 86.05% and followed by population from Amhara, 67.44%. The least percent polymorphism was observed for populations from Tigray, and SNNPR, with 34.88% and 20.93%, respectively.

5.3. Genetic Diversity as describe by Gene diversity and Shannon Index

Among cultivated safflower populations considered in the present study, Oromia showed higher gene diversity (h=0.32) and Shannon index (I=0.48) as compared to the other three populations of cultivated safflower. Safflower populations from Amhara regions showed 0.27 and 0.39 values of gene diversity and Shannon index, respectively. Tigray region also showed 0.16, 0.22 value of gene diversity and Shannon index respectively. The least gene diversity was showed by population from SNNPRs with gene diversity and Shannon index value of 0.08.and 0.11, respectively.

Table 4: Number of polymorphic loci, percent polymorphism, Nei's gen diversity and Shannon index of safflower with all primers.

Population	**No. of accessions**	**With all primers**			
		NPL	PP (%)	Gene diversity (h)	Shannon index (I)
Amhara	33	29	67.4	0.27±0.12	0.39±0.29
Oromia	28	37	86.1	0.32±0.15	0.48±0.21
SNNPRs	4	9	20.9	0.08±0.16	0.11±0.23
Tigray	5	15	34.9	0.16±0.12	0.22±0.31
Overall	70	**90**	**52.3**	**0.21±0.18**	**0.30±0.26**

Table 5: Number of polymorphic loci, percent polymorphism, Nei's genetic diversity and Shannon index of safflower with each primer.

Individual primers				
Primers	NPL	PP (%)	Gene diversity (h)	Shannon index (I)
812	8	66.7	0.24±0.20	0.36±0.28
818	6	100	0.23±0.05	0.39±0.06
873	14	93.3	0.37±0.12	0.55±0.17
880	9	90	0.36±0.17	0.52±0.22
Overall	**37**	**87.5**	**0.30±0.13**	**0.46±0.18**

NPL= number of polymorphic loci, PP (%) = percent polymorphism, h= Nei's gene Diversity and I= Shannon index information for each population and over all populations.

4.1. Analysis of Molecular Variance and Partitioning genetic diversity

When data are available from more than one population, it is usually of interest to evaluate the degree to which the total gene diversity partitions in to its within and between population components (Lynch and Milligan, 1994). In this study, analysis of molecular variance revealed that higher percentage of variation is attributed to variation

within populations (98.9%) and only 1.1% of the variation was attributed to difference among populations (Table 6).

Table 6: Analysis of Molecular Variance (AMOVA) of safflower populations in Ethiopia

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation	Fixation index	p
Among populations	3	14.235	0.04492Va	**1.1**	0.01	0.00
Within populations	66	271.022	4.10639Vb	**98.9**		0.00
Total	69	285.257	4.15131			

4.2. Clustering analysis

The four safflower populations (Amhara, Oromia, SNNPRs and Tigray) were study using ISSR markers. This generated data were used to construct a dendrogram using Un-weighted Pair-Group Method using Arithmetic Averages (UPGMA) and Neighbor-joining (NJ) in the NTSYS and free tree program, respectively. Both UPGMA and NJ based dendrogram for individual safflower appeared to recover different groups. Safflower individual assembled from different localities and regions observed to be spread all over the trees without forming strict grouping based on their geographic origin. However, in few cases some individual from Amhara and Oromia tends to form separate groups. The majority of the groups observed in UPGMA and NJ trees were intermixed individuals from SNNPR and Tigray with Oromia and Amhara populations. Generally, safflower clearly observed to recover different groups in both UPGMA and neighbor

joining analysis except in the case of eight individuals assembled from Amhara forming separate group in both UPGMA and NJ. Both trees recovered almost the same tree topology with similar groupings, although few individuals appeared to escape from groups in both cases.

