Morphological and Molecular Genetic Diversity and Cytogenetics of Cultivated Anchote (*Coccinia abyssinica* (Lam.) Cogn) from Ethiopia

A Dissertation Submitted to the Department of Microbial Cellular and Molecular Biology, School of Graduate Studies, Addis Ababa University, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology (Applied Genetics)

By

Bekele Serbessa Tolera

Addis Ababa May, 2017

Morphological and Molecular Genetic Diversity and Cytogenetics of Cultivated Anchote (*Coccinia abyssinica* (Lam.) Cogn) from Ethiopia

A Dissertation Submitted to the Department of Microbial Cellular and Molecular Biology, School of Graduate Studies, Addis Ababa University, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology (Applied Genetics)

By

Bekele Serbessa Tolera

Supervisors: Kassahun Tesfaye (PhD), Addis Ababa University (AAU), Ethiopia

Kifle Dagne (PhD), AAU, Ethiopia

Mulatu Geleta (PhD), Department of Plant Breeding, Swedish University of Agricultural Sciences, Box 101SE-23053 Alnarp, Sweden

ACKNOWLEDGMENTS

First of all, my sincere gratitude goes to Wollege University for granting me study leave with full salary and partial financial support for the research. Thanks to Addis Ababa University (Microbial, Cellular and Molecular Biology Department) for admitting me to PhD study program and for monetary support; Holeta and Ambo Agricultural research centers deserve acknowledgement for allowing me to use field experimental site. Swedish University of Agricultural Sciences (SLU-Alnarp) is thanked for providing molecular lab facility.

Dr. Mulatu Geleta and Dr. Kassahun Tesfaye took the initiative of developing a thematic research project on diversity study of anchote, together with other underutilized root crops, in finding funds, involving in many technical expertises. I received not only their valuable scientific comments and advices but also I shared their brotherly approach they do have for their students and for their colleagues. Dr. Kifle Dagne has taught me many genetics courses starting from my undergraduate study through this level. He was my MSc thesis advisor and now my PhD dissertation co-advisor). I am indebted to him for his continuous advices and encouragement.

Thanks to Dr. Gemechu Kenani and Musa Jarso (PhD prospective graduate) for training me in Biometrics and SAS software application. Dr. Zelalem G/Mariam supported me with many reading materials, and valuable scientific discussions we made, and helped me in data collection.

I thank Abel Teshome (PhD) for his peak identification and fragment size determination of my capillary gel electrophoresis results.

Grateful thanks to friends and my families including my Children: Milki Bekele and Jatani Bekele, the latter born a day after my Ph.D defense and my wife Shewareg Abdisa for their unlimited encouragement and for I shared their time to be with me.

TABLE OF CONTENTS

List of contents **page**

TABLE OF CONTENTS…CONTINUED

TABLE OF CONTENTS…CONTINUED

TABLE OF CONTENTS…CONTINUED

LIST OF TABLES

LIST OF TABLES…CONTINUED

LIST OF FIGURES

LIST OF FIGURES…CONTINUED

LIST OF APPENDIX TABLES

ABBREVIATIONS

A_1	intra-chromosomal asymmetry
A_2	inter-chromosomal asymmetry
ANC	Anchote cultivated accession (code)
APPRC	Ambo Plant Protection Agricultural Research Center
AR	chromosome arm ratio
AVRWT	average root weight
CI	centromeric index
DE	days to 50% emergence
DMC	dry matter content
DMFO	days to male flower opening
HEX	hexachlorofluorescein
IA	infinite allele
IPGRI	International Plant Genetic Resources Institute
LA	chromosome long arm
LD	leaf diameter
LGC	leaf green color intensity
LL	leaf length
MCMC	Markov Chain Mont Carol
MM	Mixer Mill
ND	Nano-Drop
NSFR	number of seed per fruit
PEDL	peduncle length
PEL	petiole length
PIC	Polymorphic information content
PL	primary lobe
RC	Root color
RD	root diameter
RL	root length
SA	chromosome short arm
SLU	Sveriges Lantbruksuniversitet (Swedish University of Agricultural Sciences)
SM	stepwise mutation model
STC	stem color
TCL	total diploid chromosome length
TF%	total form percentage
TL	total chromosome length
TP	two-phase mutational model
VG	visually observed data for a group (plot)
VP	vine pubescence
DPX	Dibutylphthalate Polystyrene Xylene
DRL	difference of range relative length
FRAWT	fruit average weight
FRD	fruit diameter
FRL	fruit length
GH	growth habit
GRC	ground coverage
HARC	Holeta Agricultural Research Center
YLD	yield
6-FAM	carboxyfluorescein

ABSTRACT

Anchote (*Coccinia abyssinica*), is a perennial climbing, monoecious root crop, endemic and indigenous to Ethiopia. It is distributed over a wide range of agro-ecologies adapted to various altitudinal ranges. The root of anchote stays for several years in the soil without being damaged and its shoot sprouts during rainy season and gives fruit, then die out as the rain season ends. The root accumulates more food nutrients every growing season, and enters into dormancy until the next rainy season resumes.

C. abyssinica is not only a valuable tuber food crop but also used as folklore medicines. Despite these importances, there are very limited works available in literature regarding its genetics, agronomy, phylogeography, and evolutionary studies.

In this study, agro-morphological and molecular marker based genetic diversity and cytogenetic characterization of anchote has been done. Data on 28 agro-morphological traits were collected under five different experimental conditions, for a total of 182 accessions collected from southwestern part of Ethiopia. The results showed variations among and within accessions. On the basis of some traits, the plants were grouped into few to several phenotypic classes. Deep green (Stcdg) and purple (Stcp) stem colors are observed as rare traits.

In molecular diversity study, a total of 47 EST-SSR markers were designed on watermelon's [*Citrullus lanatus* (Thunb.)] DNA sequences and only 13 of them amplified the target regions of which eight were polymorphic and the latter were used for diversity and population structuring analyses. Twenty four alleles were observed across eight loci, where the number of alleles per locus ranged from 2 to 6, with an average of 3. Effective number of alleles ranged from 1.06 to 4.8 with an average of 1.93. Overall allelic frequency per locus (0.007-0.967) revealed larger variations. Polymorphic information contents extracted per locus were as low as 0.0619 and the largest was 0.76. Larger Shannon indices (average = 0.633) observed were the indication of better genetic diversity existing among anchote accessions. The top allelic rich populations (Arjo-Leka Dulecha, Gimbi-Nejo, Abay Chomen-Bako Tibe and Dale Sadi-Dale wobera) are leading accessions in genetic diversity parameter (He), but it is only Arjo-Leka Dulecha population that possesses 100% polymorphic loci. Although, no linkage disequilibrium was evident in this study, three loci (WM-24, WM-34 and WM-29) showed significant deviation from Hardy-Weinberg equilibrium.

The genetic diversity (He = 0.35, 0.06-0.79) estimated shows that there is a possibility of improving anchote germplasm, especially by focusing on some accessions collected from Abay Chomen-Bako Tibe, Dale Sadi-Dale Wobera, and Gimbi-Nejo, where both allelic richness and observed genetic diversity were high. Anchote crop may be considered as panmictic population in which higher gene flow is common as it is observed in this study.

Analysis of molecular variance (AMOVA) revealed the highest proportion of genetic diversity within individuals (83.75%) and the least (7.87) among populations indicating high gene flows among populations. Wider overall loci differentiation (F_{ST} = 0.01 to 0.3, with an average of 0.11) was observed. A range of 0.01-0.127 pair wise population differentiation was observed indicating that some populations are very closely related, while others are somehow distant in kinship.

Cluster analyses based on EST-SSR data of different levels of anchote groups, i.e., populations, accessions and individual samples did not show clear geographical patterns of origin, except for very few. In general, however, three apparent clusters were obtained. From neighbor joining tree (phylogram) Dale Wobera, Gimbi, Metu, Yayo, Sayo, Gechi, Dhidhesa, and Shebe Sombo represent the older lineages (groups), while Abay Chomen and Leka Dulecha look recent population.

Although, population structuring analysis of anchote accessions gave about maximum of seven sub-groups (K = 7), there is no clear variation among the groups based on the allelic proportions of the sub-groups. Some populations were found to have experienced bottleneck. These include Sibu Sire-Wayu Tuka, Guto Gida-Diga, Gera-Shebe Sombo, and Gumay-Goma.

In cytogenetic investigation, 2n = 20 chromosomes have been evidenced. All the chromosomes are metacentric in shape. No anchote cytotypes were recognized in this study.

The result of this study has many implications for breeding and conservation strategies, specially the results from cytogenetic characterization and EST-SSR based analyses. The morphological descriptors we used may also contribute for further anchote description, variety development and improvement.

Key Word: Anchote accession, *Coccinia abyssinica*, morphological trait, cytogenetics, EST-SSR, genetic diversity, population structure

Declaration

I declare that the dissertation hereby submitted by me for the Degree of Doctor of Philosophy (PhD) in Biology (Applied Genetics) to the School of Graduate Studies of Addis Ababa University is my own independent work and has not previously been submitted by me or anybody else at another university. The materials obtained from other sources have been duly acknowledged in the dissertation.

PhD Candidate: - _____

Bekele Serbessa

1. INTRODUCTION

Anchote (*Coccinia abyssinica*) is one of the tuberous root crop species and belongs to Cucurbitaceae family (Cucurbitoideae sub-family) (Edwards *et al.,* 1995; Abera Hora *et al.,* 1995). However, the species status in the genus *Coccinia* [Wight & Am. (1834)] is not well-established (Abera Hora *et al.,* 1995); hence, some of the species in Flora of Ethiopia and Eritrea are recorded with incomplete names (Edwards *et al.,* 1995). Amare Getahun (1973) reported that anchote is endemic as well as indigenous to Ethiopia and he tried to map the biogeography of the crop covering a wide range of agro-ecologies. Nevertheless, the present distributional map of the crop needs to be revised, because it was found to occur only in southwestern parts of the county (personal observation).

Anchote is adapted to various agro-ecological zones, with an altitudinal range from 550 m a.s.l in Gambela region to 2800 m a.s.l. in Kelam Wollega zone with the annual rainfall range from 950 to 2000 mm (Westphal, 1974; Amare Getahun, 1973; Tesfaye and Abebe, 1988).

Phenologically, anchote undergoes an annual cycle of death in dry season and re-generation of its herbaceous shoot at the on setting of rainy season, from “immortal” rootstock, if maintained in the soil or stored under optimal moisture condition.

Flowers of separate sexes on the same plant have been observed so far (Edwards *et al*., 1995) with male flowers blooming earlier than the females, the condition that invites outcrossing (personal observation). The plant undergoes both sexual and asexual reproduction by seed and vegetative parts (root), respectively. The sexual reproduction involves self or cross-pollination, whereas the asexual reproduction requires planting the whole root or by splicing the root part

having rootlets. However, the main means of growing anchote is by seed as a single fruit contains hundreds of seeds (Abera Hora *et al*., 1995; Tilahun Wondimu *et al.,* 2014). Anchote is, usually, cultivated during rainy season and takes four to five months for full maturation (Daba Mengash *et al*., 2012).

Both cultivation practice and animal pests affect the production of anchote. That is why, farmers often use their home gardens for anchote cultivation not only for its soil fertility, but also to protect the crop from wild animal pests such as porcupines, pigs, warthog, and others (Abera Hora *et al.,* 1995; Dandena Gelmesa, 2010).

In western Oromia, anchote has special place and is served only during special ceremonies like Thanksgiving Day, weddings, betrothal, circumcision, birthdays, and "Meskel holiday". The common anchote dish is prepared from its root as boiled or "lanqaxaa"- a finely ground and cooked root (stew), and "mura" or chopped anchote after boiling (Abera Hora *et al.,* 1995).

As compared to other root and tuber crops, anchote is the richest in protein, iron, calcium and phosphate contents with minimum anti-nutritional content, with the latter may be further reduced during cooking processes, the character that made it a nutritionally recommended food (Habtamu Fufa and Kelbessa Urga, 1997). Anchote has been used as folklore medicine to heal bone fracture, backache, displaced joints and other diseases such as gonorrhea, tuberculosis, and cancer (Dawit and Estifanos, 1991; Dandena Gelmesa, 2010).

Genetic characterization of a crop plant gives a baseline information for its further improvement, breeding programs, cultivar release, and conservation management (Haussman *et al.,* 2004; Akhtar *et al*., 2007). There are various approaches to characterize a crop plant including

comparative anatomy, morphology, embryology, and physiology either separately or complemented by molecular markers (Weising *et al.,* 2005). For morphological characterization of accessions, descriptors have been available from the International Plant Genetic Resources Institute (IPGRI) since 1977 (Biodiversity International, 2007). Even though most agricultural studies have been focusing on morphological description, nowadays there is overwhelming shift towards molecular markers, which require small amount of sample and can be done within few days. Therefore, molecular markers are becoming very important tools for the detection and description of populations diversity and differentiation analyses (Kalinowski, 2005).

A few authors (Desta Fekadu, 2011; Tilahun Wondimu *et al.,* 2014; Abraham Bekele *et al.* 2014), made appreciable efforts to describe anchote's genetic diversity based on its nutritional contents, morphological descriptors, and ISSR markers approaches, respectively, on limited samples. Their findings showed that there are significant genetic variations between and within accessions but they could not find well-defined population structuring (differentiation), i.e., geographic distance based patterns of variation. Furthermore, the researchers used a limited sample size collected from a few geographical locations.

A study on most *Coccinia* species based on molecular phylogenetics using plastid and nuclear DNA sequence data, grouped them into four clades, where *C. abyssinica* is found to be clustered under *rehmannii* clade, which includes *C. megarrhiza, C. microphylla, C. trilobata, and C. rehmannii.* According to this report, the *rehmannii* clade started diversification during an arid period at the end of the Miocene (5.2 Ma ago) period and *C. megarrhiza* was inferred as the closest species to *C. abyssinica* and they were separated around 3.9 million years ago at the end of the warm and humid early Pliocene (Holstein and Renner, 2011). However, the time of

domestication of *C. abyssinica*, more probably by Oromo (Cushitic) people (Bula Sirika, 2016), was not well known.

The diploid chromosome number reported, so far, for six *Coccinia* species revealed that it varies between 2n = 20 to 2n = 24, but exceptionally *C. grandis* contained heteromorphic sex chromosomes (2n = 22+XX/XY). However, the chromosome number and the existence of sex chromosomes for other species including *C. abyssinica* have not been reported. For *rehmannii* clade, a reduced chromosome number (2n = 20) was implied from phylogeny not from practical observation (since only two species, *C. trilobata and C. rehmannii,* were examined to have 2n = 20) and the mechanisms of reduction was not explained (Holstein, 2012).

In general, studies concerning anchote's genetic diversity, cytogenetics, reproductive biology, adaptability, resistance to diseases and drought (except for the root being perennial, or dormant for years), soil fertility restoration potential, and other agronomic as well as breeding aspects are very limited in literature. Therefore, the aim of this study is to investigate the cytogenetic and agro-morphological and molecular variations occurring in anchote accessions collected from the major growing areas of Oromia, Ethiopia.

2. Research Hypotheses and Objectives of the Study

2.1. Research Hypotheses

- There is no significant genetic variation, based on both morphological and molecular markers, within and among anchote accessions.
- There is lower genetic variability within population confined in the same geographic area than among populations of different areas.
- Inter-accession genetic distance of anchote follows their geographical distribution/location.

2.2. General Objectives

The main objective of this study is to assess the pattern and magnitude of morphological and molecular genetic diversity of anchote accessions as well as to determine its karyotype.

2.1.1. Specific Objectives

The specific objectives are:

- ❖ to assess morphological diversity within and among anchote accessions;
- ❖ to identify anchote accession(s) with superior agronomic performance across locations and recommend them for utilization in future improvement of the crop;
- ❖ to estimate the magnitude of genetic variability, heritability and expected genetic gains for important agronomic characters;
- ❖ to determine the chromosome number, the ploidy level and the karyotype of anchote;
- ❖ to develop appropriate EST-SSR markers for anchote genome and to analyze the molecular diversity of anchote accessions using the developed markers

3. LITERATURE REVIEW

3.1. Taxonomy of Anchote (*Coccinia abyssinica*)

Cucurbitaceae (commonly called cucurbits) is one of the most diversified plant families consisting of 120–130 genera (Purseglove, 1966; Edwards *et al.*, 1995) and 940 to 980 species (Schaefer & Renner, 2011). Cucurbitaceae is probably named after the chemical cucurbitacin, which was first identified from this group of plants. Cucurbitacin is oxygenated tetracyclic triterpenoids known with bitter taste and toxic to most organisms but, at the same time, can attract some specialized herbivorous insects (Da Costa & Jones, 1971; Gibbs, 1974; Balkerna-Boomstra *et al.*, 2003).

Globally, the family of Cuccurbitaceae has two large subfamilies: Zanonioideae and Cucurbitoideae, which consists of more than 20 and 100 genera (or 75 spp. and 750 spp.), respectively. Ethiopia and Eritrea, as one flora region, are represented by more than 24 genera of Cucurbitaceae (2 from Zanonioideae and 22 from Cucurbitoideae) and 71 species (2 from Zanonioideae and 69 from Cucurbitoideae) (Edwards *et al.*, 1995).

The genus *Coccinia* belongs to subfamily Cucurbitoideae and possesses about 30 extant species in the palaeotropics, all restricted, in their distribution, to Africa except *Coccinia grandis* [(L.) Voigt.], which is disseminated to other continents such as America, Asia, and Australia (Berndt, 2007; Schaefer and Renner, 2011).

Although the species status in the genus *Coccinia* [Wight & Am. (1834)] needs further investigations (Edwards *et al.*, 1995), a total of 10 *Coccinia* species have been recorded in Ethiopia and Eritrea: eight in the Flora of Ethiopia and Eritrea (Edwards *et al.*, 1995), and two

more species, *C. ogadensis* Thulin described by Thulin, (2009), and *C. microphylla* Gilg by Holstein and Renner, (2011).

Among the eight species found in the Flora, only five were fully named according to the rules for giving scientific names to plants in the Flora of Ethiopia and Eritrea. Those correctly named include *C. schliebenii* Harms (1932), *C. adoensis* (Hochst. Ex. A. Rich.) Cogn., *C. abyssinica* (Lam.) Cogn., *C. megarrhiza* C. Jeffrey, and *C. grandis* (L) Voigt (syn. *C. indica* Wight & Arn.); and those not fully named are: *C. sp* = Bally 12989, *C. sp.* = Burger 2947A, and *C. sp*. = Gilbert & Jones 129 (Edwards *et al.*, 1995).

3.2. Origin, Domestication, and Current Distribution

Understanding a crop plant's geographic distribution and/or center of origin is very important for breeding, genetic improvement, and conservation managements activities of the crop. This is because of the nearby availability of the wild type and related species, which can provide adaptive value as well as broaden the genetic base of a crop species via outcrossing. In fact, the center of origin is, usually considered as center of diversity (Acquaah, 2007; Sebastian, 2011). Therefore, determination of center of origin of a crop plant is important to conserve its genetic diversity, especially for those species, which are vulnerable to ecosystem fragmentation (degradation) and other anthropogenic pressures.

Cucurbitaceae species are distributed in the tropics and subtropics of both the Old and New Worlds, with hotspots of diversity in Southeast Asia, West Africa, Madagascar and Mexico (Schaefer & Renner, 2011). However, Schaefer *et al*. (2009) observed the great disjunction between related genera in their geographical distribution. They reasoned out that there have been

many successful long distance dispersals (by different mechanisms, usually by birds) between Asia and Africa, back to Asia, between Africa and South America, and from Asia to Australia (Schaefer *et al*., 2009; Schaefer & Renner, 2011).

Holstein and Renner (2011) used molecular technology to estimate the center of origin and diversification dates of wild *Coccinia* species to be, in eastern and southern Africa, around 6–7 million years ago. The genus has a broad range of agro–ecology from semi–arid habitat to moist forest (Holstein and Renner, 2011).

Since Ethiopia is endowed with such highly diversified agro–ecologies, the genus *Coccinia* is suitably occurring in wide range of areas including the western, central, southeastern and northern parts of the country (Edwards *et al*., 1995). However, the well-known species for its food, medicine, and other socio-cultural values, in western and southwest Ethiopia, is only anchote [*Coccinia abyssinica* (Lam.) Cogn. (1881)] (Amare Getahun, 1973).

In different parts of Ethiopia, the plant is known by its vernacular names such as anchote (Afaan Oromo) (Amare Getahun, 1973), ushushe (Welayita), wushish (Tigrinya), shushe/ushushe (Dawuro), ajjo (Kefinya) (Wolde Michel, 1987). This indicates that, anchote was originated and endemic to Ethiopia, where it is found as cultivated and in wild form (Amare Getahun, 1973; Edwards *et al*., 1995).