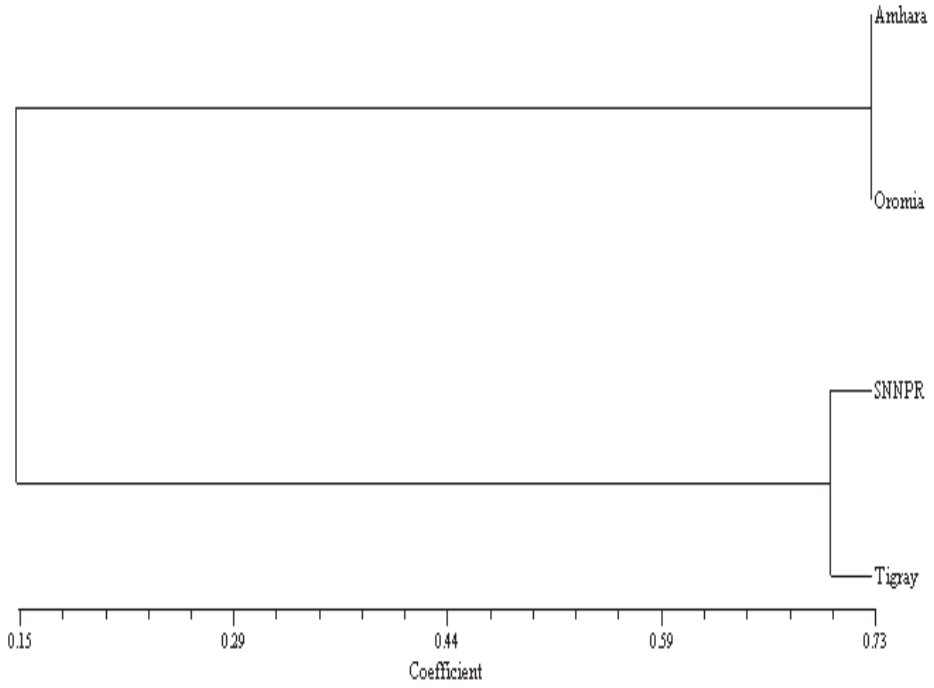
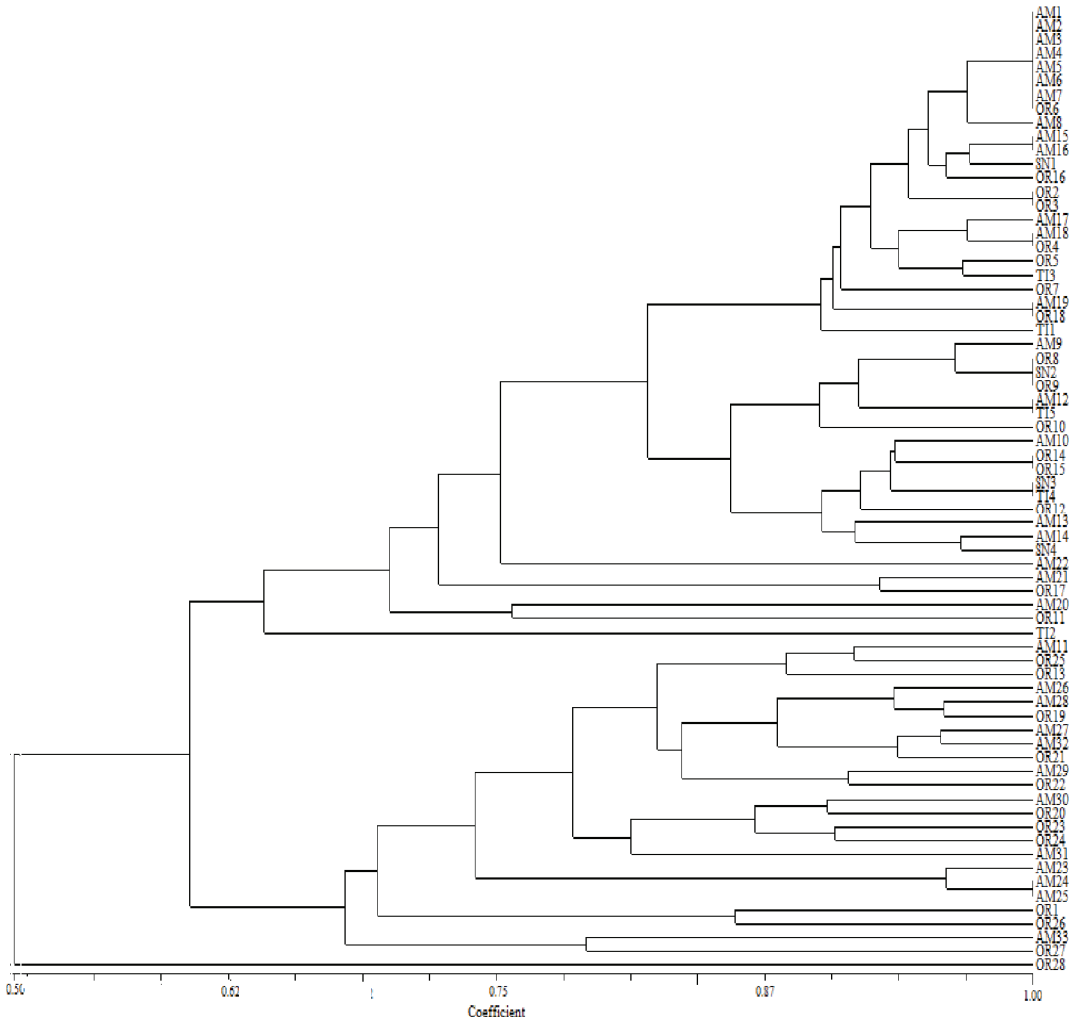
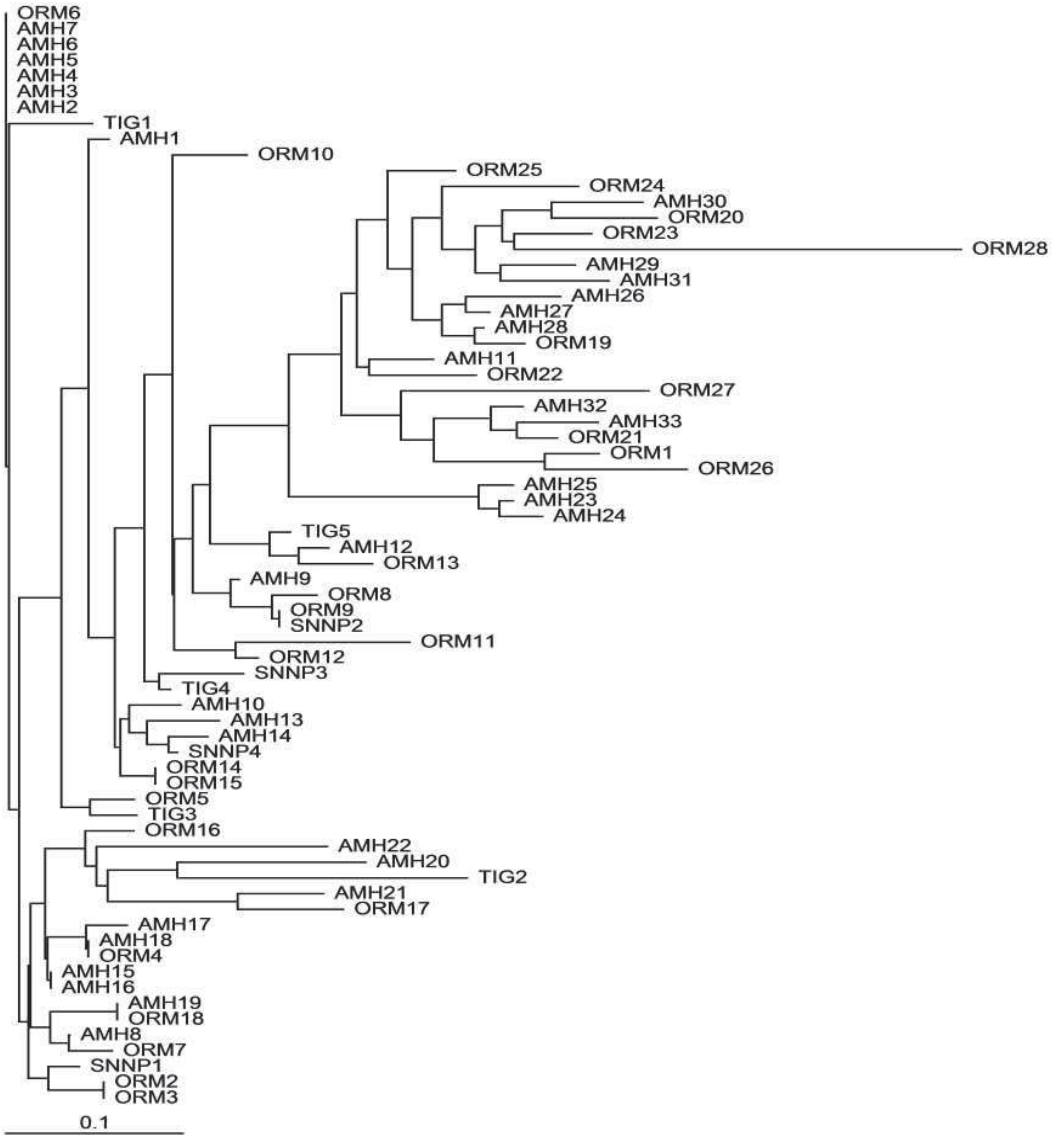


Figure 3: Dendrogram for four safflower populations obtained using UPGMA 43 ISSR-PCR bands amplified by two di-nucleotide (810 and 818), one tetra-nucleotide (873) and one penta-nucleotide (880) primers. The UPGMA algorithm is based on Jaccard's coefficients obtained after pair wise comparison of the presence-absence fingerprint.



Key*:-AM=Amhara, OR= Oromia and TI=Tigray (from individual accessions refer symbols given in table 1)*

Figure 4: UPGMA tree of 70 safflower individuals based on 43 ISSR-PCR bands amplified by four di-nucleotide (810 and 818), one tetra-nucleotide (873) and one penta-nucleotide (880) primers. The UPGMA tree is based on Jaccard's coefficients obtained after pair wise comparison of the presence-absence fingerprint data.



Key: - *AMH=Amhara, ORO=Oromia, SNN=SNNPRs and TIG=Tigray(from individual accessions refer symbols given in table 1)*

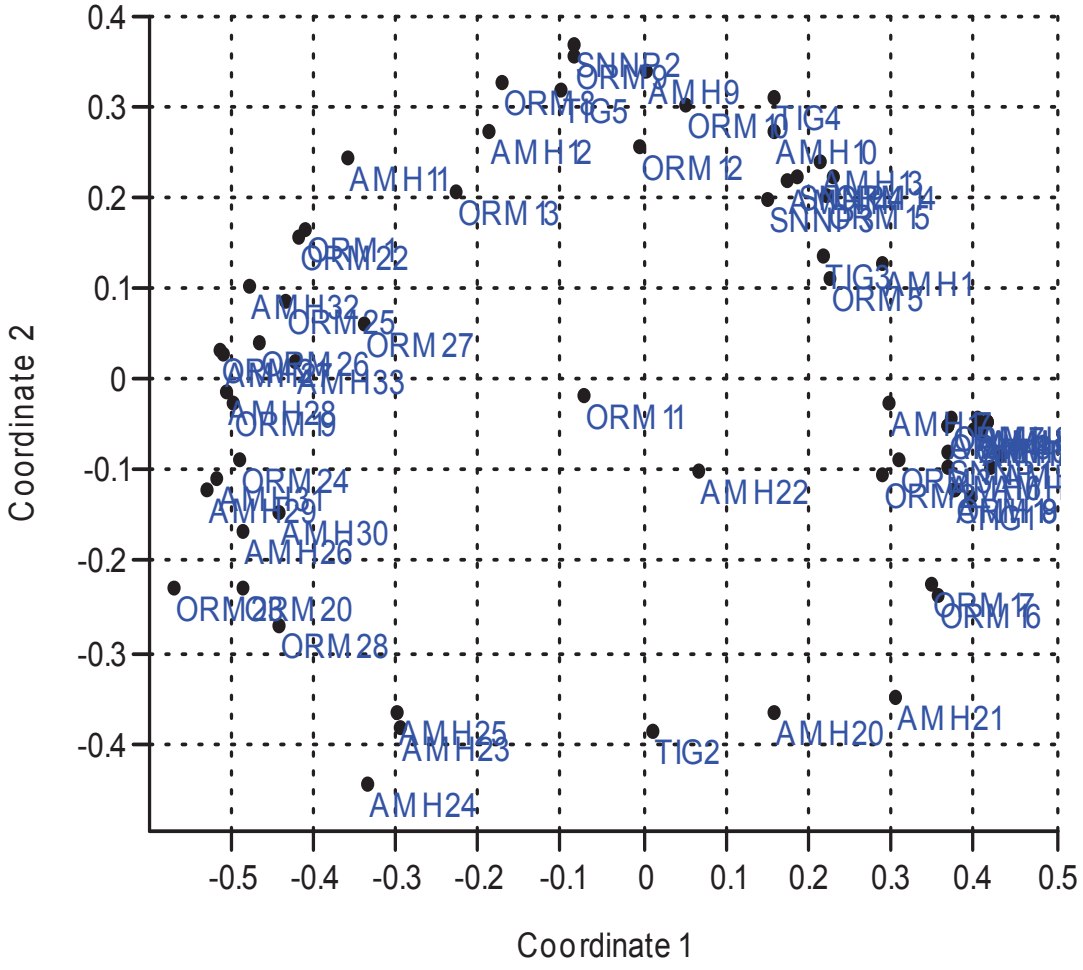
Figure 5: Neighbor-joining analysis of 70 individuals based on 43 ISSR- PCR bands amplified by four di-nucleotide (810 and 818), one tetra-nucleotide (873) and one penta-nucleotide (880) primers. The neighbor joining algorithm is based on Jaccard's coefficients obtained after pair wise comparison of the presence-absence fingerprinting

4.3. PCO Analysis

All the data obtained using four ISSR primers were used in PCO analysis using Jaccard's coefficients of similarity. The first three coordinates of the PCO having Eigen-values of 8.46, 2.92 and 1.92 with variance of 27.4%, 9.5% and 6.2%, respectively (Table 7) used to show the grouping of individuals using two and three coordinates. The four populations observed to form separate cluster and some of population were dispersed all over the 2D and 3D space. However, some individual sample from SNNPRs observed to be intermixed with Oromia and Amhara populations. Using two coordinates (Figure 6) almost similar result was observed like that of three coordinates.

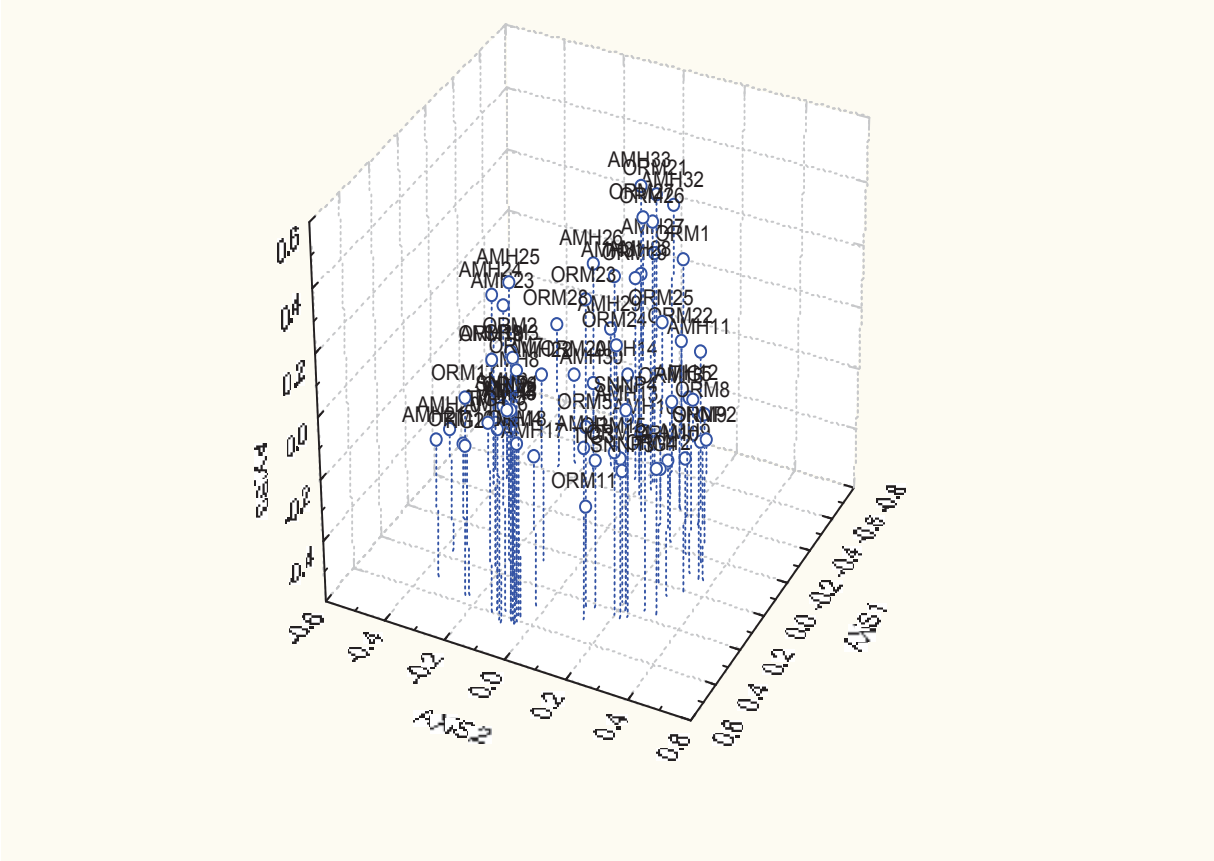
Table 7: Eigen value and explained variance in the PCO using characters used to classify 70 safflower accessions using ISSR markers.

Principal coordinate	**Eigen value**	**Explained variance (%)**
1	8.46	27.4
2	2.92	9.5
3	1.92	6.2



Key: - *AMH=Amhara, ORO=Oromia, SNN=SNNPRs and TIG= Tigray(from individual accessions refer symbols given in table 1)*

Figure 6: Two dimensional representations of principal coordinate analysis of genetic relationships among 70 individuals from four populations of safflower based on Jaccard's similarity coefficients.



Key: - *AMH=Amhara, ORO=Oromia, SNN=SNNPRs and TIG= Tigray(from individual accessions refer symbols given in table 1)*

Figure 7: Three dimensional (3D) representation of principal coordinate analysis of genetic relationships among 70 individuals of four population safflower inferred from similarity matrix using the Jaccard's index.

5. DISCUSSION

5.1. Use of ISSR marker for safflower genetic diversity study

There are more number of published reports on the use of ISSR markers than other molecular marker techniques to analyze the genetic diversity of cultivated safflower. Because ISSR markers have a better reproducibility and a much greater number of total polymorphic and discriminate fragments than RAPDs and easier detection and at lower costs than AFLPs and simpler to use than the SSR technique and less restrictive than RFLPs and may offer considerable variation among species (Mattioni, *et al*., 2002). Therefore, ISSR technique was employed in the present study. This is believed to be the first report on genetic diversity study of Ethiopia cultivated safflower using molecular markers.

Evaluation and identification of germplasms using ISSR markers are playing an important role in studies of plant genetics and breeding. ISSR marker has been used successfully in estimating the genetic diversity in rice (Gezahagn Girma, 2007), Coffee (Solomon Balami, 2007), Sesame (Dagmawi Teshome, 2011) and lentil (Edossa Fikiru *et al.*, 2011). This paper also confirms that ISSR marker is efficient in detecting polymorphism within and among populations of safflower. Hence, a total of 43 bands were amplified from only four primers, which is an average of 10.8 bands per primer. The result also showed 52.3% polymorphism for populations assembled from four regions (Amhara, Oromia, SNNPRs and Tigray). Comparisons between regions showed that, Oromia region has the highest percent polymorphic (86.1%) and followed by Amhara (67.4%) while populations from Tigray and SNNPRs were observed to have the least percent polymorphism with 34.9% and 20.9%, respectively. However, in the view

of the relatively very small number of accession included in the present study from Tigrary and SNNPR, it is not possible to make strong conclusion about the percent polymorphism in safflower from the two regions. Further study with large sample size is necessary.