The crop is cultivated in backyard, for its rootstock, particularly in southwestern part of Ethiopia, namely Wollega, IlluAbaBora, and Jimma Zones. Other parts, such as Shoa and Harerghe Zones of the Oromia Regional State have also started cultivation of anchote recently (Tesfaye and Abebe, 1988; Karin, 2002). Recent work by Bula Sirika, (2016) shows that anchote crop is also

being cultivated in other countries by Oromo Diasporas, including in various states of Canada and United States of America. However, an early research done on anchote's habitat, by Westphal (1974), indicated that commonly two agro–ecological zones have special connection with it. These regions are: 1) the southeastern part of the Ethiopian highlands, the area situated at altitudes of 1800 m a.s.l. and higher which have Alfisols as a major soil type. This area receives 950–1500 mm average annual rainfall; and 2) Southwestern part of the Ethiopian highlands, which covers areas such as Wollega, IlluAbaBora and Jimma. The latter region is known with, Oxisols, Ultisols, and Vertisols as major soil types. The area is situated between 1500–2400 m a.s.l. and receives an annual rainfall of about 1500 mm to over 2000 mm per year.

However, the crop is also found growing even in cooler and higher altitudes such as Inango Dambal (2820 m a.s.l) in West Wollega (Amare Getahun, 1973; Tesfaye and Abebe, 1988) and Walal Mountain in Kellem Wollega (Abera Hora *et al.,* 1995). Presently, the collections of *C. abyssinica* specimens in the Natural Herbarium of Ethiopia at Addis Ababa University (AAU), representing different agro–ecological zones of the Flora region, indicates that anchote is growing, even, in more wide altitudinal range, i.e., from lowest 550 m a.s.l (in Gambela region) to 2500 m a.s.l (in Debre Libanos monastery).

3.3. Importance of Cultivated Anchote

The main aim of cultivating anchote is for direct consumption of its root, but sometimes people use its leaves as a vegetable (Abera Hora *et al.,* 1995). It is among the major root and tuber crops in western and southern parts of Ethiopia (Daba Mengash *et al*., 2012). Although information on nutritional and anti–nutritional contents of anchote's leaf and seed is limited, the root has been better studied and reported as possessing a higher nutritional content than other common and

widespread root and tuber crops (Habtamu Fufa and Kelbessa Urga, 1997). Even though there is no report on nutritional composition of anchote seed, earlier study showed that most seed proteins of Cucurbitaceae have a comparable nutritive value to those of Leguminoceae (Thompson *et al.,* 1978).

The anti–nutritional contents of anchote tuber (like tannis, trypsin inhibitors, phytic acid and oxalic acid) have been found insignificant and this reflects the desirable quality of the plant's root (Habtamu Fufa and Kelbessa Urga, 1997; Abera Hora *et al*., 1995). Sometimes the foliage of the crop can also be used as feed for domestic animals (personal communication).

In many countries, different *Coccinia* species have been reported as a folklore medicine to treat various diseases such as skin diseases, diabetes, and gonorrhea (Kuriyan *et al*., 2008; Blanca *et al.,* 2011, Munasinghe *et al*., 2011; Sivaraj *et al*., 2011). As an example, the crude leaf extract of *Coccinia grandis* exhibited a broad range of anti-bacterial activities (Sivaraj *et al.,* 2011).

Similarly, in Ethiopia *C. abyssinica* has been recommended to treat individuals suffering from bone fracturing, displaced joints and other diseases such as gonorrhea, tuberculosis, and cancer (Dawit and Estifanos, 1991; Dandena Gelmesa, 2010). This may be because of its high calcium and protein contents that repairs damaged bones; however, peeling of anchote during cooking reduces its nutritional contents (Habtamu Fufa and Kelbessa Urga, 1997).

Anchote holds a very special place in the traditions and customs of the western Oromo of Ethiopia. It is served at different ritual ceremonies in different forms. For example, Reinhard and Admasu (1994) reported that in western Oromia, root of anchote is boiled and prepared with butter for the "Meskel" holiday (actually "Meskel" celebration in September is the time when

anchote is, usually, matures and harvested). It is also common that 'lanqaxaa',- anchote dish (stew) prepared from finely ground tuber, is commonly served during ceremonies marking weddings, betrothal, circumcision, birthdays, and religious celebrations like New Year and Thanksgiving day for good harvest as well as on other occasions (Abera Hora *et al*., 1995). That is why the menu of “anchote food” at various restaurants and hotels, mainly in Wollega towns, is by far more expensive than other menus (personal observation).

3.4. Reproduction and Propagation Methods

Anchote is a tuberous perennial with trailing annual shoot, of several meters long, that climbs up a support by means of its simple tendrils (Edwards *et al*., 1995). It undergoes an annual cycle of death and regeneration of herbaceous shoots, the condition known as hemicryptophytic life form (Schaefer and Renner, 2011).

The male and female flowers occur on separate nodes of the same plant. The two flowers bloom at different time (a dichogamy situation called protandry) (Edwards *et al*., 1995). This nature, obviously, invites outcrosses and prevents inbreeding. As protandrous species tend to be pollinated by bees or flies (Sargent and Otto, 2004), anchote is mainly pollinated by bees (Edwards *et al*., 1995).

The reproduction mechanism involve both sexual and asexual, since anchote can grow from seed and vegetative parts (root). The latter method involves either planting a root as a whole or by splicing it into pieces, as far as each piece has rootlets and an external covering. Asexual propagation method is usually done to establish “mother” plant, called “Gubo”, which serves as a seed source for further plantings. Stem cutting may be, sometimes employed for anchote

propagation. However, for commercial production, seed is preferred over other methods (Abera Hora *et al*., 1995).

Seed for the next year sowing are harvested from plants of good quality including size, color or other traits and stored, usually after mixing with ash, in pots or any other containers made of clay materials. Since anchote root has a dormancy period of many years, farmers keep it in the soil as a method of preserving selected anchote. For consumption, the roots are harvested by digging them up from the soil (Abera Hora *et al*., 1995).

Anchote is considered as “food of the poor” since there is no need of larger plot of land and usually growers plant it under their fence. Commonly, women are responsible in harvesting, preparing, processing and storing anchote tubers and seeds. They make seeds available for sowing well ahead of the beginning of the rainy season, while men are involved in land preparation and sowing the seed. When seed is lacking or better quality anchote is needed, women save one or few of the “tuber root(s)” and plant them so that they grow and fruit well to provide enough seeds. At the beginning of the next rainy season, the root (gubo) produces new shoots, usually more than one (Abera Hora *et al*., 1995). These newly emerging shoots depend on the reserve food in the root making the root less suitable for consumption, especially during the early growth period. After the stems are well grown, however, the shoots start manufacturing their own food, and the root again become suitable for consumption. A root produces new shoots and developed into a complete plant. A single anchote plant provides many fruits, from which hundreds of seeds are extracted. In the next growing season, the seeds will be sown and enough anchote can be produced for home consumption (Abera Hora *et al*., 1995).

3.5. Production and Productivity

Anchote is usually cultivated in rainy season (April to January) which depending on agro-ecological zones takes four to five months to mature to consumable or marketable root size, which makes anchote a short season crop.

The yield of anchote varies in different soil types and agro-ecological zones. For instance, one annual report of 2004/5 production season shows that anchote crop was produced nearly on 3,000 hectare of land in West Wollega Zone which yielded a total of 25,000 tons of tuber (Anonymous, 2011). However, maximum tuber yield (76.45 t/ha) was obtained in Ebantu agro-ecological condition (Daba Mengash *et al*., 2012) under experimental condition and even more (80 t/ha) yield was obtained on vertisol soil type at Debre-Ziet Agricultural Research Center (Desta Fekadu, 2010).

Agro-ecological conditions and other factors (like pests) can affect anchote production and productivity. The pests include both domestic and wild animals. Porcupine, warthog, and wild pig are among a few wild animal pests of the crop. The former two pests eat anchote root by digging into the soil whereas the others consume the foliage and damage the crop by trampling on it. For commercial and higher quality anchote production, the planted area should be fenced or properly protected from damage by animal pests. This is why farmers usually grow the crop around their homes. Fruit boring insects can also damage the fruit and hence, reduces the production (Abera Hora *et al.*, 1995).

3.6. Genetic Resources and Its Diversity

Genetic diversity study involves analysis of variations at individual, group, or at population levels (Mohammedi and Prasanna, 2003). The genetic diversity issue of a crop is directly related to its pollination biology. For cross-pollinating plants, genetic diversity may be maintained easily unlike that of strictly self-pollinating plants. In the latter case, breeders use different methods for crop improvement such as artificial selection, hybridization, mutation breeding, and other molecular techniques (Ngampongsai *et al.,* 2009).

Diversity analysis of germplasm collection can facilitate reliable classification of accessions and identification of subsets of important accessions with possible utility of specific breeding purpose, i.e., basic step for further improvement, breeding programs, cultivar release, and conservation management (Haussman *et al.,* 2004; Akhtar *et al.,* 2007). These genetic characterization processes (involving both marker and analytical method choices) have different approaches with objective of the study, level of resolution required, resource and time constraints (Mohammedi and Prasanna, 2003). One may wish to characterize a crop plant by comparative anatomy, morphology, embryology, and physiology separately or complemented by molecular markers, which actually vary based on the taxonomic level of the study materials (Weising *et al.,* 2005). For accessions' morphological characterization, the International Plant Genetic Resources Institute (IPGRI) descriptors have been available since its first release of descriptor list in 1977 (Biodiversity International, 2007). Even though most agricultural studies have been focused on morphological description, nowadays there is a tendency to use more stable and reliable molecular markers (Mohammedi and Prasanna, 2003).

For accurate and unbiased estimate of genetic diversity, by any of these makers, there are some basic considerations during sampling, experimentation, and bio-geographical distribution record, as well as choice of data analysis tools (Meirmans, 2015). Perennial root crops, similar to anchote, that can be stored for many years in the soil and sprouts on the setting of rainy season, constitute a reference genetic material on which as many molecular markers can be accumulated as desired, without limitation on the quantity of DNA (de Vienne *et al.*, 2003).

3.6.1. Morphological Traits for Diversity Investigation in Root and Tuber Crops

Taxonomists, farmers, and breeders use the most commonly and easily observable markers called morphological descriptors for classification and evaluation of yield or other trait of interest (Singh and Parab, 2015).

As anchote is a new crop to the rest of the world and to science as well, it is difficult to give clear diversity review. However, few studies show that anchote exhibits a great range of diversity in foliage, fruit, and root morphologies (Dandena Gelmesa, 2010; Tilahun Wondimu *et al*., 2014), which could add values as ornamental plant beside its importance as food. It is also occurring in its “wild” form that provides gene flow and ensures broad genetic diversity (Edwards *et al*., 1995) with little differentiation.

Anchote can be characterized by its unique physical features. Tuber shape is highly affected by the age, soil physical conditions, anchote type (genetic) and cultivation managements. The common shapes of anchote roots are spherical and conical, whereas other shapes are due to soil structure that affects its normal growth and acts as prop its tip or sides. The outer part (skin) of anchote root is grayish, without exception, while the edible internal part has mainly two colors, white and “reddish” (Abera Hora *et al.,* 1995).

Anchote has ovate or broadly ovate, scabrid punctuate above, crispate-setulose on veins beneath, 7.5–15 x [7.5–17] cm, margin sinuate-denticulate, shallowly to deeply palmately 5–lobed, lobes triangular to ovate, sometimes laciniate leaves. Petiole length ranges between 1.5–16 cm, spreading-setulose. Coiling tendrils are simple, usually attach with support. “Monoecious” sex (separate male and female flowers); raceme-like clustered male flowers of variable number; solitary female flowers; fruits ellipsoid, dark green with longitudinal lines of white spots, red with lines of paler spots when ripe, with variable dimensions; seeds broadly asymmetrically ovate in outline, compressed, covered in fibers (Edwards *et al.*, 1995); stem and leaf stalks are solid (Abera Hora *et al.,* 1995).

The growers select the good “variety” based on some phenotypic and nutritional based characters such as root color, root yield, and ease of cooking as well as palatable taste (source: growers). Such preference on limited traits, by growers/farmers, of the crop may result in ending up with narrow genetic diversity, and finally leads to extinction. Therefore, such artificial selection can affect the genetic diversity of anchote, which is also common for other crop (Lasalita-Zapico *et al.,* 2010). The other threat for such orphan crop is an introduction and advertisement of other, alternative, high yielder crops like Irish potato, cassava, and others (Abera Hora *et al.,* 1995).

3.6.2. Molecular Based Diversity and Population Structure Analyses

Morphological markers are characterized by their insufficiently polymorphic, dominant, and are affected by environmental conditions at different developmental stages (de Vienne *et al.*, 2003). Therefore, molecular markers have been replacing it for better characterization of genetic diversity, population structure, and evolutionary studies (Sessions, 1996).

Molecular based genetic diversity and population structure studies of a crop plant have great importance for germplasm collection, breeding programs and genetic resource conservation and management of the crop (Wen *et al.*, 2010; Lu *et al.*, 2011).

One of the breakthroughs in biological sciences is the invention of Polymerase Chain Reaction (PCR) that amplifies DNA *in vitro* (Mullis and Faloona, 1987). Since then, molecular marker techniques were separated into non-PCR based and PCR-based depending on the way in which they produce different DNA fragments.

The non-PCR markers include RFLP and its modifications and PCR-based markers include RAPD, AFLP, ISSR, SSR, *etc* (Botstein *et al*., 1980; Litt and Lutty, 1989; Williams *et al.*, 1990; Zietkiewicz *et al.*, 1994; Vos *et al.*, 1995). The PCR-based molecular markers have been used for plant genetic diversity analyses with different power of genome sampling and generating polymorphisms (Weising *et al*., 2005).

3.6.2.1. Expressed Sequence-Derived Simple Sequence Repeat (EST-SSR) Marker

Nature of SSR: Microsatellites or simple sequence repeats (SSR) are highly variable, 1-6 base pairs in length, tandemly repeated motifs of DNA distributed abundantly throughout the eukaryotic genome. Microsatellites have different repeat units such as mono-, di-, tri-, and tetra nucleotide repeats (Tautz and Renz, 1984; Temnykh *et al*., 2001). The four categories of SSRs are perfect, imperfect, interrupted, and composite. The perfect microsatellite is composed of non-interrupted repeat motifs whereas the imperfect SSRs include a base pair not being part of the repetitive motif. The interrupted SSR is to mean the repeat motifs, which include small

sequence that does not match the repeat sequence and in composite microsatellites, the motif contains two adjacent distinctive repeat sequences (Olivera *et al*., 2006).

The repeated motifs are believed to be originated from the non-repeat sequences by insertions or substitutions. Finally, the motifs will become more abundant and occurring in variant forms by several mechanisms including: errors during recombination (unequal crossing-over), and polymerase slippage during DNA replication or repairing. The crossing-over mutation effect is more pronounced than by polymerase slippage since it adds or deletes large number of repeats (de Vienne *et al.*, 2003; Olivera *et al*., 2006). Polymorphism between individuals of the same population can be caused by mismatches at priming sites and the length of repeat motifs (Zietkiewicz, 1994).

SSR marker searching follows either enrichment procedure (for whose genomic library is already constructed) (Ostrander *et al.*, 1992; Edwards *et al*., 1996) or *in silico* mining from existing databases (Temnykh *et al*., 2001), whereas the later approach being cost and time effective (de Simko, 2009). The enrichment procedure follows screening the marker by probe, sequence the positive clone, synthesize the oligonucleotide primers, and finally testing the primers with samples (Edwards *et al*., 1996). This method is very important in enriching SSR marker library, which in turn facilitate germplasm genotyping, selection, and genetic mapping. Moreover, this method is an efficient approach, especially for plants that do not have expressed sequence tag (ESTs) or genomic databases (Novelli *et al*., 2013).

SSR can occur both in genic (expressed sequence tag-EST-SSRs) and non-genic regions of a genome (G-SSRs). The advantage of EST-SSR markers over the other is many folds. Primarily, they are derived from transcripts (cDNAs) where their flanking regions are highly conserved

across taxa, hence they serve as a functional markers, inexpensive to develop, and high transferability between related species and even between genera (cross species amplification) (Peakall *et al*.,1998; Wen, 2010; Mao *et al.,* 2014). Transferability not only reduces the cost of designing specific primers but also facilitates the quick isolation of microsatellite from a taxon with low microsatellite frequencies. Transferability is limited by kinship and genome size and complexity (Olivera *et al*., 2006). Using EST-SSR markers also reduces sampling error associated with sampling of loci, since the same loci are considered for each accession in EST-SSRs amplification (Mohammedi and Prasanna, 2003). Therefore, they are potential candidates for gene tagging and comparative studies of related species (de Simko, 2009).

SSRs are markers of choice because of their high reproducibility, co-dominant inheritance, rich in information content or highly polymorphic, locus-specific (independent marker or it distinguishes between alleles), ease of automation, and relatively low cost as compared to single nucleotide polymorphism (SNP) marker. The polymorphism in length of each motif can be detected either by agarose gel or by acrylamide gel electrophoreses depending on the length between the alleles or more frequently with automatic sequencing methods (de Vienne *et al*., 2003). Because of these characteristics, EST-SSRs have been, preferentially used for a wide range of practically reliable analyses. These include genetic mapping studies, marker assisted selection (MAS), genetic diversity surveys, quantitative trait loci (QTL) analysis, germplasm maintenance programs, cultivar genome fingerprinting, marker-trait association, and population structure study (Zietkiewicz, 1994; de Simko, 2009; de Vienne *et al*., 2003; Ferra˜o *et al.,* 2014; Wang *et al*., 2014).

Although ESTs-derived SSRs are limited in number of alleles per locus (N_a) and in heterozygosity (H_e) than genomic microsatellites (G-SSR) (de Simko, 2009), both estimate genetic diversity and population structure similarly (Kim *et al.*, 2008). The other drawback of this marker is the need of *de novo* isolation from the species that are being examined for the first time (Zane *et al.*, 2002).

SSRs are important in modulating gene expression, for local adaptive value (*Adaptive Peaks Theory* of Wright's, 1932)-the frequency of SSR in a population represents the maximum local adaptive value. That means, most mutation creating variant genes of lower adaptive value. Since EST-SSRs are part of expressed region, they are important for many biological functions such as gene regulation, recombination (especially, di-nucleotides are recombination hot spot), DNA replication, cell cycle and mismatch repair. Recombination through SSR recovers genetic variations lost by genetic drift (Olivera *et al*., 2006; Ranade *et al*., 2014).

SSR mutational models: In population genetic analysis, there are four commonly used models concerning the origin of microsatellites.

The infinite allele (IA) model considers that each mutation randomly creates new allele and the proximity in terms of number of repeats does not indicate a greater phylogenetic relationship. This model is Wright's (1951) model, for which he used F-statistics for differentiation analysis.

Stepwise mutation (SM) model, in contrast, assumes that each mutation is accompanied by gain or loss of one repeat motif, i.e., in this type of mutation, two alleles differing by one repeat unit are more closely related than alleles that differ by several repeat units. Slatkin's (1995) R_{ST} (that take into account the effect of mutation), and Nei's (1973) G_{ST} are genetic differentiation

measures proposed for this model. In the investigation of population structure by these measurements, we have to be careful, because they give significantly different values (usually F_{ST} is lower than R_{ST}). The non-concordance between the two measurements is mainly due to the high and variable mutation rates of microsatellite that usually show high level of within population heterozygosity. SM model is preferred and fits for non-homoplasy markers (alleles identical not by state but by descent) like EST-SSRs, which are sampled from the same loci for all individuals.

The third model, **two-phase (TP) model**- was introduced by Di Reinzo *et al.* (1994) as an extension of SM model briefing the possibility of a rare event by which a larger number of repeats can occur rather than by the most frequent SM model. Crow and Kimura (1970) proposed another model, the **K-allele model**. This model assumes that if there are exactly k possible alleles per locus then the probability of a given allele to mutate into any one of them is μ/k-1, where μ is mutation rate (Olivera *et al*., 2006).

3.6.3. Cytogenetics as a Diversity Study Tool

Cytogenetic study of *Cucurbitaceae* was started during the early 1920s that reported the basic chromosome numbers of 11 and 12, the latter was considered as primitive basic number (Robinson and Dec- Ker-Walters, 1996). Since then, cytogeneticists and plant breeders have given better attention for the somatic metaphase chromosome and karyotype analyses for genetic diversity, mapping, breeding, evolutionary systematic, and phylogenetic analyses (Almeda, 2013; Ma *et al*., 1996; Sessions, 1996).