The separate analysis of percent polymorphic per primer showed 100% polymorphism for primer 818 followed by the tetra-nucleotide primer 873 (93.3%). Primers 880 and 812 showed high to intermediated polymorphism with 90% and 66.7% polymorphism, respectively. Comparatively, gene diversity (0.37) and Shannon index (0.55) were shown by primer 873 and followed by primer 880 has showed high gene diversity (0.36) and Shannon index value (0.52). The indication of this result is that primers 873 and 880 were much better for detecting genetic diversity of cultivated safflower. In addition, higher average percent polymorphism per primer (87.5%) was observed in this study which shows the importance of ISSR markers to reveal the molecular diversity within and among populations of safflower.

Yang *et al.*, (2007) also got a total of 429 bands using 22 ISSR primers which showed high polymorphism (82.7%) and higher average percent polymorphic per primer was observed (73.3%). In line with the results of the present study, similar result was obtained by Seyed *et al.* (2010), whereby higher percent of polymorphism (96%) was revealed with ISSR marker, which was higher than that of RAPD (81.08%). Moreover, the average number of polymorphic bands detected by each ISSR primer in Seyed *et al.* (2010) study was 5.3, which is higher than that of the short RAPD primer (4.6). The cluster results also showed that all the accessions could be distinguished by ISSR markers, which laid a solid foundation for evaluating the genetic diversity of safflower.

Overall, the present study and previous investigations on safflower with ISSR markers confirm the usefulness of this marker to study genetic diversity within and among populations of safflower.

5.2. Diversity within and among populations of safflower

An understanding of the partitioning of genetic variation within crop gene pools can provide insight into the evolution of crop lineages. In addition to its implication to conservation, partitioning the genetic variation in to its components has significant impact in the future breeding and conservation plan (Yamasaki *et al*., 2005). In this investigation, the AMOVA analysis showed very high significant genetic variation within population (98.9%) with the least variation among populations (1.1%). The one reason for high significant genetic variation within safflower population may be due to the pollination nature of safflower which is insect pollinators including bees and birds might be attracted to its colorful flower. Not only the pollination nature of safflower, but also the existence of gene flow between the cultivated safflower populations may be the main factor. Hence, the presence of extensive safflower seeds exchange system within and among different regions and also the harvest and storage mechanism of safflower by farmers have their own effect on the variation of cultivated of safflower. In other words, cultivated safflower seeds are available on the local market. Due to this reasons farmers could not store safflower seeds or landraces in their home for next growing season. If the farmer tends to stored seed and cultivate the seed and harvest, it may develop the variation among the population rather than within population. Hamrick and Godt (1989) stated regards the genetic diversity, the biological characteristics of a species, population structure, mating system (outbreeding and selfing) and multiple evolutionary process

(genetic drift: bottleneck and founder effect) are determine the level and pattern of genetic variation. Generally, According to Vavilov (1951) reported the presence of high genetic diversity in cultivated safflower considered Ethiopia to be a centre of diversity and the site of origin of cultivated safflower. According to the result of the present study, high genetic diversity within populations (Oromia and Amhara) indicate that the most effective strategy for preserving genetic variation would be to conserve a large number of individuals for more diverse population(s) and simultaneously for improvement programs would be the best approach.

5.3. Clustering analysis and relationships of safflower populations

A cluster analysis (UPGMA and neighbor joining) was used to construct a dendrogram in order to see relationships among all individuals (Figure 4 and Figure 5). Only few individuals form a discrete cluster following geographic origin. Samples from Tigray and SNNPRs were dispersed into the various groups appeared to be at random with Oromia and Amhara safflower populations. There are also intermixed groups for all the four populations. Yang *et al.* (2007) has got similar result, whereby the 24 accession from Asia were scattered all over the tree. Comparatively, the accessions originated from Europe were relatively assembled. In the same manner, Mahasi *et al*. (2009) explained the UPGMA cluster analysis showed that the 36 accessions were grouped into eight groups with individuals from different countries. Cluster 2 had the largest number of accessions originating from India, USA, Australia and Bangladesh. Cluster 3 had the second largest number of accessions from India, USA and Mexico. But proportionally accessions from India were represented at a higher level in both cluster 2 and 3. In line

with the present study, the dispersion into the various groups appeared to be at random through a few accessions formed distinct clusters based on geographic origin.

Generally, in this investigation, there are not significant regional differentiations between safflower populations. Hence, low differentiation among population from different regions

In the present study also, obtained PCO results based on two and three coordinates showed similar results except some population dispersed overall the 2D and 3D. Few groups from the PCO results have showed recovered similar groupings with that of the UPGMA and NJ results. The fact that they are grouped in the same cluster implies their close relationship in genetic characters. Therefore, their difference in cluster could imply their being originated from different sources.

6. CONCLUSION

- ✓ Previous researches based on molecular markers on cultivated safflower revealed there is genetic diversity in safflower germplasm
- ✓ It is evident from this study that the ISSR assay can be useful for safflower germplasm characterization with regarding to its genetic diversity. The results of the present study which is based on ISSR molecular markers confirm the existence of higher polymorphism based on ISSRs molecular marker. All the four primers showed polymorphism ranging from 66.7 percent to 100 percent (average 87.5%). On the average, 10.8 bands per primer were produced. Higher genetic variation is found within populations than among populations in Ethiopia safflower and also in the present study there is no clear grouping among accession collected from the same region
- ✓ In conclusion, through ISSR analysis, all the diversity parameters confirm that there is higher genetic diversity in safflower populations of Ethiopia that deserve conservation attention.

7. RECOMMENDATIONS

- ✓ This study is not exhaustive in terms of sample size, area coverage and molecular markers. Hence, more survey and sample collection and analysis with additional markers have to be carried out.

- ✓ The significant variation between populations obtained in this study has an important implication for its conservation and sustainable use. In other words, future germplasm collections should represent all growing regions by giving special emphasis to areas that are not yet represented or underrepresented. This helps to increase the range of genetic variation among accessions and thus facilitates generation of a well represented core collection. Hence, the IBC and other stakeholders should work jointly to conserve the safflower population in Ethiopia. Generally, in view of the existing genetic diversity, breeders should give attention for Oromia and Amhara populations to improve the productivity of safflower in Ethiopia.

- ✓ The ISSR molecular marker techniques used in this study revealed high genetic polymorphism in safflower. Other marker system such as SSR and AFLP along with sequence based markers should be utilized to further reveal the level of diversity in the cultivated safflower as well as its relationship with wild safflower.

8. REFERENCES

Alonso-Blanco, C., Peeters, A. J., Koornneef, M., Lister, C., Dean, C., van den Bosch, N., Pot, J. and Kuiper, M. T. (1998). Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. *J. Plant,* **14:**259–271.

Amiri, R. M., Yazdi-Samadi, B., Ghanadha, M. R. and Mishani, C. (2001). Detection of DNA polymorphism in landrace populations of safflower in Iran using RAPD-PCR technique. *Iranian J. Agric. Sci.,* **32:** 737-745.

Ash, G. J., Raman, R. and Crump, N. S. (2003). An investigation of genetic variation in *Carthamu lanatus* in New South Wales, Australia, using inter simple sequence repeats (ISSR) analysis. *Weed Res.,* **43:**208–213.

Ashri, A. (1973*). Divergence and Evolution in the Safflower Genus Carthamus L.* Hebrew University of Jerusalem, Rehovot, Israel.

Ashri, A., Knowles, P. F. (1960). Cytogenetics of safflower *Carthamus* L. species and their hybrids. *J. Agron.***, 52:**11–17.

Ashri. A, (1975). Evaluation of the germplasm collection of safflower *Carthamus tinctoriu* V: Distribution and regional divergence for morphological characters. *Euphytica., 24:651*–659.

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

Bradley, V. L., Guenther, R. L., Johnson, R.C., and Hannan, R. M. (1999). Evaluation of Safflower germplasm for ornamental use. **In:** *Perspectives on new crops and uses*. Pp. 433-435, (Janik, J. ed.), ASHS Press: Alexandria, USA.

Chapman, M. A., Chang, J. C., Weisman, D., Kesseli, R.V. and Burke, J. M. (2007). Universal Markers for comparative mapping and phylogenetic analysis in the Asteraceae (Compositae). *Theor. Appl. Genet.,* **115:** 747-755.

Chavan, V. M. (1961). Niger and Safflower, Hyderabad. *Indian Central Oilseeds Commi*ttee Publ*., **8:*** 57-150.

Dagmawi Teshome (2011). *Genetic diversity of saseme in (sesamum indicum L.) germplasm collection as revealed by ISSR marker: implication for conservation and improvement.* M. Sc. Thesis. Addis Ababa University, Addis Ababa.

Dajue, L. and Mündel, H. (1996). *Safflower, Carthamus tinctorius L. Promoting the Conservation and use of Underutilized and Neglected Crops*. Institute of Plant Genetics and Crop Plant Research, International Plant Genetic Resources Institute, Rome Italy.

Edossa Fikiru, Kassahun Tesfaye and Endashaw Bekele (2010). A comparative study of morphological and molecular diversity in Ethiopian lentil (*Lens culinaris* Medikus) landraces. *Afr. J. Plant Sci.,* **4:** 242-254.

Edwards, Sue B. (1991). Crops with wild relatives found in Ethiopia. **In:** *Plant Genetic Resources of Ethiopia.* Pp., 54-63 (Engels, J. M. M., Hawkes, J. G. and Melaku Worede, eds). Cambridge Univ. Press, Cambridge.

Ekin, Z. (2005). Resurgence of safflower (*Carthamus tinctorius* L.) utilization: *J. Agron.*, **42:** 83-87.

Excoffier, L., Smouse, P. E. and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondria DNA restriction sites. *Genetics*. **131:**479-491.

Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Longman, Harlow, UK.

Farooq, S. and Azam, F. (2002). Molecular markers in plant breeding-I: Concepts and Characterization. *J. Biol. Sci.,* **10:**1135-1140.

Flynn, C. and Bergman, J. (2001). Analytical chemistry methods used in the research and development of safflower varieties for the United States Northern Great Plains region. Proceedings of the 5th International Safflower Conference. Williston, North Dakota and Sidney, Montana.

Gezahagn Girma (2007). *Relationship between wild rice species of Ethiopia with cultivated rice based On ISSR marker*. MSc. Thesis, Addis Ababa University, Ethiopia.

Ghislain, M., Spooner, D. M., Rodríguez, F., Villamon, F., Núñez, C., Vásquez, C. and Bonierbale, M. (2004). Selection of highly informative and user-friendly microsatellites (SSRs) for genotyping of cultivated potato. *Theor. Appl. Gene.,* **108:** 881–890.

Godwin, I. D., Aitken, E. A. B. and Smith, L. W. (1997). Application of inter simple sequence repeat (ISSR) markers to plant genetics. *Electroph.,* **18:**1524–1528.

Gyulai, J. (1996). Market outlook for safflower. **In:** *Proceedings of North American Safflower Conference*. pp. 15, (Mundel, H. H., Braun, J. and C. Daniels, eds).

Hammer, O., Harper, D. A. T. and Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronic.* **4:**9.

Hamrick, J. L. and Godt, M. J. W. (1989). Allozyme diversity in plant species. **In:** *Plant Population Genetics, Breeding and Germplasm S*

ource. Pp. 43-63, (Brown, A. H. D., Clegg, M. T., Kahler, A. L. and Weir, B. S. eds.) Sunderland, MA: Sinauer.

Han, Y. and Li, D. (1992). Evaluation of safflower (*Carthamus tinctorius L.*) germplasm-Analysis in fatty acid composition of seeds of domestic and exotic safflower varieties. *Bot Res* **6:**28–35.

Hanelt, P. (1961). ZurKenntnis von *Carthanus tinctorius L*. Kulturp flanze., **9:**114–145.

Helm, J. L., Schneiter, A. A., Riveland, N. and Bergman, J. (1991). *Safflower Production*. North Dakota State University, Fargo, USA.

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

Jaradat, A. A. and Shahid, M. (2006). Patterns of phenotypic variation in a germplasm collection Of *Carthamus tinctorius* L. from the Middle East. *Euphytica*., **53:**225–244.

Johnson, R. C., Kisha, T., Foiles, C. and Bradley, V. (2005). Characterizing safflower germplasm with AFLP molecular markers. *VI Internl. Safflower Conf.,* 3-8.

Jones, C. J., Edwards, K. J., Castaglione, S., Winfield, M.O., Sala, F., Wiel, C., van d., Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P. and Karp, A. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in Plants b.y a network of European laboratories. *Mol. Breed.* **3:** 381–390.

Kaftka, S. (1965). *Safflower Production in California.* Oakland, CA: University of California, Communication Services Publications, California.

Karp, A., Kresovich, S., Bhat K. V., Ayand, W. G. and Hodgkin, T. (1997). *Molecular Tools in Plant Genetic Resources Conservation: A guide to the technologies.* International Plant Genetic Resources Institute, Rome, Italy.

Kearsey, M. J. and Pooni, H. S. (1996). *The Genetical Analysis of Quantitative Traits.* Chapman and Hall, London.

Khidir, M.O. and Knowles, P. F. (1970). Cytogenetic studies of *Carthamus* species (Compositae) with 32 pairs of chromosomes. II. Intersectional hybridization. *Can. J. Genet. Cytol.,* **12:** 90–99.

Knowles, P. F. (1980). *Hybridization of Crop Plants: Safflower*. American Society of Agronomy-Crop Science of America. USA

Knowles, P. F., Miller, M. D., Henderson, D. W., Foy, C. L., Carlson, E. C., Klisiewicz, J. M., Goss, J.R., Jones, L.G. And Edwards, R.T. (1965). *Safflower-Circular.* University of California, Davis, California.

Knowles, P.F. (1969). Centers of plant diversity and conservation of crop germplasm: Safflower. *Econ. Bot.,* **23:**324–329.

Landaua, S., Friedmana, S., Brennera, S., Bruckentalb, I., Weinbergc, Z.G., Ashbellc, G., Henc,Y., Dvasha, L., and Lehsem, Y. (2004). The value of safflower (*Carthamus tinctorius* L.) Hay and silage grown under Mediterranean conditions as forage for dairy cattle. *Livestock Prod. Sci.,* **88:** 263-271.

Li, D., M. Zhou, and Rao, V. R. (1993). *Characterization and Evaluation of Safflower Germplasm*. Geological Publishing House, Beijing.

Litt, M, Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a di-nucleotide repeat within the cardiac muscle actin gene. *Amr. J. Hum. Genet.,* **44:** 397–401.

Lynch, M. and Milligan, B.G. (1994). Analysis of population genetic structure with RAPD markers. *Mol. Ecol*., **3:** 91-99.

Mahasi, M. J., Wachira, F. N., Pathak, R. S. and Riungu T. C. (2009). Genetic polymorphism in exotic safflower (*Carthamus tinctorious* L.) using RAPD markers. *J. Plant Bre. and Crop Sci.,* **1:** 8-12.

Malik, A. And Srivastava, A.K. (2009). Karyotypic analysis of different populations of *Carthamus tinctorius* L. (Asteraceae). *Genetica.* **40:** 84-88.

Mattioni, C., Casasoli, M. and Gonzalez, M.(2002). Comparison of ISSR and RAPD markers to characterize three *Chilean Nothofagus* species. *Theor. Appl. Genet.,* **104:** 1064 – 70.

Morgante, M., Hanafey, H. and Powell, W. (2002). Microsatellites are preferentially associated with non repetitive DNA in plant genome. *Nature Genet,* **30:**194–200.

Mündel, H.H., Blackshaw, R.E., Byers, J.R., Huang, H.C., Johnson, D.L., Keon, R., Kubik, J., McKenzie, R., Otto, B., Roth, B., and Stanford, K. (2004). *Safflower Production in the Canadian Prairies*. Agriculture and Agri-Food Canada, Lethbridge Research Centre, Lethbridge.

Nagaraj, G. (1993). *Safflower Seed Composition and Oil Quality: A Review*. Third Internl. Safflower .Beijing, China.

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

Powell, W., Machray, G. C.and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Tren. Plant Sci.,* **1:215**–222.