In order to study the diversity or identity of a group of species, the quick, easy, reliable and well-established method is to investigate the diploid chromosome count and/or ploidy level (Kazem,

$\overline{Sd}$ $\bar{X}$ $\overline{Sd}$

$\bar{X}$

Even though the genus *Coccinia* is found to be paraphyletic (Reddy, 2009), the genetic investigation in general and cytogenetics in particular is very limited, only a few somatic chromosome number counts are reported. For instance, six *Coccinia* species were reported to have diploid chromosomes (2n) ranging from 20 to 24 (Holstein and Renner, 2011). In some relative genera of *Coccinia*, one study showed chromosome number count variation between closely related species (*Cucumis melo* 2n = 24 and *Cucumis sativus* 2n = 14) (Samuel *et al*., 1995). The diploid chromosome number 2n = 24 seems conserved throughout the family since it is common in a wide range of species in sister genera too (Sutar *et al*., 2013). This conserved chromosome number (2n = 24) throughout the family indicates the basic chromosome number might be twelve (x = 12) as shown in early study (Robinson and Dec-Ker-Walters, 1996). To our knowledge, there was no cytogenetic work done on anchote crop, partly because of the endemicity of the crop and partly due to its being rare garden plant (orphan crop).

Therefore, it is essential to establish the correct diploid chromosome number of anchote, which is useful for its systematics, breeding, and genetic studies.

4. MATERIALS AND METHODS

4.1. Anchote Germplasm Collection

Anchote seed (germplasm) sample collection regions were predetermined by the availability and long history of traditional cultivation practices and usage by native people of western Oromia (Ethiopia) (Amare Getahun, 1973; Abera Hora *et al*., 1995). Therefore, the major growing areas, namely, West Wollega, Kelam Wollega, East Wollega, IluAbaBora, and Jimma Zones were covered in order to collect as many genotypes as possible for the aim of genetic diversity study (Figure 1).

For the present study, 182 accessions (ANC1-ANC182) of anchote seeds were collected from the above mentioned zones (Table 1), in year 2013. The seed samples were acquired from households in the form of gift or purchase. A total of 20 additional accessions (ANC183-ANC202) were acquired from a researcher at Bako Agricultural Research Center, making the total number of accessions 202. The accession code ANC represents the material name (ANC = anchote cultivated) and the serial number indicates the collection order (1 = first, 2 = 2^{nd}, 3 =3^{rd}…etc).

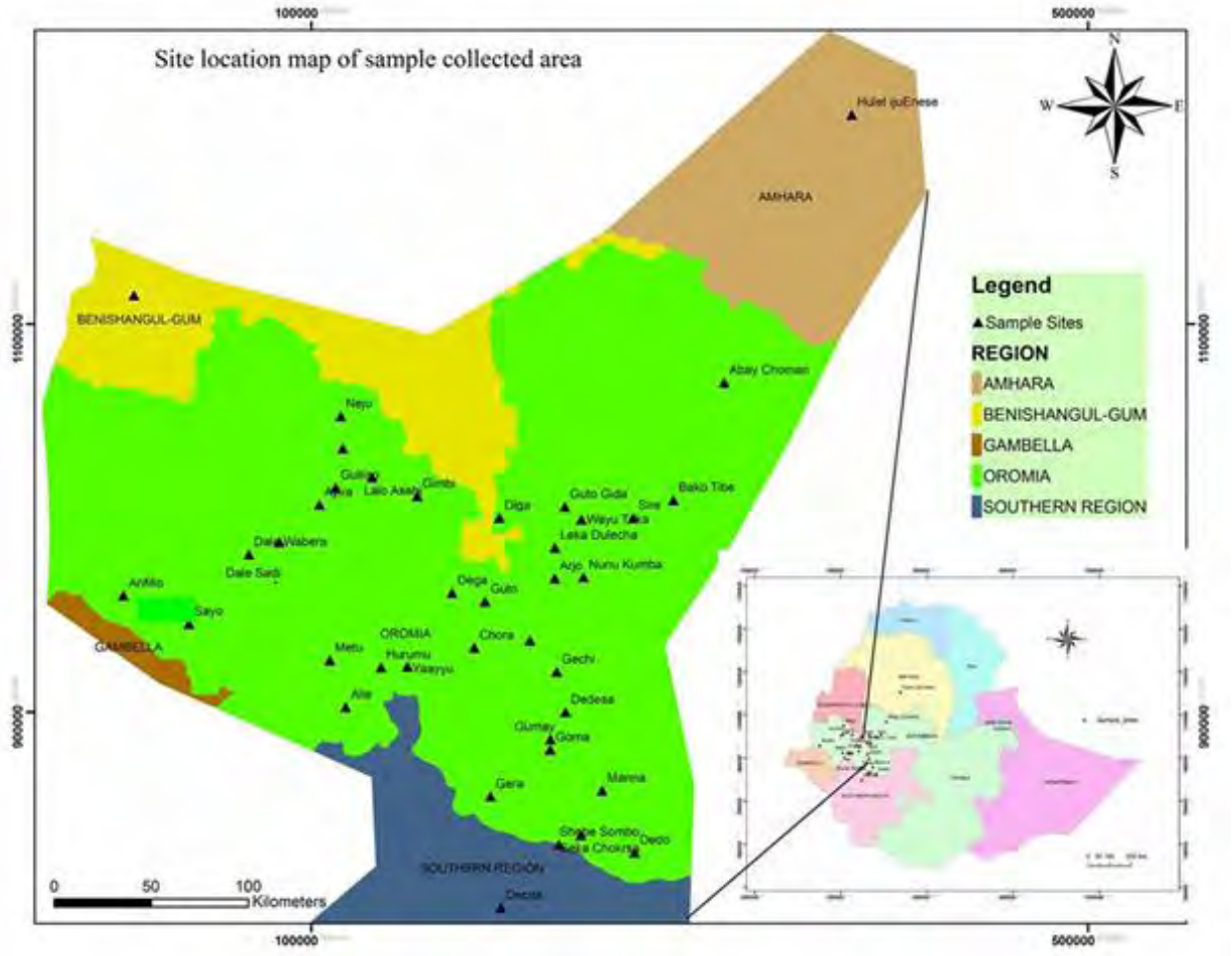

Site location map of sample collected area
100000
500000
1100000
900000
N
W
E
S
Legend
Sample Sites
REGION
AMHARA
BENISHANGUL-GUM
GAMBELLA
OROMIA
SOUTHERN REGION
Hulet ijuEnese
AMHARA
BENISHANGUL-GUM
Abay Chomen
Neyo
Gulliso
Ano
Lalo Asabi
Gimbi
Diga
Guto Gida
Bako Tibe
Sire
Wayu Tuka
Leka Dulecha
Dale Wabera
Dale Sadi
Arjo
Nunu Kumba
Dega
Guto
Ambo
Sayo
GAMBELLA
OROMIA
Chora
Metu
Hurumu
Yaaryu
Gechi
Alle
Dedesa
Gumay
Goma
Gera
Manna
Shebe Sombo
Sigja Chokorsa
Dedo
SOUTHERN REGION
Decha
0
50
100
Kilometers

Table 1. Anchote accessions passport data. Germplasm (seed) sample collection sites (Zone and Woreda), co-ordinate, altitude and sample size

Zone	Woreda	Co-ordinate	Altitude	Accession Code	Sample size	Total/zone
	Arjo (Ar)	8045'18.34"N 36029'53.66"E	2510.638	ANC 123 – ANC 142	20	
	Diga (Dg)	9001'32.40"N 36027'14.53"E	2215.896	ANC 152 – ANC 162	11	
	Guto Gida (Gg)	9001'22.30"N 36028'24.35"E	2242.414	ANC 143 – ANC 151 & ANC 188(DIGGA–1)	10	
	Leka Dulecha (Ld)	8053'51.92"N 36029'11.29"E	2161.946	ANC 163 – ANC 172	9	
	Nunu Kumba (Nk)	8045'47.32"N 36038'15.59"E	2344.217	ANC 180 & ANC 181	2	
	Sibu Sire (Ss)	9001'58.13"N 36051'59.67"E	1818.742	ANC 182, ANC 185(223093), ANC 189(KUWE)	3	
	Wayu Tuka (Wt)	9004'30.37"N 36030'47.77"E	2320	ANC 173 – ANC 179 & ANC 194(KICHI)	9	
East Wollega	Guto Wayu (Gw)	-	-	ANC 183	1	65
Horo Guduru Wollega	Abay Chomen (Ac)	9031'28.09"N 37021'04.60"E	1980	ANC 187 (90802–1) & ANC 190(90802)	2	2
	Alle (Al)	8008'57.04"N 35032'12.79"E	1950.415	ANC 044 – ANC 047, ANC 200(223108–1)	5	
	Chora (Ch)	8021'39.86"N 36007'14.56"E	2021.129	ANC 058 – ANC 066	9	
	Dega (De)	8034'06.69"N 36006'36.03"E	2266.798	ANC 067 – ANC 072 & ANC 202(223104)	7	
	Dhidhesa (Dh)	8004'53.55"N 36027'32.67"E	1902.257	ANC 082 – ANC 087	6	
	Gechi (Gc)	8020'51.11"N 35051'53.91"E	1642.872	ANC 073 – ANC 081	9	
	Hurumu (Hr)	8020'51.31"N 35041'49.24"E	1815.084	ANC 049 – ANC 053	5	
	Metu (Mt)	8018'00.01"N 35035'00.65"E	1703.222	ANC 039 – ANC 043 & ANC 048	6	
Illu-AbaBora	Yayo (Yy)	8020'09.42"N 35049'20.82"E	1598.981	ANC 054 – ANC 057	4	51
Kaficho Shakicho	Decha (Dc)	6029'48.33"N 35007'58.05"E	2150	ANC 186(240407B) & ANC 191(24407G)	2	2
West Shoa	Bako Tibe (Bt)	9007'05.17"N 37003'30.65"E	1780	ANC 184(220563-1) & ANC 197(220563–1)	2	2
	Dedo (Dd)	7030'58.80"N 36052'39.12"E	2112.569	ANC 108 – ANC 115	8	
	Gera (Gr)	7030'00.14"N 36004'00.00"E	2164.69	ANC 098 – ANC 107	6	
	Goma (Go)	7051'26.25'N 36035'18.16"E	1671.523	ANC 094 – ANC 097	8	
	Gumay (Gu)	8001'00.00"N 36031'00.00"E	1621.536	ANC 088 – ANC 093	6	
	JARC (Jm)	-	-	ANC 121-ANC122	2	
	Seka Chokorsa (Sc)	7036'05.68"N 36043'31.35"E	1843.43	ANC 116 – ANC 119	4	
Jimma	Shebe Sombo (Sh)	7030'22.25"N 36030'48.40"E	1776.07	ANC 120	1	35
	Anfilo (An)	8028'19.64"N 34035'26.55"E	1496.568	ANC 027 & ANC 028	2	
	Dale Sadi (Ds)	8053'59.12"N 35013'28.30"E	1516.685	ANC 012 – ANC 015	4	
	Dale Wobera (Dw)	8055'46.01"N 35003'09.30"E	1606.296	ANC 016 – ANC 020	5	
Kelam Wollega	Sayo (Sy)	7036'05.68"N 36043'31.35"E	1843.43	ANC 021– ANC 026, ANC 029 & ANC 030	8	19
	Gida Gebo (Gd)	9039'52.53"N 5036'13.38"E	2359	ANC 192&193, ANC 199, ANC 201	4	
	Ayira (Ay)	9005'50.16"N 35023'46.86"E	1625.803	ANC 008 – ANC 011	4	
	Gimbi (Gm)	9010'15.15"N 35050'08.88"E	1967.789	ANC 031 – ANC 034 & ANC 195(GM) – ANC 198(GM)	7	
	Guliso (Gl)	9010'14.89'N 35028'21.77"E	1574.597	ANC 006 & ANC 007	2	
	Lalo Asabi (La)	9010'02.79"N 35041'04.15"E	1822.399	ANC 001 – ANC 005	5	
West Wollega	Nejo (Nj)	9030'08.26"N 35030'10.22"E	1870.862	ANC 035 – ANC 038	4	26
Total						202

4.2. Field Experiment for Agro-Morphological Diversity Evaluation

4.2.1. Experimental Site, Layout and Experimental Design

For agro-morphological characterization of anchote accessions, two sites, namely, Ambo Plant Protection Research Center (APPRC) and Holeta Agricultural Research Center (HARC) were selected for their logistic advantages. HARC and APPRC are located at 28 Km and 114 Km West of Addis Ababa, respectively.

HARC is located at 09° 02' N latitude and 38° 34' E longitude at an altitude of 2400 m.a.s.l with an average annual temperature of 21° C and annual rainfall of 900 to 1100 mm. The soil type is a mixture of Nitosols and Vertisols (source: HARC).

APPRC is located at 8° 57' N and 38° 07' E coordinates in a temperate highland region at an altitude of 2175 m.a.s.l. The Center has an average temperature of 18.7° C with minimum of 10.13° C and maximum of 27.63° C. The annual average rainfall is 1,242.9 mm. The period from June-October accounts for more than 77% of the total annual rainfall. The soil type of the Center is mainly Vertisols (67% clay, 18% silt, 15% sand, and 1.5% organic matter) (source: APPRC).

Plants were grown for two cropping seasons at HARC and one season at APPRC. During the first cropping season, plants were grown from seeds at both sites. During the second growing season plants were grown only at HARC from seeds, because of seed shortage and so plantation at APPRC couldn’t be repeated. These seasons make three experimental environments for agro-morphological characterization (APPRC season one = environment one, HARC season one = environment two, and HARC season two = environment three).

At HARC plants were raised, both during the main cropping season as well as during off-season, from roots ("gubo") harvested, from plants grown from the previous season, for fruit and seed characterization.

All the five experiments were laid out according to an alpha lattice design, with two replications each consisting of two incomplete blocks following Abd El-Shafi and Abo-Hegazy (2013) recommendation. Each accession was grown in a single row of 20 plants in a plot of two-meter by 40 cm, with 40 cm space between plants. The space between the adjacent plots was 40 cm wide whereas the inter-block distance was one meter and two meters between replications. The total experimental plot of land, for each experimental environment, was 646.4 m^2. For irrigation based experiment, the same experimental design, as described above was followed, except that the spacing between plants was 20 cm and the number of plants per plot was reduced to ten.

Fertilizer was applied in planting holes before sowing at the rate of 46/20 kg/h (N/P) as recommended for anchote by Girma Abera and Hailu Gudeta (2007). Different agronomic practices were also carried out following seed germination, including thinning (where necessary), hand weeding, earthing-up (top dressing), erecting support (stalking) poles, and irrigating the land before harvesting the root. For root plantation, the plot of land was irrigated before planting. Irrigation continued from plantation time until harvesting time at different intervals depending on the developmental stages of the plant.

4.2.2. Morphological Data Collection

Among the 202 anchote accessions planted, 20 accessions either failed to germinate or failed to survive until one month of age. These include ANC 16, ANC 20, ANC 42, ANC 44, ANC 64,

ANC 66, ANC 83, ANC 121, ANC 122, ANC 153, ANC 155, ANC 163, ANC 165, ANC 166, ANC 185, ANC 190, ANC 196, ANC 197, ANC 198, and ANC 199. In addition, 53 accessions were not represented at all the three experimental environments and thus excluded from analyses.

Descriptors for morphological characterization of anchote accessions have not yet been established well. Therefore, descriptors from different sources, i.e., those developed for related families of root and tuber crops were combined with some root and fruit descriptors unique to anchote such as root color, root length and diameter (Table 2).

Most of the descriptors used were adopted from Esquinas-Alcazarand *et al*. (1983), standard descriptor sets of Cucurbitaceae devised by the International Plant Genetic Resources Institute (IPGRI), with little modifications.

Totally, 28 (nine qualitative and 19 quantitative) traits were scored for this study (Table 2). Measurements and observations of quantitative and qualitative traits were taken. Qualitative data were recorded visually, on a group or a plot basis (VG) or as absence or presence (VS) of a trait's value assigned for a plot (Table 2). Except days to 50% emergence (DE), ground cover (GC), and days to male flower opening (DMFO), the rest discrete traits were scored by visual inspection of each phenotypes occurring in each accession (presence or absence of each phenotype $^{(VS)}$), whereas the record assigned for the former three traits represent group's phenotype (group observation $^{(VG)}$) (Table 6).

Quantitative traits were scored on five randomly selected individual plants per plot (accession) and averaged.

Qualitative traits include: days to 50% emergence (converted to interval category) (DE) (juvenile survival trait, is number of days from sowing date till the emergence of 50% of the plants), days to male flower opening (DMFO) (is number of days from sowing date till the emergence of the first male flower of any plant of the accession), primary leaf-lobe (PL) (major lobe of a plant), leaf green color intensity (LGC), growth habit (GH), ground coverage (GRC), stem pubescence density (VP), root color (RC), and stem color (STC) (Table 2).

Measurements of the following quantitative traits were taken. Number of main branch(s) (VN), leaf length (from apex tip to petiole) (LL), leaf diameter (longest distance across leaf's main rib) (LD), leaf length to diameter ratio (LLDR), petiole length (from stem node to leaf base) (PEL), internodes length (length of the third inter-node, starting from plant's base) (INL), petiole length to internodes ratio (PINR), root length (RL), root diameter (RD), leaf length to diameter ratio (RLDR), root yield (YLD) (calculated based on the fresh weight of sample of five plants per plot), dry matter content (DMC) ($1 - MC$ (where $MC = \frac{\text{Fresh weight} - \text{oven dry weight}}{\text{fresh weight}} \times 100$), average root weight (AVRWT), fruit average weight (FRAWT), fruit length (FRL), fruit diameter (FRD), fruit length to diameter ratio (FRLDR), peduncle length (PEDL), and number of seeds per fruit (NSFR).

Table 2. List of the morphological descriptors used to characterize anchote accessions

Character	Phenotypic classes (for discrete qualitative traits)	Code
Phenological traits		
days to 50% emergence	Early (DEE), medium (DEM), and late (DEL)	DE
Days to male flower opening	Early (DMFOE), medium (DMFOM), and late (DMFOL)	DMFO
Vine (stem) traits		
Number of main vine (branching)		VN
Vine color	Light green (STCLG), green (STCG), deep green STCDG), and purple (STC (STCP)	STC
Vine tip pubescence	No (VP0), sparse (VP3), heavy (VP5), and very heavy (VP7)	VP
Inter-node length (cm)		INL
Growth habit	Bushy (GH1) and runner (GH2)	GH
Ground coverage	50% (GRC3), 50-70% (GRC5), 70-90% (GRC7), >90% (GRC9)	GRC
Leaf traits		
Primary lobe	No (PL0), weak (PL1), medium (PL2) and deep (PL3)	PL
Leaf color	Light green (LGCo1), Green (LGCo2), deep green (LGCo3)	LGC
Leaf length (cm)		LL
Leaf diameter (cm)		LD
Leaf length: Leaf diameter		LLDR
Petiole length (cm)		PEL
Inter-node length: Petiole length		PINR
Root traits		
Root length (cm)		RL
Root diameter (cm)		RD
Root Length: Diameter ratio		RLDR
Root color	White (RCW) and reddish (RCR)	RC
Root yield (t/ha)		YLD
Average root weight (gm)		AVRWT
Dry matter content (%)		DMC
Fruit and seed traits		
Fruit weight (gm)		FRWT
Peduncle length (cm)		PEDL
Seed number per fruit		NSFR
Fruit length (cm)		FRL
Fruit diameter (cm)		FRD
Fruit length to diameter ratio		FRLDR

Note: DE: early<10 days, medium 10-13 days, late >13 days; **DMFO**: early<97 days, medium [97-127] days, late >127 days; **VC**: Light green = 1, green = 2, deep green = 3, and purple = 4; **VP**: No = 0, sparse = 3, heavy = 5, and very heavy = 7; **PL**: No lobe = 0, weak lobe = 1, medium lobe = 2, strong lobe=3; **LGC**: light green=1, green = 2, deep green = 3; **GH**: bushy = 1, runner = 2; **GRC**: <50% = 3, 50-70% = 5, 70-90% = 7, >90% = 9; **RC:** Reddish = 1, white = 2.

4.2.3. Morphological Data Analyses

Morphological data analysis was carried out on anchote materials grown from seeds at APPRC and HARC during the main cropping seasons and material grown at HARC from roots harvested from plants grown from seeds the previous year. The analysis was done on the basis of individual experimental environment as well as by combining the experimental environments.