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

Rohlf, F.J. (2000). NTSYS-pc, Numerical Taxonomy and Multivariate Analysis System, Version 2.1. Exeter Software, New York.

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

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

Schierwater, B. and Ender, A. (1993). Different thermostable DNA polymerases may apply to different RAPD products. *Nucl. Acids. Res.,* **21:**4647–4648.

Sehgal, D. and Raina, S. N. (2005). Genotyping safflower (*Carthamus tinctorius*) cultivars by DNA fingerprints. *Euphytica.*, **146:**67–76.

Seyed, A. S., Seyyed, S. P., Mohmmad, T. and Mahmoud K. (2009). Assessment of genetic variation among safflower (*Carthamus tinctorius* L.) accessions using agro- morphological traits and molecular markers. *J. Food, Agri. Enviro.*, **8:** 616 – 625.

Smith, J.R. (1996). *Safflower.* AOCS Press, Champaign, USA.

Smith, D.N. and Devey, M. E. (1994). Occurrence and inheritance of microsatellites in Pinusradia. *Genome,* **37:** 977–983.

Sneath, P. H. A. and Sokal, R. R. (1973). *Numerical Taxonomy*. Freeman. San Francisco.

Solomon Balami (2007). *Genetic Diversity Analysis of the Wild Coffea arabica L. Populations from Harenna Forest, Bale Mountains of Ethiopia, Using Inter Simple Sequence Repeats (ISSR) Marker*. M.Sc. Thesis, Addis Ababa University, Addis Ababa.

Soranzo, N., Provan, J. and Powell, W. (1999). An example of microsatellite length variation in the mitochondrial genome of conifers. *Genome,* **42:** 158-161.

Stat soft, Inc (2001). Statistica data analysis system, Statistica software.

Stobart, A. K. and Stymne, S. (1985).The interconversion of diacylglycerol and phosphatidylcholine during triacylglycerol production in microsomal preparations of developing cotyledons of safflower (*Carthamus tinctorius* L.). *Biochem. J.,* **232:** 217-221.

Vavilov, N. I. (1951). *The Origin, Variation, Immunity and Breeding of Cultivated Plants*. The Ronald Press Co., New York.

Vilatersana, R., Garnatje, T., Susanna, A. and Garcia-Jacas, N. (2005). Taxanomic problems in *Carthamus* (Asteraceae): RAPD markers and sectional classification. *J. Bot. Linn. soc.,* **147:**375–385.

Vithanage, V., Anderson, K. A. and Thomas, M. (1995). Use of molecular markers in crop improvement of Macadamia. *The Sixth Conf. of the Australasian Council* on Tree and Nut Crops. Lismore, Australia.

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

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). *DNA Fingerprinting in plants. Principles, Methods, and Applications*. 2nd edition. CRC Press, Tylor and Francis group. Boca Raton.

Weising, K., Nybom, H., Wolff, K. and Meyer, W. (1995). *DNA Fingerprinting in Plants and Fungi.* CRC Press, Boca Rato.

Weiss, E. A. (1971). *Castor, Sesame and Safflower*. Barnes and Noble, Inc., New York.

Weiss, E. A. (2000). *Oil Seed Crops*. Blackwell Science Ltd, Oxford, UK.

Weiss, E.A. (1983). *Oilseed Crops*. Longman Inc., New York.

Williams, J.G.K., Hanafey, M.K., Rafalski, J. A. and Tingey, S.V. (1993). Genetic analysis using random amplified polymorphic DNA markers. *Meth. Enzy. mol.,* **218:** 705–740.

Wu, K., R., Jones, L., Dannaeberger and Scolnik, P.A. (1994). Detection of microsatellite polymorphisms without cloning. *Nucl. Acids Res*., **22:** 3257–3258.

Yamasaki , M. , M. I., Tenaillon , I. V., Bi , S. G., Schroeder , H., Sanchez- Villeda , J. F., Doebley, B. S. and McMullen, M. D. (2005). A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell,* **17:** 2859 – 2872.

Yang, Y.X., Wu, W., Zheng, Y. L., Chen, L., Liu, R. J. and Huang, C.Y. (2007). Genetic diversity and relationships among safflower (*Carthamus tinctorius* L.) analyzed by inter-simple sequence repeats (ISSRs). *Genet. Resour. Crop. Evol*., **54:1043**–1051.

Yeh, F.C., Yang, R.C., Boyle, T.J.B., Ye Z. and Mao, J.X. (1999). POPGENE, the user friendly Shareware for population genetics analysis, version 1.31, Molecular Biotechnology center, University of Aleberta, Canada.

Zhang, L.P. (1997). *Safflower: a Versatile Plant*. IV Internl. Safflower Conf. USA.

Zhang, Z. W. (2000). Studies on genetic diversity and classification of safflower (*Carthamus tinctorius* L.) germplasm by isozyme techniques. *J. Plant Gene. Res.,* **1:** 6–13.

Zhaomu, W. and Lijie, D. (2001). *Current Situation and Prospects Of Safflower Products Development In China.* 5th Internl. Safflower Conf. USA.

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

Zohary, D. and Hopf, M. (2000). *Domestication of Plants in the Old World: The Origin and Spread of Cultivated Plants in West Asia, Europe and the Nile Valley*. Oxford University Press, Oxford, UK.

9. APPENDICES

Appendix-1 DNA extraction protocol

1. Introduction

This protocol describes a modified CTAB method for the isolation of DNA from plant material. A maximum of 24 samples can be isolated at once using this protocol.

2. Materials

CTAB	-	3 x 700 µl per sample
ß-Mercapto-Ethanol		
Chloroform		
Isopropanol	-	100 %, 4°C
Ethanol	-	100 % and 70% in sterile ddH_2O
TE	-	1x, p.a. grade
NH_4Ac	-	7.5 M Solution, sterile
NaAc	-	3 M Solution, sterile
Water bath	-	65°C
Centrifuge	-	4°C
Cut 1000 µl tips	-	Sterile

3. Procedure

1. Poor CTAB solution (700 µl per sample) in a 15ml-tube and add 0.2 vol % Mercapto-ethanol (use fume hood!). Mercapto-ethanol is stored at 4°C.
2. Aliquot CTAB in 1,5 ml Eppendorf-caps and warm in water bath up to 65°C.
3. Weigh in 100 mg fresh leave material (50mg dry material) per sample. Pulverise thoroughly using a clean mortar and pestle. For fresh material add liquid nitrogen or quartz sand for dry material. First grind down slightly, then more powerful (cells have to be crashed). Use safety goggles!
4. Transfer the powder into an Eppendorf cap with warm CTAB solution immediately (use a new, clean spatula for each sample)
5. Add 700 µl of warm CTAB solution to the powdered sample (open the caps carefully), dissolve the powder and incubate the sample for 30 minutes at 65°C.
6. Centrifuge for 5 minutes at 15000 rpm.
7. Transfer the supernatant (only clear liquid) in a new Eppendorf-cap. Use blue pipette tips which are cut.
8. Add new CTAB solution (700 µl) to the tissue pellet and stir slightly with a new 1000 µl pipette tip, incubate 30 min at 65°C. Step 6 and 7 are repeated. The same is carried out for a third extraction. Each fraction proceeds with step 9 and is treated separately.
9. Add 600 µl Chloroform to the cap with supernatant and shake carefully a few times upside down. This chloroform step should be carried out immediately.

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

11. Centrifuge for 5 min at 15000 rpm.

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

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

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

15. Centrifuge 10 min at 15000 rpm.

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

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

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

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

20. Dissolve pellet in 100 µl TE (1x, p.a. grade) and store at 4°C. (BREAK POSSIBLE)
21. Add cooled 7.5 M NH_4Ac-solution (4°C, half of the solution volume). Mix carefully.
22. Add cool Ethanol 100 % (double of the solution volume). Mix carefully. Freeze for more than 2 h at -20°C. (BREAK POSSIBLE)
23. Centrifuge 30 min at 15000 rpm. Aspirate fluid carefully.
24. Add 200 µl Ethanol 70%. Rinse the inner cap surface by turning the cap.
25. Centrifuge 10 min. at 15000 rpm. Aspirate liquid and dry pellet at room temperature. Dissolve the pellet in 100 µl TE (1x, p.a. grade)
26. Repeat steps 21 to 24 with 3 M Na-AC-solution (4°C, half the volume) then proceed with step 28.
27. Centrifuge 10 min. at 15000 rpm. Aspirate liquid and dry pellet at room temperature. Dissolve the pellet in 100 µl TE (1x, p.a. grade)

- Cleaning the mortar and pestle:
    - Rinse the mortar and pestle with water
    - Clean the mortar and pestle in a 1:10 Klorix-bath for 24 hours
    - Rinse with ddH_2O

Autoclave the mortar and pestle wrapped in aluminium foil at 134°C

Appendix-2: Gel pictures of four primers (812, 818, 873 and 880).

- Gel picture of primer 812:

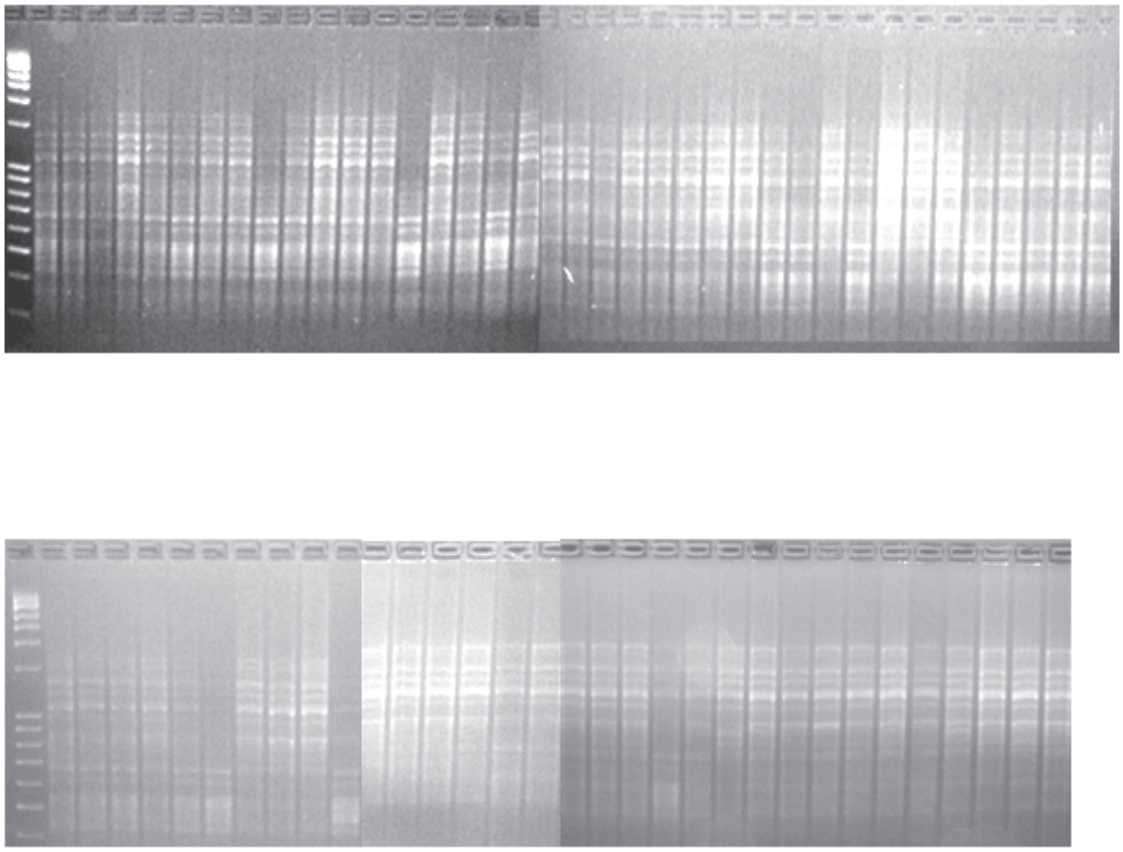


Figure 1: ISSR fingerprinting of primer 812 *Key: The first lane is DNA ladder, lanes 1-33 samples from Amhara, 34-61 samples from Oromia, 62-65 samples from SNNRP and 66-70 samples from Tigray. The accessions sequences correspond to that given in table 1*

❖ Gel picture of primer 818:

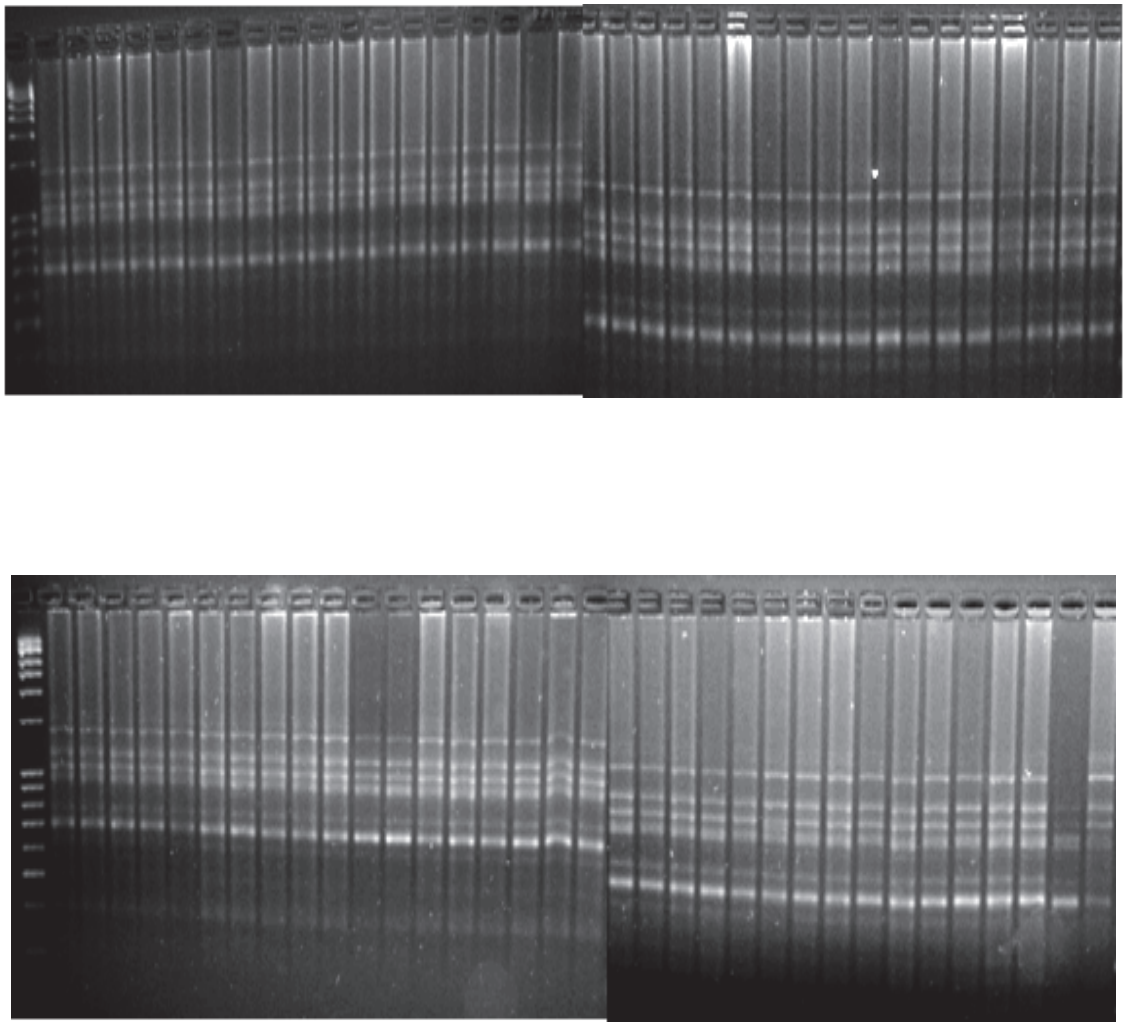


Figure 2: ISSR fingerprinting of primer 818. *Key: The first lane is DNA ladder,lanes 1-33 samples from Amhara, 34-61 samples from Oromia, 62-65 samples from SNNRP and 66-70 samples from Tigray. The accessions sequences correspond to that given in table 1*