The statistical analyses done include: descriptive statistics, the analysis of variance (ANOVA), computation of correlation coefficients, estimation of components of variances, discriminate analysis, principal component analysis, and clustering using Statistical Analysis Software (SAS) ver. 9.1. The model assumed was General Linear Model (GLM).

Minitab ver. 14.13.0.0 software was used for clustering as well as principal component analysis of qualitative traits, which were converted into binary matrix, based on Average Linkage and Squared Euclidean Distances (Minitab, 1998).

The qualitative data were converted into both frequency and binary forms. The frequency data were used to compute Shannon-Weaver diversity index (H) using the following formula:

$$\mathbf{H} = -\sum_{i=1}^{n} \boldsymbol{pi} \ln \boldsymbol{pi},$$

Where, H is the Shannon-Weaver diversity index of a given character which lastly standardized to H' as a ratio of H by In(n), n is the number of phenotypic classes in a given categorical character and *pi* is the relative proportion of the total number of entries (N) in the i^{th} class *i* (1, 2…n). The converted binary data were used to calculate the Euclidean distances, which further used for constructing clusters.

$$d_{i,j} = [(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots (x_p - y_p)^2]^{1/2}$$

to the smallest) before being

$\frac{\textbf{larger MSE}}{\textbf{Small MSE}}$, i.e., F ratio less

Descriptive statistics were used to compare accessions by their means, range, coefficients of variations of each trait; to construct graphs, cross-tabulation with Chi-square tests, frequency tables, and for principal component analysis. Mean separation for each categorical variable such as accessions, Woredas and populations as well as for experimental conditions were done using Duncan Multiple Ranges Test (DMRT).

Separate and combined ANOVAs, by using general linear model (GLM) procedure of the SAS software package (SAS, 2002-2004) were done to evaluate the variations due to different factors such as experimental environment, accessions, and by the interactions between the accession and experimental condition.

Correlation and regression coefficients (with the formulae **r = Covxy/sqrt (δ_x^2+ δ_y^2), b = cov xy/var x,** respectively) were analyzed to estimate relationships between the variables (Gomez and Gomez, 1984).

Genotypic and phenotypic correlations coefficients were computed using PROC CANDISC procedure of SAS software. Genotypic correlation between traits indicates the magnitude and direction of correlated responses to selection, the relative efficiency of indirect selection, and permits calculation of optimal multiple trait selection indices (Falconer and *Mackay, 1996).

The procedure PRINCOMP was used to perform principal component analysis (PCA) using quantitative traits. PCA, using correlation matrix, was conducted to project the data into lower dimensions and to show genetically similar genotypic clusters. It was also used to show the traits, which were responsible for the significant variations in anchote accessions. The accessions

were then plotted on two dimensions of the first two PCs (PC1 and PC2) to see pattern of their distribution.

Accessions were regrouped, for cluster analysis, into 35 Woredas of their origin, 16 populations (two or more closer woredas, expected to share common markets as a population) and six altitudinal groups. The latter was based on Agrawal's (1996) formula for altitudinal class (W): W = (L-S)/K, where L and S are the largest and smallest altitude, respectively, K is constant [$(1+3.32\log_{10}n$, n is the number of entries or sample size].

Cluster analyses using standardized means of quantitative data were done to determine the nature of accessions' grouping as a function of genotypes (accessions), geographic origins and altitudinal range.

Components of variances were calculated with restricted maximum likelihood (REML) method, which does not suffer from the problems of the possibilities of obtaining the estimates outside of parameter bounds and ignoring the estimator's distributional properties when data are unbalanced or missing (Holland, 2006), by the following formulae:

1. Genotypic variance $(\delta^2_g) = (MS_g\text{-}MS_{gl})/rl$
2. Genotype by location interaction variance $(\delta^2 gl) = (MS_{gl} - MS_e)/r$
3. Phenotypic variance $(\delta^2_p) = \delta^2_g+(\delta^2_{gl}/l)+\delta^2_e/rl = MSg/rl$
4. $\delta^2_e = MS_{e,}$

where MSg- mean sum of square of genotype, MSgl- mean sum of square of genotype by location (experimental environment) interaction, MSe- mean sum of square of error.

Genotypic coefficient of variance (GCV), phenotypic coefficient of variance (PCV), environmental coefficient of variance (ECV), and variance of interaction between genotype and environment (GECV) were derived from the components of variances. The broad sense heritability (h^2) and genetic advance (GA) were also calculated as follows:

GCV = $\sqrt{(\delta^2_g)}$*100/grand mean, PCV = $\sqrt{(\delta^2_p)}$*100/grand mean, ECV = $\sqrt{(\delta^2_e)}$*100/grand mean, GECV = $\sqrt{(\delta^2 gl)}$*100/grand mean, $h^2 = (\delta^2_g)/(\delta^2_p)$, R = I*$h^2$*($\sqrt{(\delta^2_p)}$ where I = selection differential (2.06 for selecting 5% genotypes), GA% = (R/grand mean)*100 (Singh and Chaudhary, 1985).

4.3. Molecular Markers Based Genetic Variation Analyses Methods

4.3.1. Leaf Sample Collection and Genomic DNA Extraction

Leaf sample collection

Bulked samples of young leaves from three to five individual plants of 182 anchote accessions, grown from seed at Ambo and Holeta sites, in the year 2013/14, during the main cropping season, were collected for total genomic DNA extraction following Gilbert *et al*. (1999). The leaf samples were collected in zipper bags with 1:10 ratio of leaf samples to silica gel that facilitate the drying of leaves and then stored under room temperature at Addis Ababa University, Arat-Kilo Campus (Cytogenetic Lab), until used for DNA extraction.

In addition to the above 182 accessions, leaf samples from four accessions (ANC 204, (Al), ANC 216 (Mn), ANC 227 (Hi), and ANC 232 (As), which were replanted from anchote roots (gubo) in a glasshouse were used. Two additional samples (A–1 and A–2) collected from Ambo town

and look like wild relatives of anchote, and might belong to one of *Coccinia* species, *Diplocyclos palmatus,* were also included for molecular analysis as out–groups.

In addition to the 186 bulk leaf samples mentioned above, fresh leaf samples were separately collected from 150 individual plants, representing 30 accessions which we grouped into 15 populations on the basis of the proximity of their collection sites. The collection and drying methods followed the same procedure as for bulked samples of 186 accessions explained above.

DNA Extraction

Silica gel dried leaves were ground to fine powder using mixer mill (MM 400). For 186 bulked leaf samples and the two samples (possibly belongs to *Diplocyclos*) (Table 3), the total genomic DNA was extracted, at Addis Ababa University, using a modified triple cetyltrimethyl ammonium bromide (CTAB) extraction technique (Borsch *et al*., 2003). The second extraction was selected based on concentration and quality of DNA as measured by Nano-Drop (ND-1000 Spectrophotometer) and gel test-on agarose gel electrophoresis.

Genomic DNA extraction, from leaf samples of 150 individual plants collected separately and one sample collected from a wild relative of anchote (A-2), was done in Sweden at Swedish University of Agricultural Sciences (SLU) following Geleta *et al.* (2012) with minor modifications, i.e., the extraction was done only once. These samples comprising of 30 Woredas constitute 15 populations. The intactness and quantity of extracted g-DNA was tested on 1.5% (w/v) agarose gel electrophoresis first, and using Nano-Drop before and after RNase treatment. However, four samples (ANC 28D, ANC 92A, ANC 182E and ANC187E) were discarded due to their low quality and quantity of genomic DNA (Table 3).

Table 3. Anchote accessions (146 separate plant's leaf samples and 33 accessions of bulk leaves samples plus three relatives) with population code from which genomic DNA was extracted

Sample code	Population code	Sample size	Sample code	Population code	Sample size	Sample code	Population code	Sample size
10A-10E	Ay-La	10	143A-143E	Gg-Dg	10	ANC 58	Illu	Cont'd (8x)
5A-5E			157A-157E			ANC 68		
15A-15E	Ds-Dw	10	171A-171E	Ar-Ld	10	ANC 75		
19A-19E			129A-129E			ANC 86		
21A-21E	Sy-An	9	178A-178E	Ss-Wt	9	ANC 89	Ji	7x
28A-28D			182A-182D			ANC 94		
33A-33E	Gm-Nj	10	184A-184E	Ac-Bt	9	ANC 99		
36A-36E			187A-187D			ANC 111		
39A-39E	Mt-Al	10	A-2	Relatives	1	ANC 117		
45A-45E			ANC1	Ww	6x	ANC 191		
52A-52E	Hr-Yy	10	ANC 6			ANC 216		
54A-54E			ANC 9			ANC 123	Ew	9x
61A-61E	Ch-De	10	ANC 32			ANC 145		
72A-72E			ANC 37			ANC 160		
74A-74E	Gc-Dh	10	ANC 193			ANC 167		
87A-87E			ANC 13	Kw	3x	ANC 172		
92B-92E	Gu-Go	9	ANC 17			ANC 180		
96A-96E			ANC 27			ANC 187		
98A-98E	Gr-Sh	10	ANC 41	Illu	8x	ANC 227		
120A-120E			ANC 47			ANC 232		
108A-108E	Dd-Dc	10	ANC 50			A-1	Relatives	1
186A-186E			ANC 55			A-2	Relatives	1

NB: Letter A-E = Five different individual plants (A, B, C, D & E) of the same accession; Arabic numeral preceding letters or without letter 1, 2, 3…represents accession numbers, whereas A-1 and A-2 represent wild "relative" of anchote. Letter "x" after numerals indicates the unknown number (3-5) of bulked leaves
Ww = West Wollega (La, Gl, Ay, Gm, Nj & Gd); Kw = Kellam Wollege (Ds, Dw & An); Ew = East Wollega (Gg, Dg, Ld, Wt, Nk, Ac, Hi, Dc, Mn, As); Ji = Jimma (Gu, Go, Gr, Dd, Sc, Dc, & Mn); Illu = IlluAbabBora (Mt, Al, Hr, Yy, Ch, De, Gc, & Dh); ANC = anchote accession.

4.3.2. EST-SSRs Screening and PCR

Forty-seven expressed sequence tagged (EST–SSR) primer pairs were designed from DNA sequences of watermelon [*Citrullus lanatus* (Thunb.) 2n = 22)], retrieved from public databases and used for the estimation of genetic variation existing among and within anchote accessions. EST-SSRs mining was done using the perfect SSRs repeats with minimum size of 12 bp and maximum 28 bp. Primer3 (Rozen and Skaletsky, 2000) software was used to design the primer.

The parameters for primer design were as follows: product length ranges from 163 to 400 bp, primer length ranges from 18 to 27 bp, and primer melting temperature ranges from 57.13^{o} C to 62.25^{o} C. All of the EST-SSR markers were coded as WM- followed by Arabic numeric to indicate watermelon origin and specific code for each primer, respectively.

EST-SSRs fragment amplification was done according to Don *et al.*, (1991) with minor modification of 25 µl total reaction volume with the following proportions: 2.5 µl of 1xPCR buffer (10 mM Tris–HCl pH 8.3, 50 mM KCl including 1.25 mM $MgCl_2$), $MgCl_2$ (1.5 µl), dNTPs (0.3 µl), forward and reverse primers (0.75 µl each), 1.0 U (0.2 µl) Dream Taq polymerase (Sigma, Germany), template DNA (2.5 µl = 25 ng), and 16.5 µl ddH_2O (Table 4).

PCR amplifications were performed in 96-well plates using Thermal Cycler (S 1000^{TM}) machine with the following touchdown PCR (Td 58^{o}-48^{o} C) conditions: initial 95^{o} C for 3 min preheating, denaturation at 94^{o} C for 30 sec, annealing at 58^{o} C for 45 sec, and extension at 72^{o} C for 45 sec. The later three steps (complete PCR cycles) were repeated nine times by reducing the annealing temperature by 1^{o} C for each subsequent cycle. The rest 30 cycles were continued, each with 94^{o} C denaturing, 58^{o} C annealing, 72^{o} C for extension, and additional 3 min at 72^{o} C for the last

extension. The PCR reaction mix and products had been stored at 4^{o} C and -20^{o} C for short and long-term storages, respectively.

PCR amplified products stained with ethidium bromide were tested on 1.5% agarose gels in 1xTAE buffer. The gels were photographed under gel documentation system and the fragment sizes were compared with 50-bp DNA ladder (GeneRuler™, Fermentas Life Sciences), which is used as molecular size standard.

The larger bands (expected fragment length: WM-23, WM-28, WM-37, WM-42, WM-43, and WM-45), were excised, under UV radiation, and purified using extraction kit (QIA-quick gel). The purified DNA was used as a template for second round of PCR using the same primer pairs and PCR profile used before. These reamplified PCR products were also separated on a 1.5 % (w/v) agarose gel and purified by the same method described above. These purified amplicons were sequenced and the sequence data were analyzed using Web Sat, web-based front-end, in searching for the repeat motifs (Martins *et al*., 2009).

After screening, the functional EST-SSRs (thirteen primer-pairs) the forward primers were labeled on 5′-end with carboxyfluorescein (6-FAM™) or hexachlorofluorescein (HEX™) fluorescent dyes, whereas the reverse primers were tailed with GCTTCT to reduce polyadenylation and improve genotyping. Fluorescent dyes help in visualizing and the reverse primer's tail is intended to prevent non-template addition by *Taq* polymerase to the PCR products (Ballard *et al.*, 2002). Thirteen labeled primer pairs were used for PCR amplifications of 182 samples by each primer pairs (Table 6), which include a total of 146 anchote samples and one sample of *Diplocyclos* (A-2) whose DNA was extracted from separate plants (at SLU, Sweden) and 33 anchote samples together with two *Diplocyclos* samples (A-1 and A-2) of which

bulked DNA samples were extracted at AAU, Ethiopia. Negative control (distilled water) was included during PCR to check or to control contamination related errors. For confirmation of each primer pairs, five samples of each PCR products, together with 50 bp DNA ladder were re-tested by 1.5% agarose gel electrophoresis.

The amplified products (26 plates- two plates for each primer in order to accommodate 182 samples and ten negative controls) were multiplexed into six micro-plates. Multiplexing was undertaken based on the fragment sizes difference of at least 50 bp and the color of labeled primers. Capillary gel electrophoresis was done using ABI Prism® 3730xl genetic analyzer (Applied Biosystems), at the Department of Plant and Environmental Sciences, University of Copenhagen, Denmark. The peak identification and fragment size determination were done using GeneMarker software ver. 2.4.0 (2012). The sizes of PCR products were determined based on the Genescan-500 LIZ internal size. Genotypic data were archived in Excel tables for further analyses after arranging into different formats by convert software (Glaubitz, 2005).

The larger DNA fragments obtained from PCR, with six EST-SSRs marker primers, were sequenced and proved, by Web Sat online software, that they have no repeat units and we, therefore, unable to design EST-SSR primers, out of those sequences, to develop them for anchote diversity and differentiation analyses. The criteria to determine the repeat motifs include di-nucleotides $\geq$ 6, tri-nucleotide $\geq$ 4, tetra and above nucleotides $\geq$ 3. However, using our query sequences we retrieved very few similar sequences from public database (NCBI) for phylogenic analysis, hence we were not able to proceed to further analyses. The sequences are made ready for submission to database.

Table 4. PCR conditions of EST-SSR markers used for diversity analyses of *C. abyssinica*

Solution	Amount (µl) for one sample
dH_2O	16.50
1xPCR buffer (with 1.25mM $MgCl_2$)	2.50
$MgCl_2$ (1.5mM)	1.50
dNTPs (100mM)	0.30
Forward primer	0.75
Reverse primer	0.75
Taq polymerase	0.20
Sub total	22.50
Sample DNA	2.50
Total	25

Table 5. Functional EST-SSR primer pairs size, annealing temperature and source

No.	Locus name	product size	Forward primer (labeled)	T_a (^{O}C)	Reverse primer	T_a (^{O}C)	Gene I.D. or Origin
1	WM-5	189	[6FAM]GCAAGCTGCCTTCATTCTCT	59.72	GCTTCTAAAAGTGGGGGTGATAAGAGG	59.33	4521831
2	WM-8	390	[HEX]GCGTCAAGAAATTCCCTCTC	58.86	GCTTCTTTGAAACTTGGAATGGATTGG	59.79	327341253
3	WM-11	372	[6FAM]CAACGAAAAAGAAAACACATCC	58.65	GCTTCTAAGGAGAAGCTTCCAGGACA	59.01	327340444
4	WM-12	284	[HEX]ATCGAGGGATATGTGGTGGT	59.09	GCTTCTCCCCCATTTACTCAATTTCC	58.23	327340438
5	WM-14	297	[FAM]GAAAAAGAAAACACATCCCTTCC	60.21	GCTTCTGAGGCTGCTCCTCATCATCA	61.49	327340277
6	WM-15	283	[6FAM]ATGTGAAATGAGGAAGGAAAAG	57.35	GCTTCTATACCCAAATCTATATCCCAAAAA	57.13	327340267
7	WM-24	240	[6FAM]AGCAGCTGCAGCAGTGAAAT	61.29	GCTTCTGATGACCTTAAAAGATATTCTACAAGC	57.40	327339800
8	WM-25	242	[HEX]AACCTCTCACACACAAACAAAAA	58.71	GCTTCTCGATGCGATTGAGAACGAA	60.91	327339787
9	WM-29	300	[6FAM]TGGGTTGTTGCACAAATTGA	60.95	GCTTCTAAAATGAATTCCCAAGAAAGATG	58.55	327339617
10	WM-30	236	[HEX]GACAATCCAATACCCTTTACATTT	57.53	GCTTCTTTTGCCTTCATGTTTGCGTA	60.25	327339440
11	WM-32	400	[6FAM]CCTCCAAACTCATCATTACCC	58.37	GCTTCTTCAATCTCAATGGACATGAAAGA	59.56	327339252
12	WM-34	245	[6FAM]TTGGCTTTTGGATTCTGCTT	59.82	GCTTCTTAAGCCAAGACCCCTTGATG	60.07	327339192
13	WM-46	300	[HEX]CAAAATGAACACTATGACAACGAA	59.10	GCTTCTCATGAAGAAGATCAAGGAAAGGA	59.71	327338173

4.3.3. EST-SSR Data Analyses

4.3.3.1. EST-SSRs Based Genetic Diversity and Population Differentiation Analyses

Analyses of genetic variation and differentiation of anchote populations were performed using various software packages. Most of them compute similar genetic parameters, which in fact help us to verify our results and some of them compute specific and/or additional information.

Genetic diversity parameters such as observed and effective number of alleles accessed from each marker (N_a and N_e, respectively) were done using POPGENE ver. 32 software package (Yeh *et al*., 1999); percentage of polymorphic loci, Shannon–Weaver diversity index (I), homozygosity (H_o), heterozygosity (H_e), random segregation and distribution (Hardy-Weinberg equilibrium) of each genotype within populations for each locus, Nei's genetic identities (J_i) and genetic distances (Ds), population differentiation (F-statistics), gene flow (as a number of migrant per generation between pairs of populations-N_m), and dendrogram based on Nei's (1972) genetic distance by unweighted pair-group method with arithmetic average method (UPGMA) were also computed by the same software. The resulting dendrograms were constructed and visualized by MEGA ver.5.2.2 software (Tamura *et al*., 2011).

Polymorphic information contents (PIC) for each primer pair (or eight loci) were estimated using POWERMARKER ver. 3.25 software (Liu and Muse, 2005) to evaluate the efficiency of EST-SSR marker for accession discrimination with the formula: **PIC = 1-$\sum_{i=1}^{k} pi^2$,** where k is total number of alleles detected for an SSR marker, *pi* is the frequency of i^{th} allele (Anderson *et al*., 1993).

Population genetic diversity analyses including hierarchical Analysis of Molecular Variance (AMOVA) and population differentiation, genotype assignment, and linkage disequilibrium were calculated by ARLEQUIN ver. 3.01 software (Excoffier *et al.*, 2006).

The population structure analysis by clustering method and determination of the correct K were done using STRUCTURE ver. 2.3.4 software using admixture model. The model assumed that, there are K populations each of which has unique sets of allele frequencies for each locus. Therefore, each individual was assigned to population based on its membership coefficient (Q) for each cluster.

The STRUCTURE running parameters were: burn-in periods of 50,000 plus 450,000 length Markov Chain Mont Carol (MCMC) repetitions each of which replicated ten times for each K (K = 1 to 15) as recommended by Gilbert *et al.* (2012) and Evanno *et al.* (2005). Optimum K value (maximum ΔK), estimation of different genetic groups, was determined by the STRUCTURE HARVESTER online software. It was determined based on the rate of change of the log-likelihood between successive K values. The obtained ΔKs were plotted against the K number of groups, so that the maximum (peak) was selected as correct cluster (or K value) (Evanno *et al*., 2005). Clusters alignment, across the replicates, was done using CLUMPP software (Jakobsson and Rosenberg, 2007) and the population clusters were visualized by DISTRUCT software (Rosenberg, 2004). DISTRUCT software uses the outputs of STRUCTURE HARVESER (which is reprocessed by CLUMPP for individual file and population file separately) to cluster populations into predetermined Ks in left-to-right or top-to-bottom graphs separated by colors. It also printed out the labels atop/or below the graphs (Rosenberg, 2004).

Factor correspondence analysis (FCA) and population pair wise distances were done by GENETIX ver. 4.05.2 software (Cavalli-Sforza and Edwards, 1967).

F_{STAT} software ver. 2.9.3.2 program was used to calculate allele frequency per locus and population, gene diversity per locus and population, allelic richness (based on minimum sample size-to standardize the values) and their random distribution test (testing for Hardy-Weinberg proportion using the statistic F_{IT}), testing for population differentiation using G-test per locus (Goudet *et al.,* 1996). It was also used for fixation index (F_{IS}) estimation per locus and population as well as for overall loci per population, Nei's estimation of heterozygosity, Weir & Cockerham (1984) estimation of F_{IT}, F_{ST} and F_{IS}, relatedness, inbreeding corrected following Pamilo (1985), components of variances among samples, among individuals within samples, and within individuals, Jackknpifing over population and loci, Bootstrapping over loci.

Allele frequency based correlation (F_{IS}) and diversity (Weir and Cockerham, 1984) as well as Hardy-Weinberg exact tests by locus and by population were also computed using GENEPOP ver. 4.0.10. Additionally, frequency of null alleles per locus per population and genotypic linkage disequilibrium were done using the same software (Rousset, 2007).

Excel add-ins (GENAIEX) ver.6.502 software (Peakall and Smouse, 2012) was applied for Shannon statistics, assignment of individuals into appropriate population by log-likelihood method, pair wise population assignment by graph, co-dominant genotypic distance that was used for Principal Coordinates Analysis (PCoA, using Dice similarity coefficient). An analysis of molecular variance (AMOVA), test for the significant genetic variation among groups and pair wise population differentiation (F_{ST}), and pair wise gene flow between populations (N_m) were also calculated using this software.

BOTTLENECK software ver.1.2.02 (INRA, 1999) was used to test for any significant heterozygosity excess (deviation from mutation-drift equilibrium) that would be considered as the result of recent drastic change in effective population size. Heterozygosity excess can be computed from the difference between the observed heterozygosity and the heterozygosity expected from observed number of alleles (Cornuet and Luikartt, 1996).

We used Dissimilarity Analysis and Representation for windows (DARwin) ver. 6.0.112 software (2014) for clustering using UPGMA and neighbor joining (NJ) methods of individual genotypes.

Clustal X ver. 2.0.11 software was used for DNA sequence alignment and for saving the outputs in different formants (Larkin *et al.,* 2007). BioEdit software was also used for edition and sequence saving (Hall, 1999).

4.4. Cytogenetics and Karyotype of Anchote

4.4.1. Collection of Root Tips and Slide Preparation

Anchote seeds were germinated on Petri-dish using wet filter paper. Actively growing radicles were collected into a vial.

The collected radicles were subjected to standardized pretreatment, fixation, and maceration before using the tips for slide preparation.

Accordingly the radicles treated by colchicines (0.1%) for about three hours at room temperature. The radicals were fixed in3:1 ethanol: glacial acetic acid for about 12-24 hrs. After thorough rinsing in distilled water, the radicles were macerated in mecerozyme containing 4%

cellulase and 4% pectinase at 37° C in a water bath for about one hour or more until the tips were detached from radicles due to the enzyme digestion.

After decanting the maceration enzymes the tips were rinsed with distilled water, pipetted onto a glass slides and meshed in drops of fresh fixative. The slides were kept away for air dry at room temperature.

Staining was done by immersing air dry slide preparation in staining jar using Giemsa's stain (in phosphate buffer, pH 6.8) for 15-30 minutes or more until satisfactory staining was obtained. The slides were rinsed in distilled water, air-dried, and mounted under a 22x55mm cover slip in-Dibutylphthalate Polystyrene Xylene (DPX) mounting medium (Kifle Dagne, and Heneen, 1992).

Slides with better chromosome spreads were selected and photomicrographs of metaphase plates were taken under x100 objective oil immersion microscope. The correct diploid (2n) chromosome number was determined from several mitotic metaphase chromosome counts.

4.4.2. Karyotypic Analysis

Metaphase spread chromosome pictures were scanned into a computer and used for karyotypic analysis using Micro-Measure version 3.3- computer program. Based on the obtained arm ratio values, the centromeric positions were determined using Levan *et al*. (1964) recommendation, in which chromosomes are divided as M, m, sm, st, t and T. This respectively refers to chromosomes having the centromere exactly at median point, median region, submedian region, subterminal region, terminal region and exactly at terminal point.

5. RESULTS

5.1. Agro-Morphological Traits Variation

A total of 181 and 149 anchote accessions were used to assess genetic variability for qualitative and quantitative agro-morphological traits, respectively, in this study and significant variability within and among anchote accessions have been observed. The influence of geographic origin (variances between anchote populations) and experimental environments was also immense for some traits.

5.1.1. Qualitative Traits

5.1.1.1. Description of Phenotypic Traits

Different anchote morphotypes have been observed in this study. These were growth characters, leaf, stem, and root traits (Figures 2-7). These figures show the morphological diversity between and within accessions, since each phenotype of a trait has been observed in most of the accessions.

Qualitative data, collected visually, are presented in the form of table and bar charts, revealing frequencies and relative percentage of each phenotype occurring in anchote accessions in combined forms and percentage of accessions possessing a particular phenotypic trait (Table 6 and Figure 8). High Shannon-Weaver indices (H') was observed in most qualitative traits, whereas some of the traits are occurring in similar proportion, almost in all accessions (with more than 75% rates).


A
B
C


D

A B C D

E F G

H I


A
B
C
D
E
F


A
B


A
B
C


A
B
C
D
E

Table 6. Frequency and percentage of qualitative morphological traits observed in 181 anchote accessions during its different phenological stages

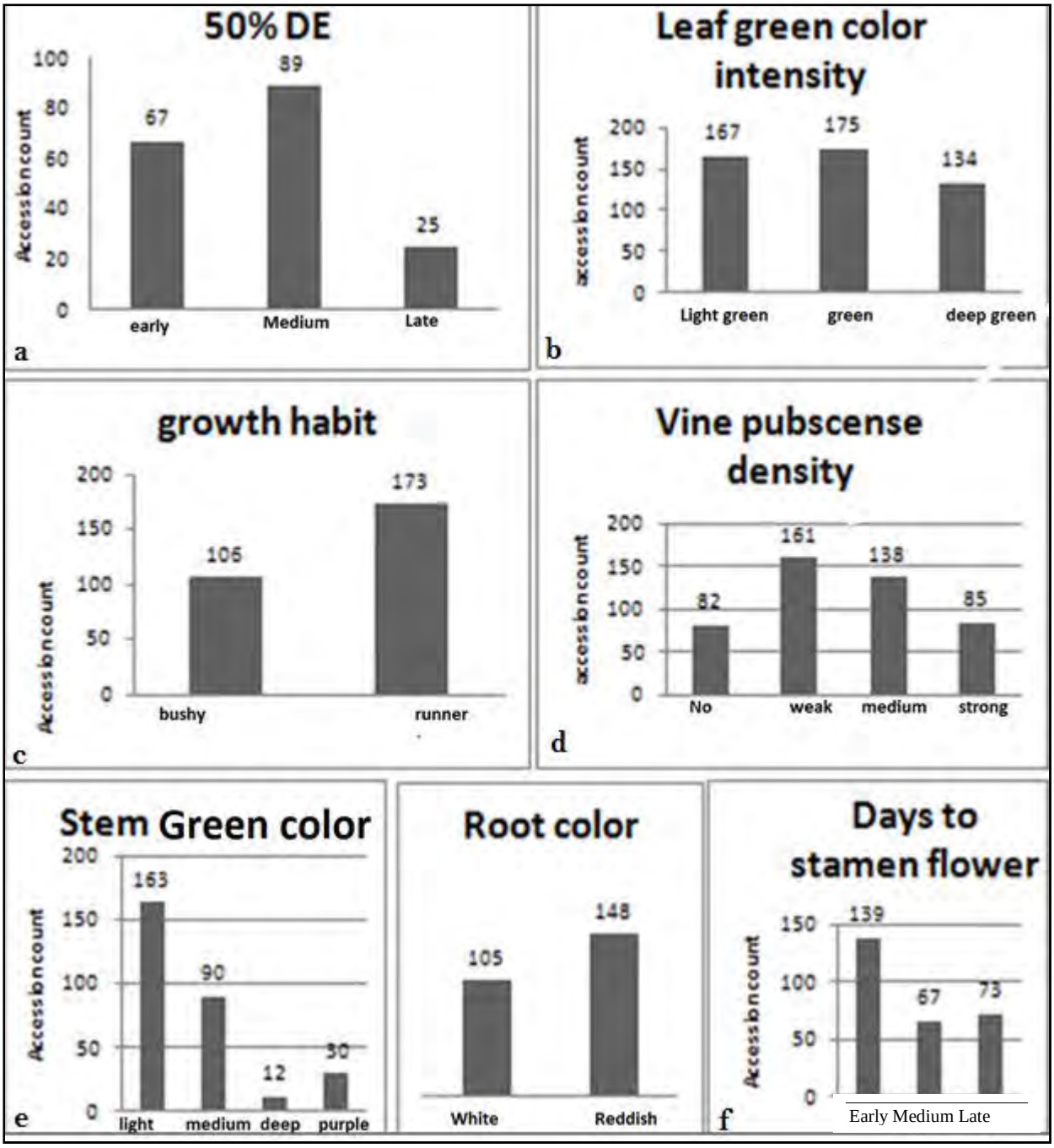
Trait#	Phenotypes*	code	No. of accessions (Freq.)	Relative % of each phenotype	% of accessions possessing a phenotype	H'
50% DE^{VG}	Early (< 10 days)	1	67	37.000	37.000	0.901
	Medium (10-13 days)	2	89	49.200	49.200	
	Late (> 13 days)	3	25	13.800	13.800	
PL^{VS}	No lobe	1	152	37.530	84.000	0.965
	Shallow lobed	2	165	40.740	91.000	
	Deep lobe	3	88	21.730	41.000	
GH^{VS}	Bushy type	1	106	37.990	58.600	0.601
	Runner	2	173	62.010	95.600	
GC^{VG}	Ground cover<50%	0	0	0	0	0.592
	Ground cover (50-70]%	3	8	4.420	4.400	
	Grong cover (70-90]%	5	66	36.460	36.500	
	Ground cover>90%	7	107	59.120	59.100	
VP^{VS}	No or very sparse	0	82	17.600	45.000	0.967
	Sparse	3	161	34.550	89.000	
	Dense	5	138	29.610	76.000	
	Very dense	7	85	18.240	47.000	
RC^{VS}	Reddish	2	148	58.500	82.000	0.981
	White	1	105	41.500	58.000	
DMFO ^{VG}	Early	1	27	16.270	16.270	0.928
	Medium	2	72	43.370	43.370	
	Late	3	67	40.360	40.360	
STC^{VS}	Light green	1	163	55.250	90.000	0.757
	Green	2	90	30.510	50.000	
	Deep green	3	12	4.070	6.600	
	Purplish	4	30	10.170	17.000	
LGC^{VS}	Light green	1	167	35.080	92.270	0.992
	Green	2	175	36.760	96.690	
	Deep green	3	134	28.150	74.030	

#VG = visually observed data for a group (plot), VS = visually observed data for each sampled plants per plot
*50% DE = days to 50%emergence, **PL**= primary lobe, **GH** = Growth habit, **GC** = Ground coverage, **VP** = vine pubescence, **RC** = root color, **DMFO** = Days to male flower opening, **STC** = Stem color, **LGC** = leaf green color intensity, **RC** = root color, H' = standardized estimate of Shannon-Weaver diversity index

Table 7. Frequency (%) distribution of eight qualitative traits (24 phenotypes) within each anchote population

Population	Trait																							
	DE			PL			LGCo			GH		VP				STC				RC		DMFO		
	E	M	L	1	2	3	1	2	3	1	2	0	3	5	7	LG	G	DG	P	W	R	E	M	L
Ay-La	78	22	0	21	42	37	31	35	35	31	69	30	39	22	9	62	23	0	15	40	60	11	67	22
Ds-Dw	57	43	0	24	35	41	38	44	19	42	58	25	44	25	6	50	29	7	14	22	78	43	57	0
Sy-An	20	70	10	40	32	28	35	35	31	47	53	14	34	34	17	43	33	5	19	50	50	0	30	70
Gm-Nj	56	33	11	41	41	18	36	32	32	40	60	29	33	29	8	43	50	7	0	45	55	33	33	33
Mt-Al	22	22	56	39	44	17	40	45	15	47	53	18	27	23	32	46	38	0	15	36	64	13	50	38
Hr-Yy	56	44	0	41	41	18	35	35	31	46	54	15	35	31	19	50	50	0	0	46	54	11	44	44
Ch-De	8	77	15	43	46	11	34	34	31	31	69	8	28	36	28	57	29	5	10	37	63	14	57	29
Gc-Dh	29	29	43	41	45	14	32	37	32	43	57	6	35	32	26	63	32	0	5	58	42	57	43	0
Gu,Go, &Mn	50	50	0	32	42	26	35	35	30	38	62	12	29	32	27	65	25	10	0	47	53	14	36	50
Gr,Sc,& Sh	18	64	18	35	48	17	37	37	27	48	52	3	27	33	37	61	11	6	22	46	54	0	18	82
Dd-Dc	40	50	10	35	43	22	36	36	28	47	53	0	31	35	35	47	21	5	26	31	69	10	20	70
Gg-Dg	40	50	10	40	38	21	33	39	27	26	74	24	37	29	10	69	28	0	3	34	66	10	65	25
Ac-Bt	50	0	50	67	33	0	50	50	0	50	50	50	0	0	50	25	50	0	25	50	50	50	50	0
Ss,Wt,& Nk;	8	77	15	43	43	14	31	38	31	28	72	32	42	26	0	50	33	4	13	44	56	15	31	54
Ar-Ld	50	46	4	39	35	26	37	36	27	32	68	21	39	30	10	70	24	3	3	34	66	8	31	62
Gl-Gd	20	60	20	22	56	22	42	42	17	29	71	40	50	10	0	18	45	18	18	56	44	60	20	20
Total (%)	37	49	14	38	41	22	35	37	28	38	62	18	35	30	18	55	31	4	10	42	58	16	40	44

DEE, DEM & DEL (early, medium &late days to 50% emergence);PL1, PL2 &PL3 (weak, medium &strong primary lobe); LGCO (light 1, medium 2 &deep 3 leaf green color intensity); GH1&GH2 (bushy & runner growth habit); VP0,VP3,VP5,VP7 (weak, medium, strong vine pubescence density); STCLG, STCG, STCDG,& STCP (light, medium, deep & purple stem color); RCR&RCW (reddish &white root color); DFOE, DFOM & DFOL (early, medium &late flower initiation)


50% DE
Accessioncount
100
80
60
40
20
0
67
89
25
early
Medium
Late
a
Leaf green color intensity
accessioncount
200
150
100
50
0
167
175
134
Light green
green
deep green
b
growth habit
Accessioncount
200
150
100
50
0
106
173
bushy
runner
c
Vine pubscense density
accessioncount
200
150
100
50
0
82
161
138
85
No
weak
medium
strong
d
Stem Green color
Accessioncount
200
150
100
50
0
163
90
12
30
light
medium
deep
purple
e
Root color
105
148
White
Reddish
Days to stamen flower
Accessioncount
150
100
50
0
139
67
73
Early Medium Late
f

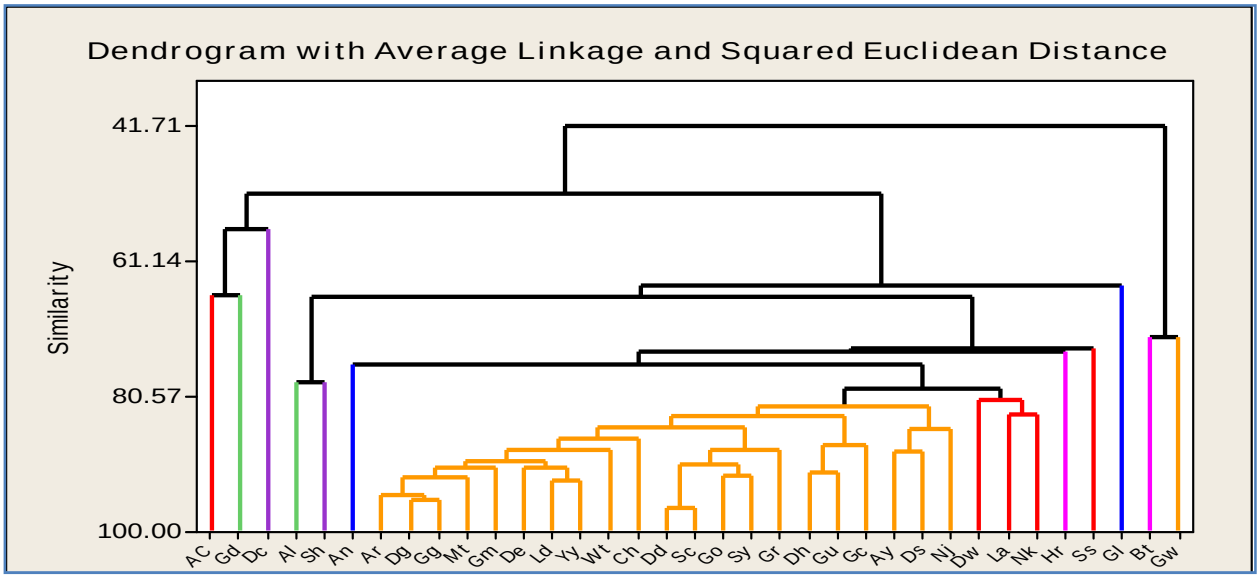
5.1.1.2. Cluster analysis Based on Qualitative Traits

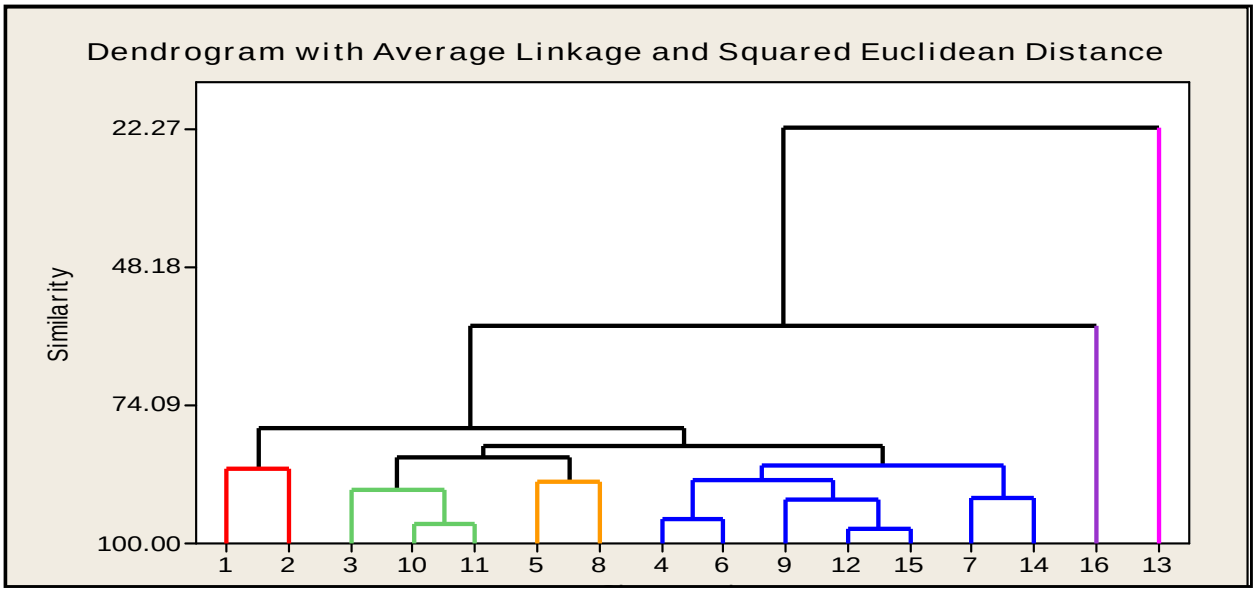
The qualitative trait data in the form of category was converted into binary data matrix in order to minimize biasness, because a variable with more categories receives more weight. Therefore, cross-tabulation and chi-square test as well as clustering based on the matrix computed from the frequencies of binary data assigned for each population was done by average linkage between groups (Figure 9 A-C).