❖ Gel picture of primer 880

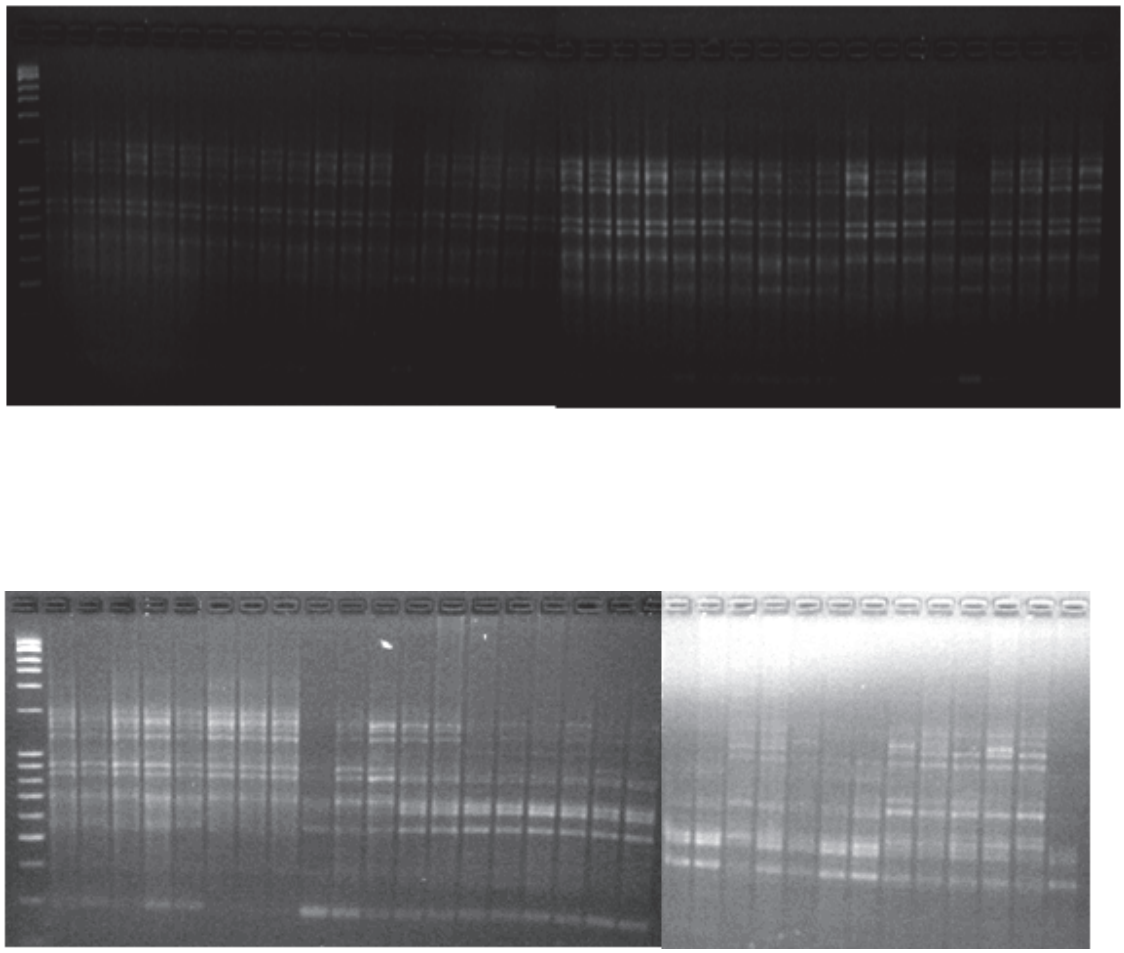


Figure 3: ISSR fingerprinting of primer 880. *Key: The firs tlane is DNA ladder, lanes 1-33 samples from Amhara, 34-61 samples from Oromia, 62-65 samples from SNNRP and 66-70 samples from Tigray. The accessions sequences correspond to that given in table 1*

❖ Gel picture of primer 873

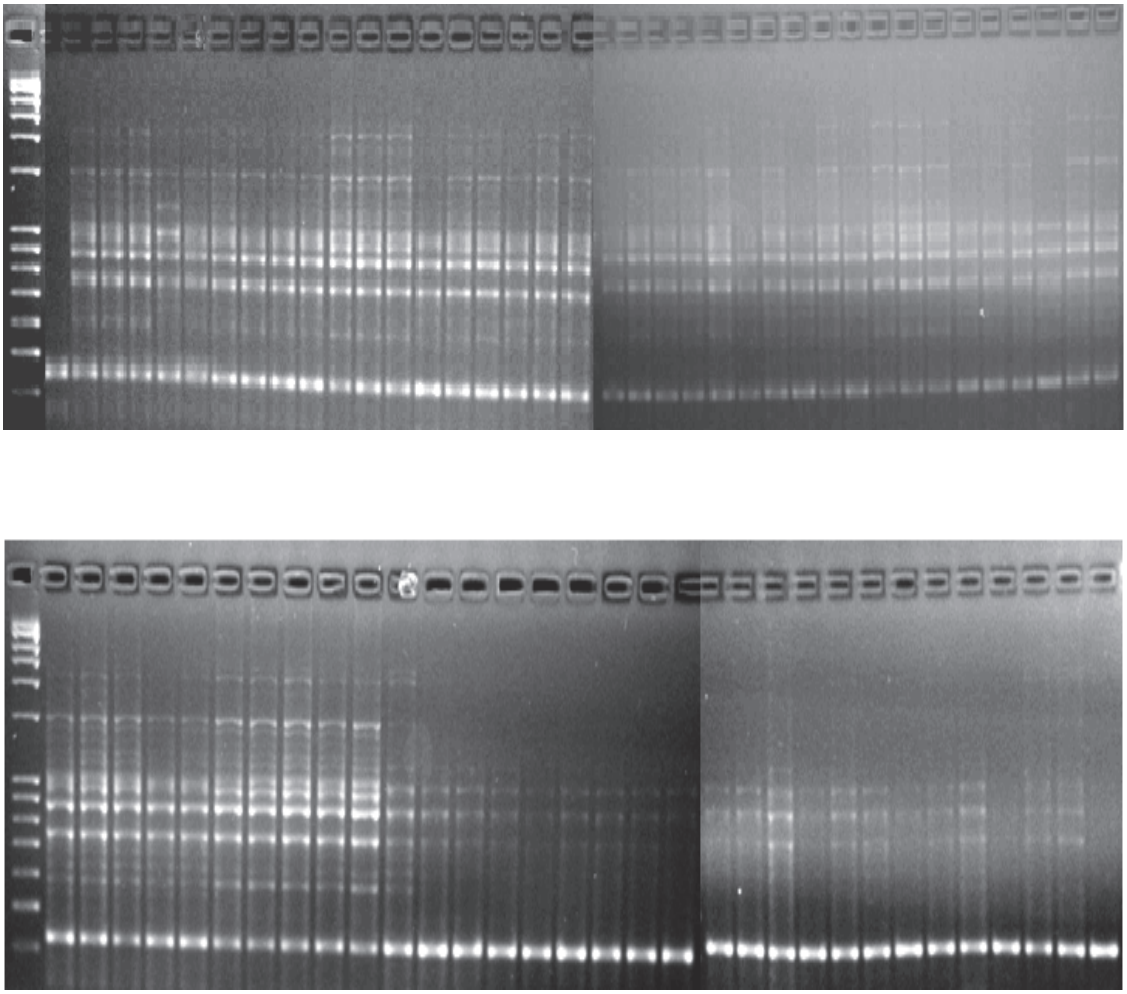


Figure 4: ISSR fingerprinting of primer 873. *Key: The first lane is DNA ladder,lanes 1-33 samples from Amhara, 34-61 samples from Oromia, 62-65 samples from SNNRP and 66-70 samples from Tigray. The accessions sequences correspond to that given in table 1*

Appendix-3 A: Photo picture of the Ethiopia cultivated safflower during growing in the greenhouse. (Photo by Baye Wodajo, 2012)

B: Pictures during laboratory sessions