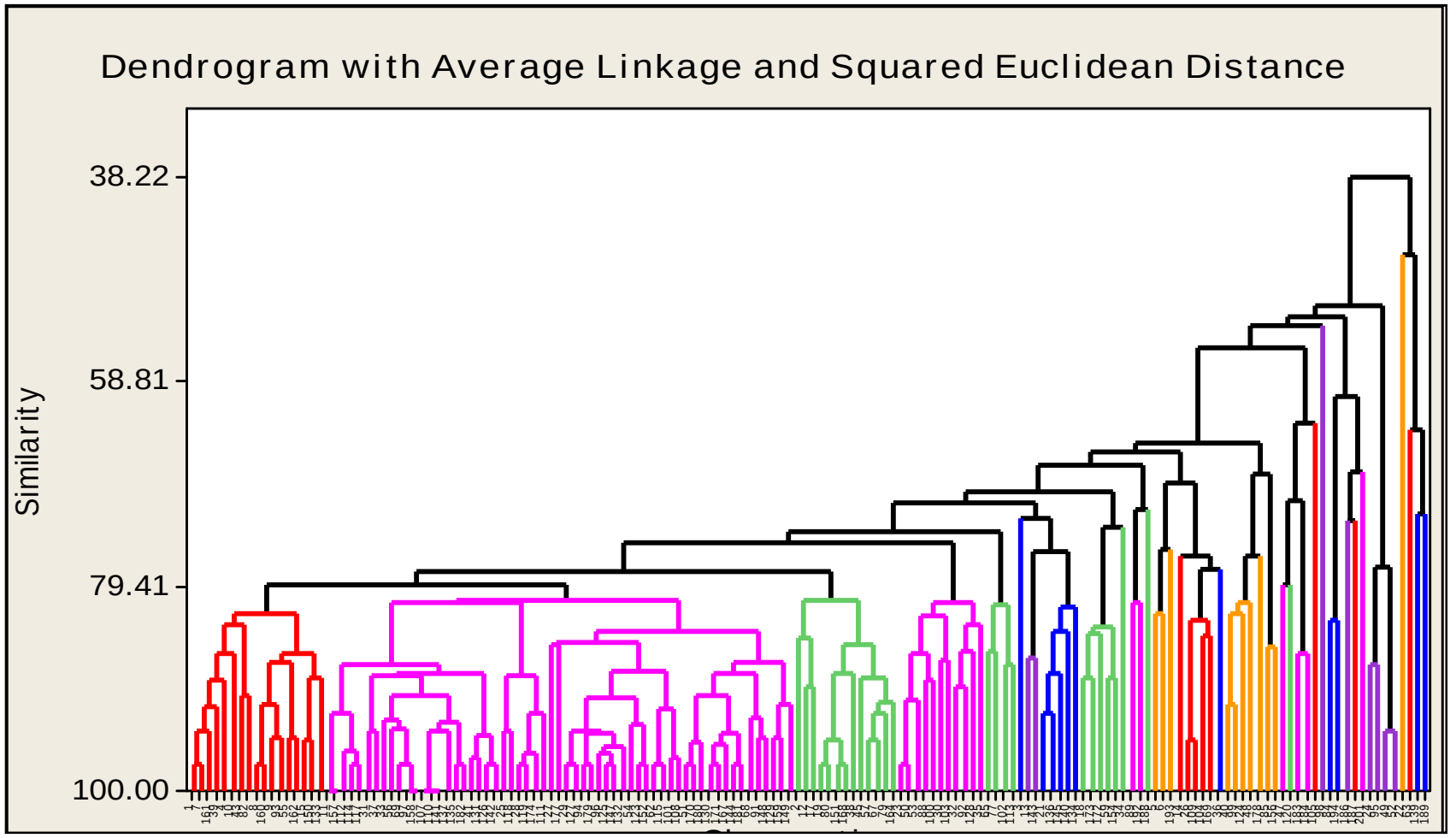
The chi-square test results (not shown) indicated that there are significant variations among populations ($p < 0.05$) in the following phenotypic traits: early days to 50% emergence (DEE), primary lobe (PL1 and PL3), growth habit (GH1), vine pubescence (VP0 and VP7), stem color (STCP), and days to male flower opening (DFOE, DFOM, and DFOL).

For district (Woreda) groups the tests were only significant for limited traits including late days to 50% emergence (DEL), primary lobe (PL1&PL2), leaf green color intensity (LGCo1& LGCo3), growth habit (GH1), vine pubescence (VP0,VP3,VP5, &VP7), stem color (STCLG, STCG, & STCP), and days to male flower opening (DMFOE and DMFOL).

Cluster analysis of 16-predetermined anchote populations, 35 Woredas, and 181 accessions showed, certain relationship between accessions geographic distribution, whereas accessions of some Woredas or populations did not form cluster with others, even at loose (80%) level of similarity. Thirteen clusters were formed for anchote accessions grouped under 35 Woredas (eleven woredas formed eleven separate groups, Ds, La, and Nk formed one group, and the rest were clumped together) (Figure 9A).


Dendrogram with Average Linkage and Squared Euclidean Distance
Similarity
41.71
61.14
80.57
100.00
AC Gd Dc Al Sh An Ar Dg Gg Mt Gm De Ld Yy Wt Ch Dd Sc Go Sy Gr Dh Gu Gc Ay Ds Nj Dw La Nk Hr Ss Gl Bt Gw


Dendrogram with Average Linkage and Squared Euclidean Distance
Similarity
22.27
48.18
74.09
100.00
1 2 3 10 11 5 8 4 6 9 12 15 7 14 16 13


Dendrogram with Average Linkage and Squared Euclidean Distance
Similarity
38.22
58.81
79.41
100.00

5.1.2. Quantitative Traits

5.1.2.1. Quantitative Data Exploration and Descriptive Statistics

Separate (by accession, Woreda, population, and experimental conditions) as well as combined mean performances of agro-morphological traits of anchote accessions, both grown from seed and root, show large variation. The comparison of range, standard deviation, and mean separation showed variations between all classes (accessions, origin, and experimental environment).

The mean separation between accessions, Woredas, and populations resulted in maximum of five mean groups while the minimum was two. In all the groups, the average root weight (AVRWT) showed largest variation followed by root yield (YLD) and dry matter content (DMC) (Appendix Tables 1-4).

The mean performances of some traits, such as LL, LD, PEL and INL, of anchote grown from roots were superior to those grown from seeds. However, the ratio (shape) variables, LLDR and PINR decreases and increases, respectively, i.e., the leaf diameter increased relative to leaf length and petiole length increase relative to inter-node length in anchote grown from root ("gubo"). Fruit and seed traits characterization was also done from root planted anchote at single (off-) season using irrigation. The mean number of seed per fruit was 112$\pm$25, with larger range of 42-177. Totally, 182 accessions were used for this evaluation and mean performances of certain traits have been computed (Appendix Table 5).

The two extremes, top 5-15 % of accessions with highest and lowest mean performance, for each trait are sorted out, in preparation for identification of the higher and lower performing

accessions (Appendix Table 6). Altitudinal based group mean performance was also done in searching for variation that arise due to elevation difference. The highlander anchote accessions were late emerging and were poor in most traits, but highest in DMC (Appendix Table 7).

5.1.2.2. Analysis of Variance and Components of Variances

For analysis of variance, the common assumptions such as normal distribution and variance heterogeneity were checked as indicated in section 4.2.3. Then, the data from three experimental environments were combined for multivariate analysis according to the technique used for analysis of variance (ANOVA) for the alpha lattice design (Patterson and Williams, 1976).

Combined ANOVA for 149 anchote accessions grown from seed at three experimental environments with two replications showed significant variations in most of the traits among accessions (Table 8). Separate and combined ANOVAs for some agro-morphological traits of fruit and seed of 182 anchote accessions, which were grown from root, are presented in Tables 9A & B. Except LLDR and INL the rest agro-morphological traits showed significant variations between accessions. Fruit and seed traits ANOVA indicated similar performances among for accessions.

Components of variances, phenotypic and genotypic coefficients of variances, broad sense heritability, and predicted genetic gains which were estimated following Singh and Chaudhary, (1985) methods are summarized in Table 10. Generally, moderate heritability of most traits were obtained. The most responsive trait upon selection is YLD, i.e., 21% improvement can be achieved through selection of this trait. In other way, only 1.4% response is possible, if selection will made following the performance of INL.

Table 8. Summary of combined ANOVA for anchote accession diversity study in alpha lattice design with 149 accessions and two replication at three environments (year 1 APPRC = Env1, Year 1 HARC = Env2, year 2 HARC = Env3)

Character	F-test							
	Env	Rep(Env)	Accn	Bl(rep)	Accn x Env	CV	S.E.	C.D
VN	ns	*	*	**	ns	40.12	0.87	1.70
AVRWT	**	*	**	**	ns	48.87	35.12	68.83
YLD	**	**	**	**	ns	42.39	7.45	14.61
DMC	**	**	*	ns	ns	27.69	5.38	10.54
LL	**	ns	**	*	ns	12.73	0.97	1.89
LD	**	ns	**	ns	ns	12.89	0.98	1.93
LLDR	**	*	**	ns	ns	6.77	0.07	0.13
PEL	**	ns	**	**	ns	18.29	1.84	3.60
INL	**	**	**	ns	**	29.36	0.96	1.89
PINR	**	**	**	ns	**	28.27	1.02	2.00
RL	**	**	**	**	ns	18.37	1.76	3.46
RD	**	**	**	**	ns	21.88	0.91	1.78
RLDR	**	ns	**	ns	ns	26.36	0.64	1.26

** = highly significant ($P < 0.01$), * = significant ($P < 0.05$) and ns = non significant ($P > 0.05$)
S.E = standard error of the difference of the accession means
C.D. = critical difference at 5% level of significance ($P = 1.96$) for the degree of freedom of error mean square (d.f. MSe > 120)

Table 9. Combined and separate ANOVAs for 182 anchote accessions, grown from root for two seasons at HARC (season 1 = off-season & season 2 = main season) with two replications for agro-morphology and seed traits

A

ANOVA for combined (two) seasons						
Trait	F-test					CV (%)
	Season	Rep(Season)	Bl(Rep)	Accn	Accn x season	
LL	**	ns	ns	**	*	13.44
LD	ns	ns	ns	**	*	12.12
LLDR	**	ns	ns	ns	ns	8.2
PEL	**	ns	ns	**	**	21.31
INL	**	*	ns	ns	ns	29.23
PINR	**	ns	*	**	**	31.63

B

ANOVA for single off- season				
Trait	Rep	Bl(rep)	Accn	CV (%)
FRL	ns	ns	ns	10.21
FRD	*	ns	ns	15.18
FRLDR	*	ns	ns	18.34
PEDL	ns	ns	ns	32.06
AVFRWT	ns	ns	*	22.12
NSPFR	ns	ns	ns	23.44

** = highly significant (P < 0.01), * = significant (P < 0.05) and ns = non significant (P> 0.05)
S.E = standard error of the difference of the accession means
C.D. = critical difference at 5% level of significance (P = 1.96) for the degree of freedom of error mean square (d.f. MSe>120).

Table 10. Combined estimated components of variances for anchote accessions based on morphological traits

Variable	Mean	Components of variances											h^2	predicted Genetic gain ($R=i.h^2.\delta p$)	R%	S.E	C.D.
		δ^2envt	δ^2rep(Envt)	δ^2Bl(Rp)	δ^2geno.	δ^2Geno*Envt	δ^2Error	δ^2pheno.	G.C.V	P.C.V	E.C.V	G.E.C.V					
VN	2.16	0.00	0.00	0.06	0.06	0.00	0.75	0.18	10.86	19.67	40.17	0.00	30.50	0.27	12.35	0.00	0.00
AVRWT	72.04	73.61	61.35	88.58	115.92	37.65	1187.90	326.45	14.95	25.08	47.84	8.52	35.50	13.22	18.35	8.58	16.82
YLD	17.70	5.10	3.00	6.21	7.83	3.19	54.85	18.03	15.81	23.99	41.84	10.08	43.40	3.80	21.46	2.26	4.43
DMC	19.39	0.71	1.24	0.00	1.69	0.00	26.36	6.08	6.70	12.72	26.48	0.00	27.70	1.41	7.26	0.84	1.65
LL	7.58	0.11	0.00	0.02	0.23	0.02	0.94	0.39	6.27	8.23	12.77	1.99	57.90	0.75	9.82	0.33	0.64
LD	7.63	0.02	0.00	0.02	0.28	0.02	0.99	0.45	6.95	8.82	13.05	1.81	62.10	0.86	11.29	0.15	0.30
PEL	10.03	0.05	0.01	0.09	0.65	0.00	3.36	1.21	8.04	10.96	18.27	0.00	53.70	1.22	12.13	0.23	0.45
INL	3.27	0.91	0.15	0.00	0.01	0.22	0.84	0.22	3.13	14.43	27.96	14.29	4.70	0.05	1.40	0.95	1.87
RL	9.61	2.29	0.36	0.49	0.73	0.04	2.87	1.22	8.88	11.49	17.63	1.97	59.80	1.36	14.14	1.52	2.97
RD	4.14	0.10	0.06	0.03	0.11	0.05	0.77	0.26	8.05	12.22	21.18	5.38	43.40	0.45	10.93	0.31	0.60

5.1.2.3. Correlation and Regression Coefficients

Simple correlation coefficients show strong relationships between some agro-morphological traits, while some traits have no relationship at all or are negatively related with one another. The effect of relationships of each trait was also evaluated by regression coefficients. The dry matter content has no significant association with any of the traits. Most traits showed significant relationship with yield trait (YLD) (Table 11).

Genotypic and phenotypic correlation coefficients estimated for each pair of traits revealed that most phenotypically related traits are also related genotypically, but not vice versa. IN and RL are the only traits with no genotypic relation to YLD (Table 12).

Table 11. Correlation coefficients with significance test (lower left including diagonal value = 1.00) and regression coefficient (upper left of the diagonal) test between agro-morphological traits in 149 anchote accessions grown at three different environments

Character	Character												
	VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINLR	RL	RD	RLDR
VN	1.00	**	**	ns	**	**	ns	*	**	**	*	**	ns
AVRWT	0.334**	1.00	**	ns	**	**	ns	ns	**	**	**	**	**
YLD	0.325**	0.995**	1.00	ns	ns	**	*	**	**	*	**	**	**
DMC	0.04 ns	-0.08ns	-0.08ns	1.00	ns	ns	ns	*	**	**	ns	*	ns
LL	0.217**	0.489**	0.491**	-0.03ns	1.00	**	**	**	**	**	**	**	**
LD	0.206**	0.506**	0.506**	-0.06ns	0.845**	1.00	**	**	**	ns	*	**	*
LLDR	0.00 ns	-0.05 ns	-0.05 ns	0.05ns	0.214**	-0.28**	1.00	**	**	**	**	ns	**
PEL	0.170**	0.416**	0.413**	-0.12**	0.524**	0.532**	-0.03ns	1.00	**	**	*	**	**
INL	0.140**	0.300**	0.299**	0.150**	0.364**	0.238**	0.222**	0.01ns	1.00	**	ns	**	ns
PINLR	-0.07 ns	-0.109*	-0.109*	-0.16**	-0.113**	0.01ns	-0.217**	0.394**	-0.889**	1.00	ns	ns	*
RL	0.143**	0.385**	0.374**	-0.01 ns	0.171**	0.104*	0.112**	0.095*	0.392**	-0.317**	1.00	*	**
RD	0.222**	0.663**	0.663**	-0.04 ns	0.450**	0.495**	-0.095*	0.304**	0.179**	-0.05ns	0.088*	1.00	**
RLDR	-0.05ns	-0.206**	-0.22**	0.04 ns	-0.199**	-0.28**	0.157**	-0.147**	0.168**	-0.206**	0.622**	-0.68**	1.00

VN = maximum branching number, AVRWT = average root weight, YLD = root yield, DMC = root dry matter content, LL = leaf length, LD = leaf diameter, LLDR leaf length to its diameter ratio, PEL, petiole length, INL = Internal node length, INLR, petiole to internodes lengths ratio, RL = root length, RD = root diameter, RLDR root length to its diameter ratio;** = highly significant (P < 0.01), * = significant (P < 0.05)

Table 12. Genotypic correlation coefficient (lower left of the diagonal) and phenotypic correlation coefficient (upper right of the diagonal).

Variable	VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINR	RL	RD	RLDR
VN	*1.00*	*0.31*	*0.30*	*-0.01*	*0.20*	*0.19*	*0.01*	*0.17*	*0.09*	*-0.06*	*0.11*	*0.21*	*-0.06*
		<.0001	*<.0001*	*0.86*	*<.0001*	*<.0001*	*0.90*	*<.0001*	*0.04*	*0.15*	*0.01*	*<.0001*	*0.15*
AVRWT	0.40	*1.00*	*1.00*	*-0.08*	*0.50*	*0.50*	*-0.01*	*0.40*	*0.28*	*-0.16*	*0.38*	*0.63*	*-0.19*
	<.0001		*<.0001*	*0.07*	*<.0001*	*<.0001*	*0.85*	*<.0001*	*<.0001*	*0.00*	*<.0001*	*<.0001*	*<.0001*
YLD	0.39	1.00	*1.00*	*-0.07*	*0.50*	*0.50*	*-0.01*	*0.39*	*0.28*	*-0.16*	*0.37*	*0.63*	*-0.20*
	<.0001	<.0001		*0.10*	*<.0001*	*0.09*	*0.90*	*<.0001*	*<.0001*	*0.00*	*<.0001*	*<.0001*	*<.0001*
DMC	0.08	-0.20	-0.19	*1.00*	*-0.06*	*-0.07*	*0.03*	*-0.11*	*0.09*	*-0.12*	*-0.02*	*-0.03*	*0.00*
	0.33	0.01	0.02		*0.20*	*<.0001*	*0.57*	*0.02*	*0.04*	*0.01*	*0.62*	*0.46*	*0.92*
LL	0.13	0.55	0.55	-0.20	*1.00*	*0.85*	*0.26*	*0.53*	*0.36*	*-0.14*	*0.17*	*0.43*	*-0.20*
	0.11	<.0001	<.0001	0.02		*<.0001*	*<.0001*	*<.0001*	*<.0001*	*0.00*	*<.0001*	*<.0001*	*<.0001*
LD	0.17	0.56	0.57	-0.21	0.89	*1.00*	*-0.28*	*0.54*	*0.22*	*-0.02*	*0.10*	*0.46*	*-0.27*
	0.04	<.0001	<.0001	0.01	<.0001		*<.0001*	*<.0001*	*<.0001*	*0.70*	*0.02*	*<.0001*	*<.0001*
LLDR	-0.13	-0.10	-0.10	0.02	0.12	-0.33	*1.00*	*-0.02*	*0.24*	*-0.23*	*0.13*	*-0.07*	*0.14*
	0.12	0.24	0.22	0.79	0.14	<.0001		*0.60*	*<.0001*	*<.0001*	*0.00*	*0.13*	*0.00*
PEL	0.08	0.43	0.42	-0.22	0.60	0.60	-0.08	*1.00*	*0.02*	*0.39*	*0.06*	*0.27*	*-0.15*
	0.33	<.0001	<.0001	0.01	<.0001	<.0001	0.37		*0.70*	*<.0001*	*0.15*	*<.0001*	*0.00*
INL	0.13	0.18	0.17	0.02	0.20	0.20	-0.03	-0.14	*1.00*	*-0.83*	*0.35*	*0.15*	*0.14*
	0.12	0.03	0.04	0.78	0.01	0.01	0.70	0.09		*<.0001*	*<.0001*	*0.00*	*0.00*
PINR	-0.12	-0.01	0.00	-0.09	0.09	0.10	-0.02	0.55	-0.82	*1.00*	*-0.34*	*-0.08*	*-0.19*
	0.15	0.93	0.96	0.28	0.26	0.22	0.77	<.0001	<.0001		*<.0001*	*0.06*	*<.0001*
RL	0.05	0.23	0.21	0.01	-0.02	-0.04	0.05	-0.14	0.40	-0.46	*1.00*	*0.09*	*0.66*
	0.57	0.01	0.01	0.90	0.78	0.61	0.59	0.10	<.0001	<.0001		*0.04*	*<.0001*
RD	0.27	0.61	0.61	-0.26	0.49	0.50	-0.09	0.32	0.01	0.12	-0.16	*1.00*	*-0.64*
	0.00	<.0001	<.0001	0.00	<.0001	<.0001	0.27	<.0001	0.89	0.14	0.05		*<.0001*
RLDR	-0.14	-0.26	-0.27	0.16	-0.32	-0.34	0.08	-0.30	0.25	-0.39	0.76	-0.72	*1.00*
	0.10	0.00	0.00	0.06	<.0001	<.0001	0.36	0.00	0.00	<.0001	<.0001	<.0001	

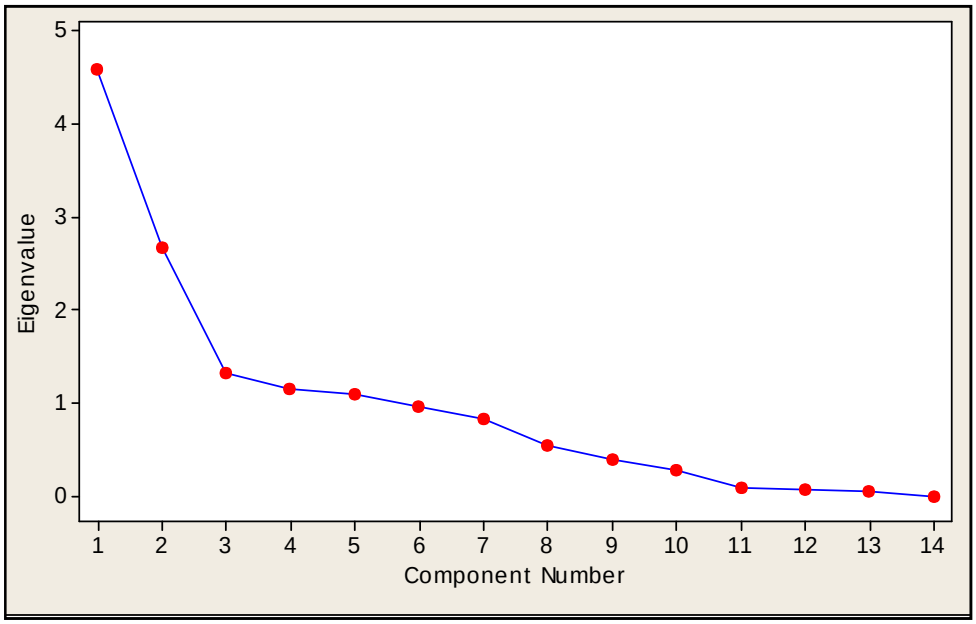
NB: Genotypic correlation coefficients are directly taken from the calculation of correlations between group (accession) and phenotypic correlation coefficients equates the correlation coefficient computed as a total sample correlations coefficient. $P < 0.001$ is significant correlation

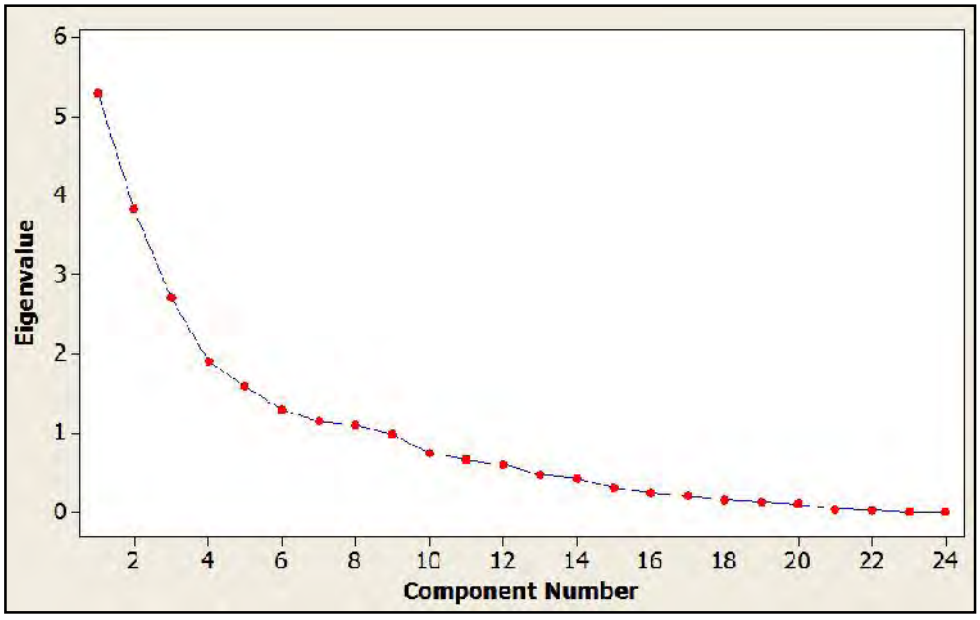
5.1.2.4. Principal Components

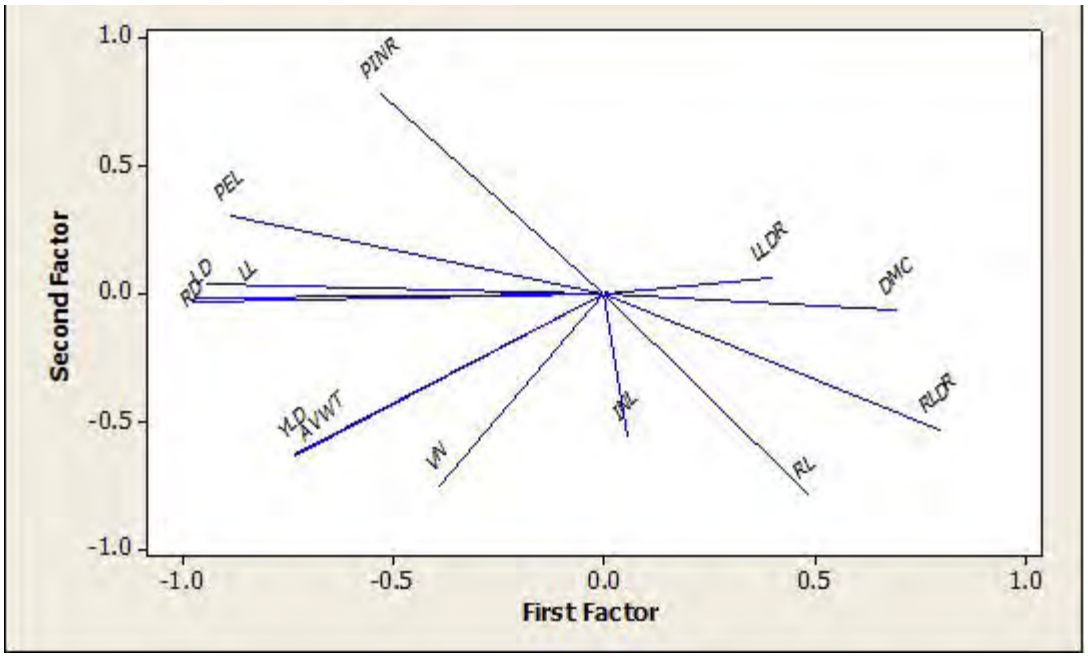
The principal components analyses (PCA) revealed the presence of a wide range of diversity that involves agro-morphological traits of anchote accessions.

Five PCs, with Eigen value of one and above (Figure 10 A), extracted from 13 agro-morphological quantitative traits, explained 77.2% of the total variations among 149 anchote accessions. The first PC alone accounted for more than 32% of the variation, followed by second (>19%) and the third (>9%) PCs. Accordingly, loading plots show the traits' relationships (correlations) (Figure 11). The first PC resulted mainly from variations in leaf length (LL), leaf diameter (LD), average root weight (AVRWT), yield (YLD), root diameter (RD), and petiole length (PEL), while the second PC is based on variations in inter-node length (INL), root length (RL) and the ratios (or shape of root (RLDR) and petiole length to inter-node ratio (PINLR) (Table 13).

For anchote grown from root, the phenotypic data (presence or absence), that were converted into binary matrix were also used for PC analysis. Nine PC_S (Figure 10 B), accounted for more than 80% of the total variations, where the first three PCs separately explain about 22%, 16%, and 11 % of the variation, respectively (Table 14). PC1 is mainly constituted by stem colors (STCLG & STCG), leaf green color intensities (LGCO1 & LGCO3), whereas PC2 is due to variation in stem color (STCDG), days to 50% emergence (DEL), and their negative counterpart vine pubescences (VP7 & VP3). PCA shows not only the effect of traits in discriminating accessions but also divided or grouped accessions into different categories, based on those discriminating traits. Accessions collected from Arjo, Ayira, Gechi and Gumay tend to cluster together according to their Woreda of origin (Figure 11 A), otherwise no clear grouping between


Eigenvalue
0
1
2
3
4
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Component Number


Eigenvalue
0
1
2
3
4
5
6
2
4
6
8
10
12
14
16
18
20
22
24
Component Number


Second Factor
First Factor
1.0
0.5
0.0
-0.5
-1.0
PINR
PEL
RLD
LL
LLDR
DMC
RLDR
YLD
AVWT
VW
DRL
RL

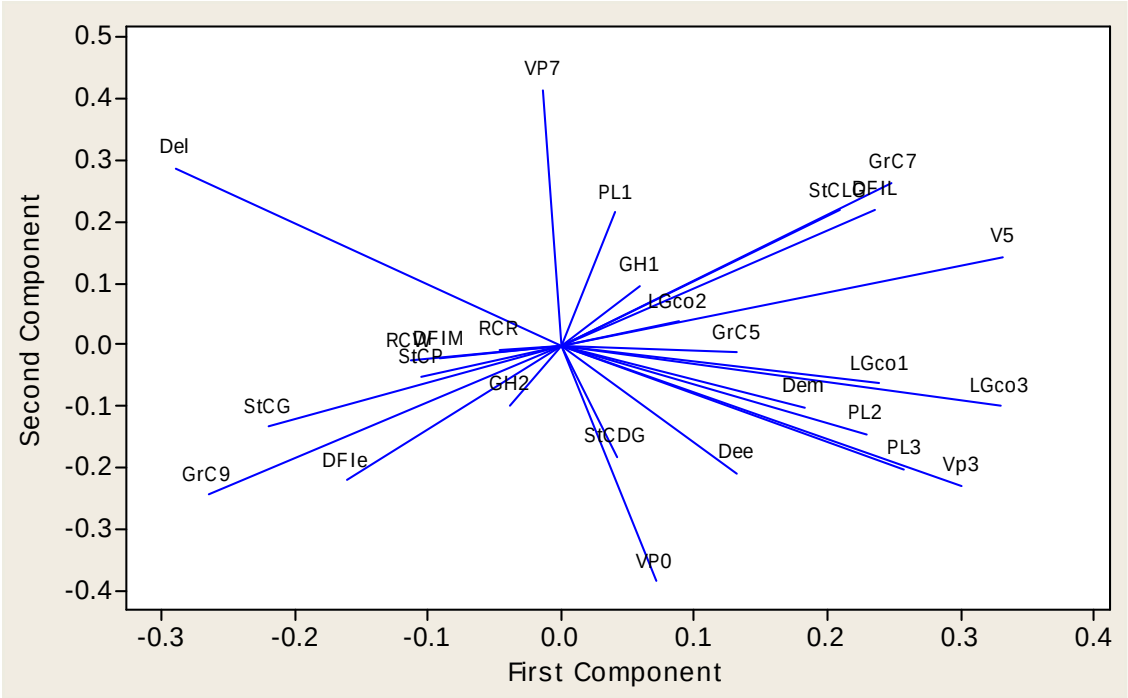

Second Component
First Component
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
VP7
Del
GrC7
PL1
StCLDFIL
V5
GH1
LGco2
RCR
GrC5
StCP
LGco1
GH2
Dem
LGco3
StCG
PL2
StCDG
PL3
Dee
GrC9
DFIe
Vp3
VP0

Table 13. Eigen vectors and eigen values of the first five Principal Components of 13 agro-morphological quantitative traits of 149 anchote accessions

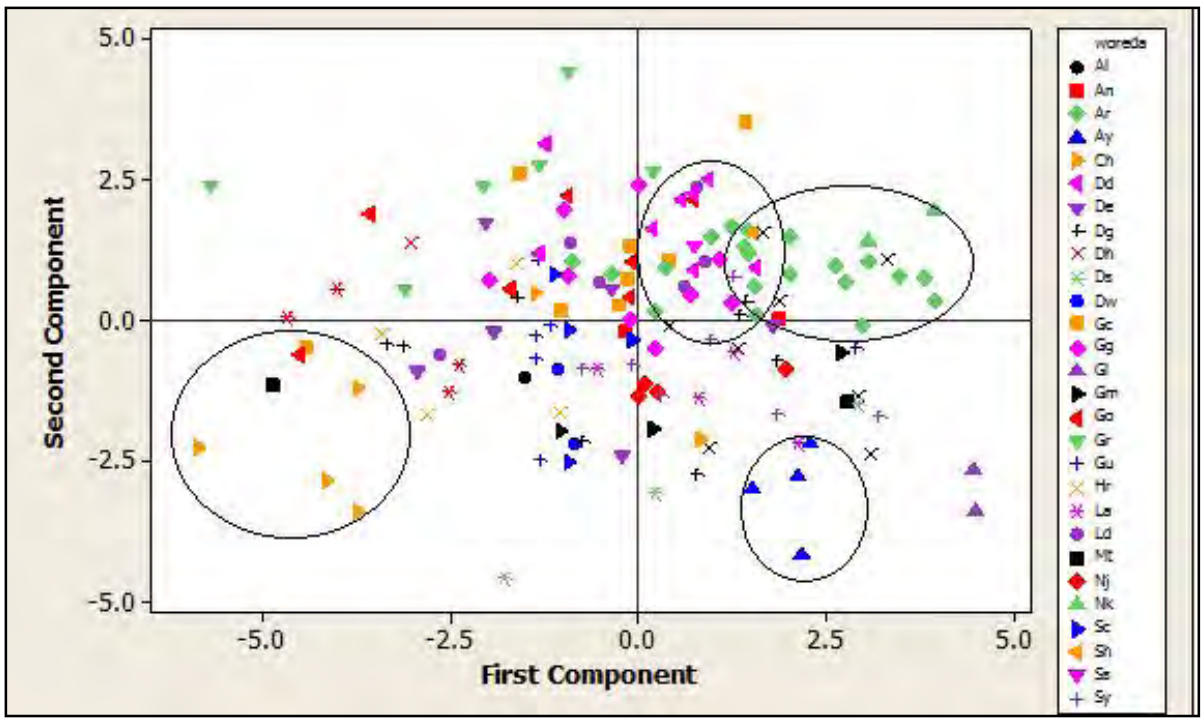
	Eigen vector				
Trait*	PC1	PC2	PC3	PC4	PC5
VN	-0.170	-0.190	0.538	-0.134	-0.060
AVRWT	-0.388	-0.207	0.211	0.119	-0.116
YLD	-0.378	-0.213	0.195	0.138	-0.104
DMC	0.135	-0.040	0.086	-0.616	-0.255
LL	-0.371	-0.045	-0.410	0.002	0.091
LD	-0.406	-0.048	-0.306	-0.067	0.005
LLDR	0.149	0.020	-0.211	0.208	0.198
PEL	-0.353	0.123	-0.297	0.002	-0.164
INL	-0.074	-0.406	-0.224	-0.274	0.497
PINLR	-0.119	0.482	-0.068	0.217	-0.400
RL	0.017	-0.510	-0.044	0.235	-0.361
RD	-0.355	0.125	0.306	-0.017	0.190
RLDR	0.235	-0.424	-0.204	0.185	-0.360
Eigen value	4.5784	2.6665	1.3239	1.1550	1.0861
% total variance	32.7	19.0	09.5	08.3	07.8
% cumulative variance	32.7	51.7	61.2	69.5	77.2

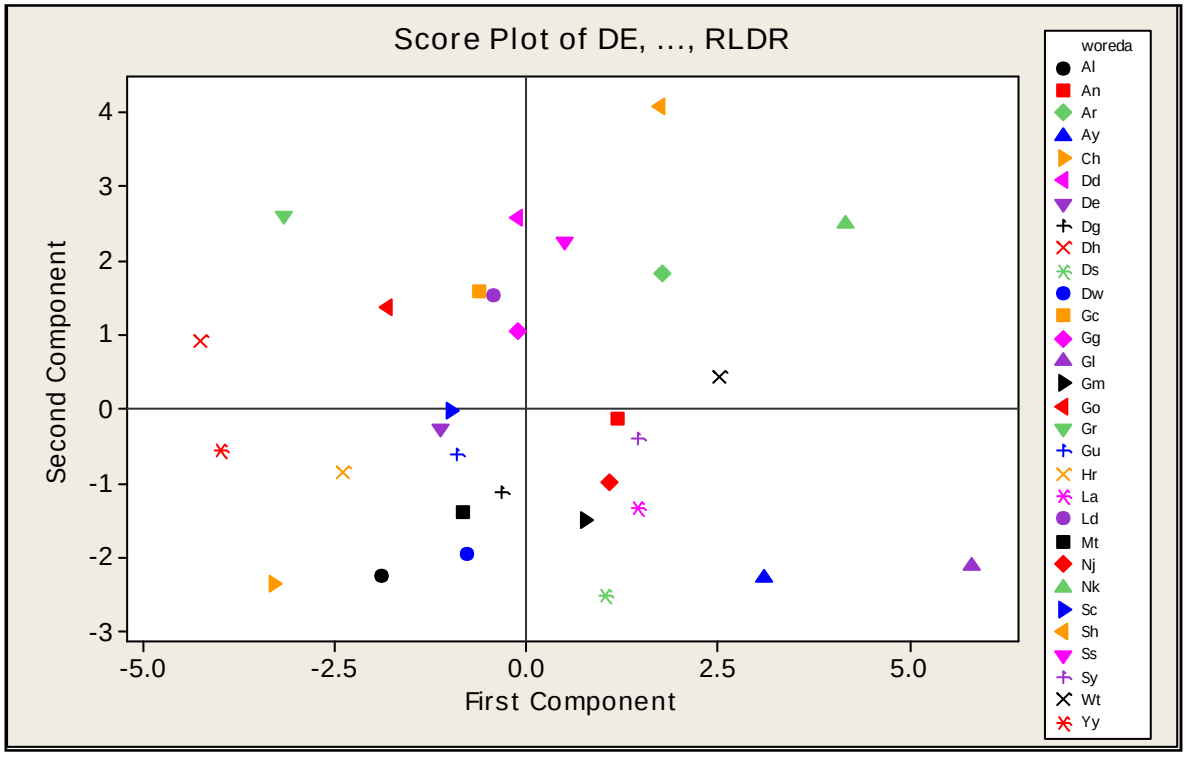
*VN= maximum branching number, AVRWT=average root weight, YLD= root yield, DMC= root dry matter content, LL=leaf length, LD=leaf diameter, LLDR leaf length to its diameter ratio, PEL, petiole length, INL= Internal node length, PINLR, petiole to internodes lengths ratio, RL= root length, RD=root diameter, RLDR root length to its diameter ratio

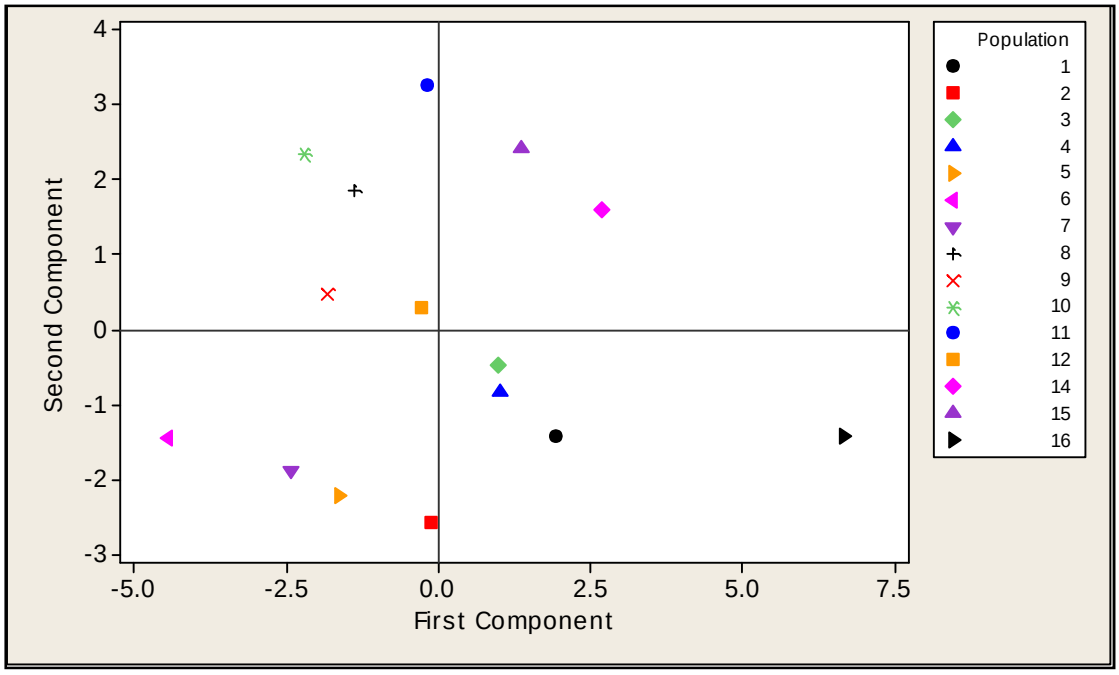
Table 14. Eigen vectors and eigen values of the first nine Principal Components of 24 phenotypic qualitative characters of 156 anchote accessions

Phenotype*	Eigen vector								
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
dee	0.132	-0.21	-0.007	-0.505	-0.059	-0.059	-0.06	-0.028	0.139
dem	0.183	-0.1	-0.021	0.561	0	0.041	-0.079	0.016	-0.123
del	-0.289	0.288	0.026	-0.031	0.055	0.019	0.128	0.012	-0.02
pl1	0.041	0.215	-0.023	-0.074	-0.025	0.044	-0.584	0.458	0.087
pl2	0.229	-0.15	-0.26	0.109	0.037	-0.135	0.059	-0.036	-0.292
pl3	0.258	-0.2	0.189	-0.108	0.006	0.18	0.018	-0.251	0.007
lgco1	0.239	-0.06	-0.212	-0.208	0.078	0.11	-0.082	0.106	0.189
lgco2	0.089	0.038	0.098	-0.267	0.163	0.168	0.08	0.381	-0.597
lgco3	0.33	-0.1	0.15	0.051	0.14	-0.059	-0.076	-0.05	-0.048
gh1	0.059	0.096	-0.189	-0.311	0.182	0.249	-0.187	-0.492	-0.134
gh2	-0.038	-0.1	0.061	0.041	-0.389	0.248	0.386	0.206	0.097
vp0	0.132	-0.01	0.125	-0.193	-0.133	-0.536	0.319	-0.052	-0.01
vp3	0.249	0.264	-0.199	-0.073	-0.004	-0.027	0.181	0.2	0.047
vp5	-0.265	-0.24	0.155	0.116	0.037	0.159	-0.248	-0.174	-0.041
vp7	0.071	-0.38	-0.008	-0.157	0.021	0.055	0.019	0.199	0.11
stclg	0.301	-0.23	0.084	0.183	0.018	-0.1	0.012	0.052	0.035
stcg	0.332	0.143	-0.06	0.139	0.037	-0.075	-0.179	-0.182	0.079
stcdg	-0.014	0.414	0.017	-0.008	0.142	-0.023	0.139	-0.134	-0.178
stcp	0.209	0.219	0.335	0.099	-0.068	0.018	-0.099	0.031	0.132
rcw	-0.22	-0.13	-0.068	0.105	0.302	0.034	0.024	0.2	0.174
rcr	0.042	-0.18	-0.101	0.071	0.247	0.444	0.272	0.042	-0.129
dfie	-0.105	-0.05	-0.337	0.085	-0.298	-0.103	-0.038	-0.173	-0.297
dfim	-0.113	-0.03	-0.336	0.016	0.264	-0.053	0.158	-0.085	0.434
dfil	-0.046	-0.01	0.064	-0.091	-0.575	0.276	-0.096	-0.123	0.032
Eigen value	5.882	4.299	3.075	1.959	1.609	1.417	1.231	1.21	1.121
% total variance	0.218	0.159	0.114	0.073	0.06	0.052	0.046	0.045	0.042
% cumulative	21.8	37.7	49.1	56.4	62.3	67.6	72.1	76.6	80.7

*The phenotype code is defined in Table 2.


5.0
2.5
0.0
-2.5
-5.0
Second Component
-5.0
-2.5
0.0
2.5
5.0
First Component
woreda
Al
An
Ar
Ay
Ch
Dd
De
Dg
Dh
Ds
Dw
Gc
Gg
Gl
Gm
Go
Gr
Gu
Hr
La
Ld
Mt
Nj
Nk
Sc
Sh
Se
Sy


Score Plot of DE, ..., RLDR
4
3
2
1
0
-1
-2
-3
Second Component
-5.0
-2.5
0.0
2.5
5.0
First Component
woreda
Al
An
Ar
Ay
Ch
Dd
De
Dg
Dh
Ds
Dw
Gc
Gg
Gl
Gm
Go
Gr
Gu
Hr
La
Ld
Mt
Nj
Nk
Sc
Sh
Ss
Sy
Wt
Yy


Second Component
First Component
4
3
2
1
0
-1
-2
-3
-5.0
-2.5
0.0
2.5
5.0
7.5
Population
1
2
3
4
5
6
7
8
9
10
11
12
14
15
16

Membership discriminate analysis, of the 149 accessions revealed that the first cluster consists of 92 accessions; the next top three (second to fourth) contained 15, 18 and 17 members, respectively. Accessions ANC 103 and ANC 178 failed to form cluster with any of the groups (Table 15).

Thirteen and seven Woredas were grouped into the first and second clusters, respectively, but five of the accessions did not cluster under either of the groups. The third and fourth clusters consisted of three and two members, respectively (Table 16).

Fifteen populations were clustered into five groups. Cluster 3, 5 and 2 contain five, four and three members, respectively, whereas the 4^{th} cluster has two members and the population Gl-Gd was not grouped with any other member (Table 17).

Table 15. Membership of 149 anchote accessions in eight clusters using standardized mean of 13 agro-morphological traits

Cluster	No. of accessions	Accessions (ANC. No.)
1	92	1, 2, 4, 17, 19, 21, 23, 26, 27, 28, 29, 33, 35, 53, 69, 70, 71, 73, 74, 76, 80, 81, 88, 90, 94, 95, 97, 99, 100, 101, 102, 104, 106, 108, 109,110, 111, 112, 113, 114, 115, 118, 120, 123, 124, 125, 126, 127, 128, 130, 131,133, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 147, 148, 149, 150, 151, 152, 157, 158, 159, 164, 168, 169, 170, 171, 172, 173, 174, 175, 177, 179, 180, 181, 182
2	15	5, 6, 7, 8, 9, 10, 11, 13, 24, 32, 41, 58, 91, 129, 162
3	18	12, 14, 31,36, 38, 52, 54, 75, 87, 89, 92, 107, 116, 117, 119, 161, 167, 176
4	17	22, 39, 50, 51, 55,56, 57, 59, 61, 62, 65, 67, 72, 77, 96, 98, 160
5	3	45, 60, 156
6	2	30, 68
7	1	178
8	1	103

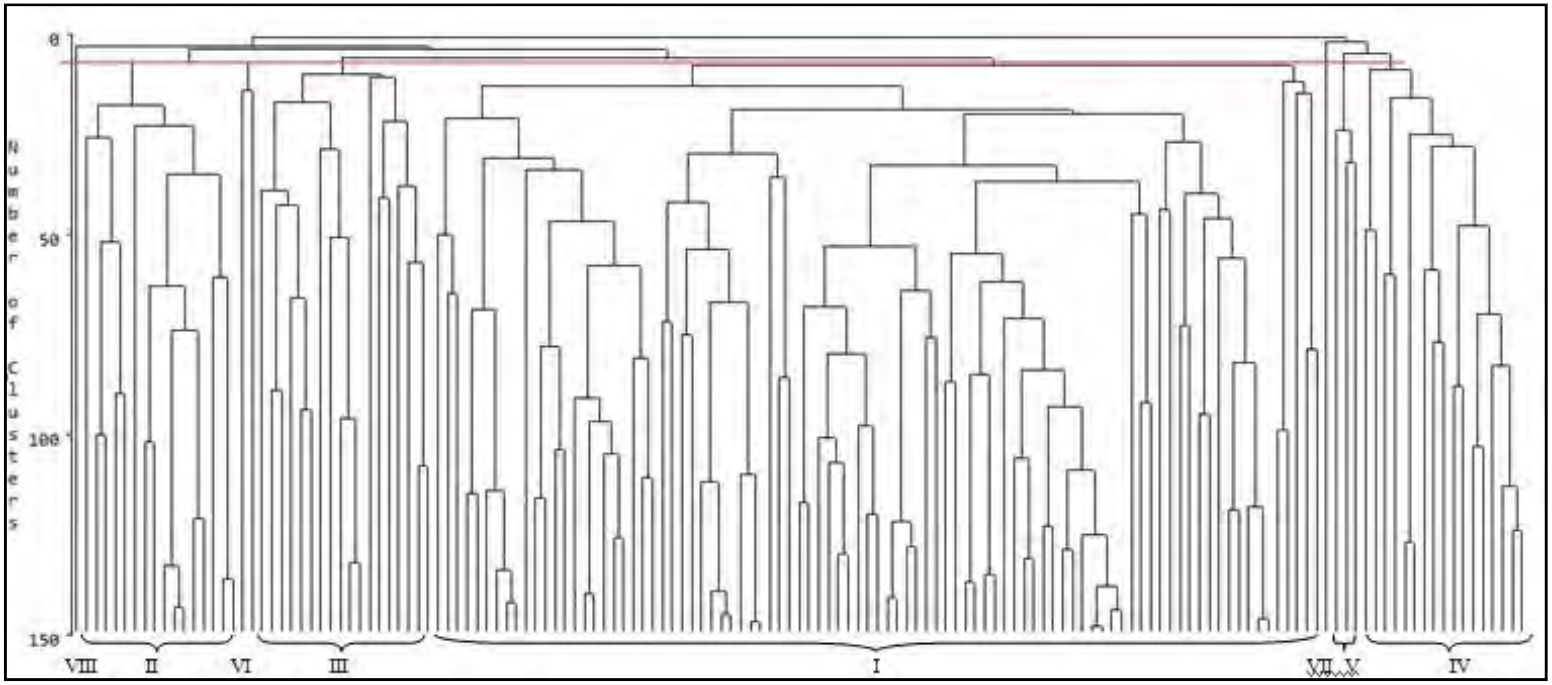
Table 16. Membership of 30 woredas (of anchote accessions' origin) in nine clusters

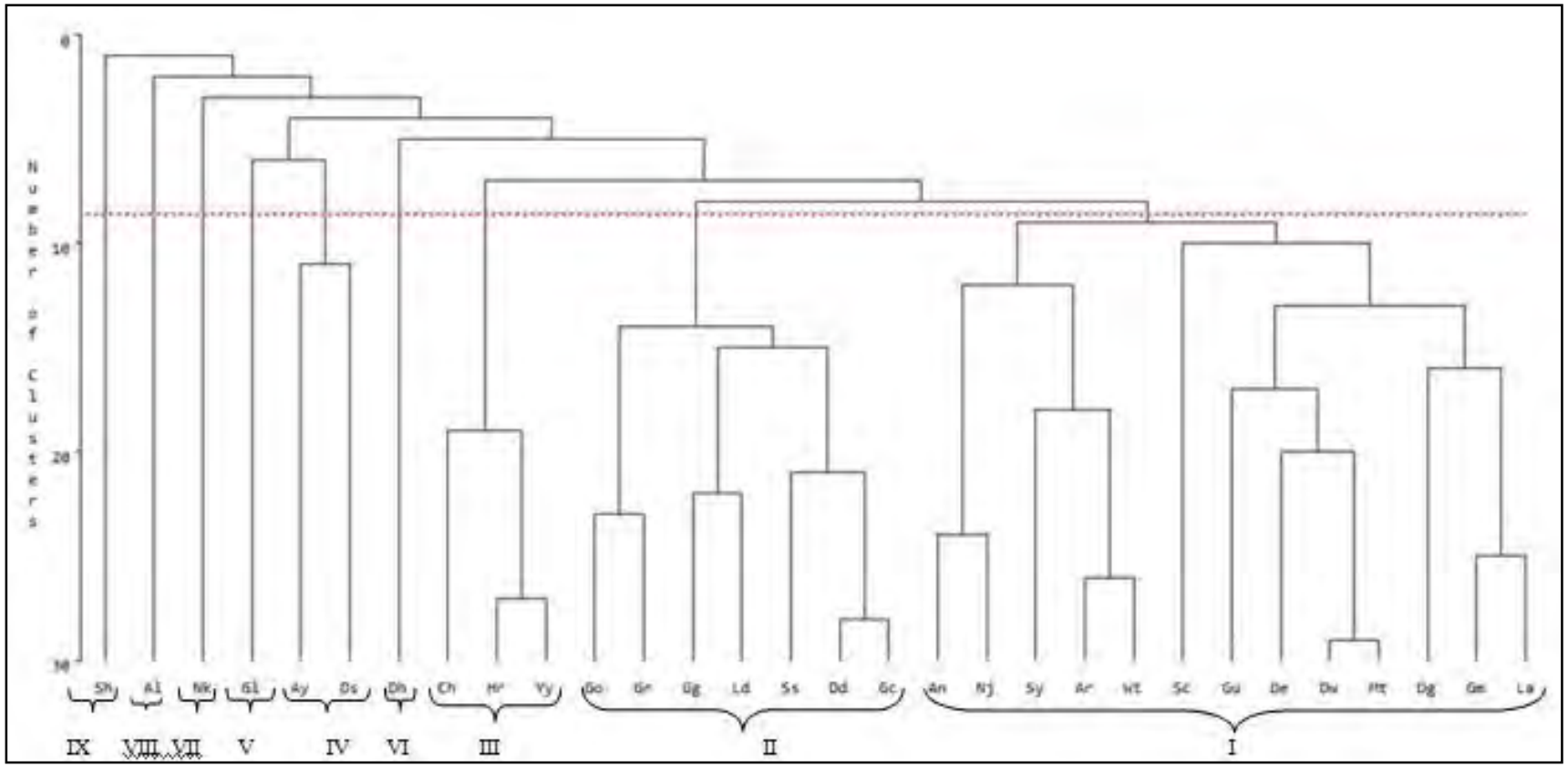
Cluster	Cluster								
	1	2	3	4	5	6	7	8	9
Woreda in a cluster	Dw, Mt, Ar, Wt, Gm, La, An, Nj, De, Sy, Gu, Dg, Ss	Dd, Gc, Go, Gr, Sc, Gg, Ld	Hr, Yy, Ch	Ay, Ds	Gl	Dh	Nk	Al	Sh

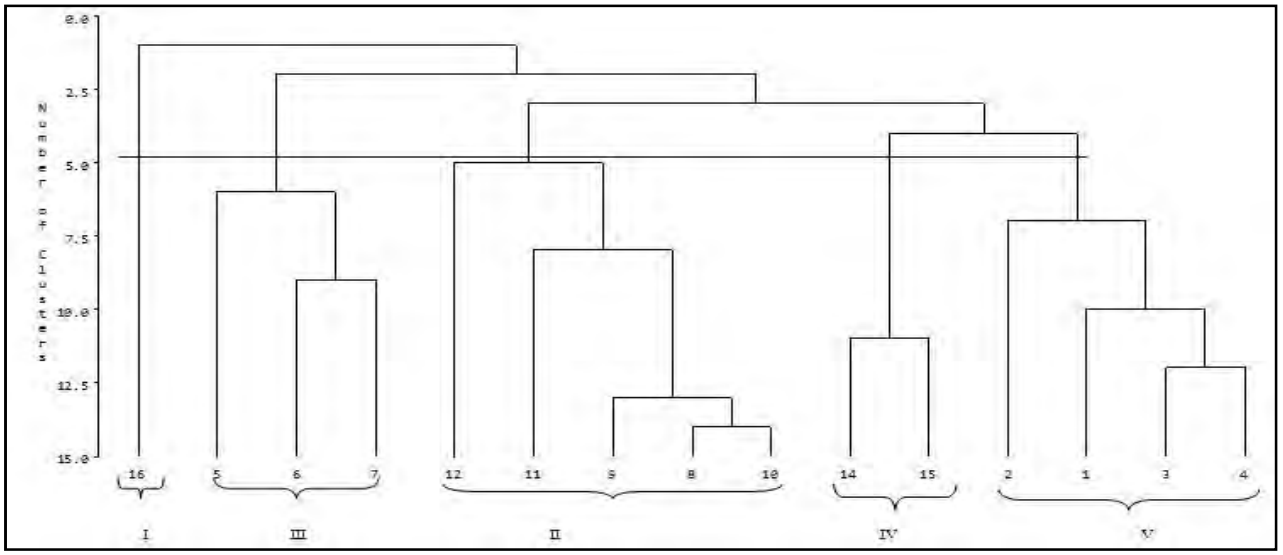
Table 17. Membership of 15 anchote populations in five clusters

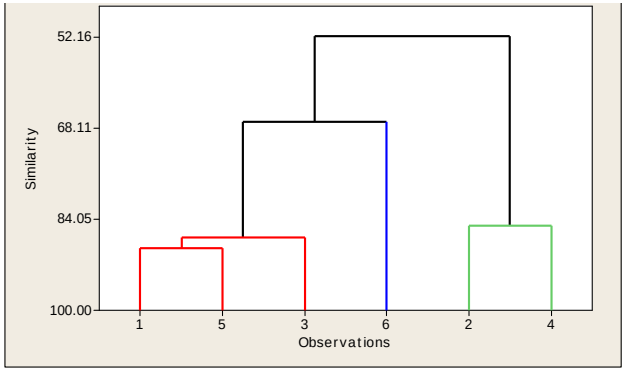
Cluster	1	2	3	4	5
Population in a cluster	16	5,6,7	8,9,10,11,12	14,15	1,2,3,4

Pop:1 = Ay-La; 2 = DS-DW; 3 = Sy-An;4 = Gm-Nj, 5 = Mt-Al; 6 = Hr-Yy; 7 = Ch-De; 8 = Gc-Dh; 9 = Gu,Go,&Mn; 10 = Gr,Sc,&Sh;11 = Dd-Dc; 12 = Gg-Dg; 14 = Ss,Wt,& Nk; 15 = Ar-Ld; 16 = Gl-Gd;


0
50
100
150
Number of Clusters
VIII
II
VI
III
I
VII
V
IV


Number of Clusters
0
10
20
30
Sh
Al
Nk
Gl
Ay
Ds
Dh
Ch
Hr
Yy
Go
Gr
Gg
Ld
Ss
Dd
Gc
An
Nj
Sy
Ar
Wt
Sc
Gu
De
Du
Ft
Dg
Gm
La
IX
VIII
VII
V
IV
VI
III
II
I


Distance of Clusters
0.0
2.5
5.0
7.5
10.0
12.5
15.0
16
5
6
7
12
11
9
8
10
14
15
2
1
3
4
I
III
II
IV
V


Similarity
52.16
68.11
84.05
100.00
1
5
3
6
2
4
Observations

5.1.2.6. Genetic Distance and Discriminant Analysis

The pair wise Mahalanobis (1936) distances (D^2), between clusters of accessions revealed that the largest distance obtained was between cluster 5 and 7 and the smallest is between cluster 3 and 4 (Table 18). In general, between Woredas and populations the Mahalanobis genetic distances (D^2) show greater increment as compared to accession clusters (Table 19 & 20).

Discriminant analysis of anchote accessions showed that 86 (57.71%) out of 149 accessions were correctly classified under their respective group made by cluster procedure of SAS software, while the rest accessions were scattered all over the groups. However, the percentage of accessions classified correctly varies greatly. The clusters whose members were 100% correctly clustered include clusters 5-8. The first and the second clusters accounted for about 57.61 % and 73.33% of their accession members which are correctly assigned, respectively, whereas the third and fourth have similar percent of correct members (44.44 and 41.18%, respectively) (Table 21).

Discriminant analysis using the Woredas (origin of the accessions) as grouping variable revealed that only 11 Woredas out of 30 (36.67%) were correctly classified in their respective Woreda. Only the third cluster was fully, and correctly, clustered but none of the accession was correctly clustered under clusters 5-9. About 40% of the accessions can be grouped under cluster two (Dd, Gc, Go, Gr, Gg, Ld, Sc), specially most of the members of cluster one (Table 22).

Population wise discriminant analysis also showed that there was no member that correctly clustered under clusters one and four (Table 23). Generally, discriminant analyses based on different groupings of anchote accessions showed that the correct assignment of members in each cluster decline when the groups inclusiveness increases.

The last row of each cluster (Table 21, 22, and 23) represents the percent of total members that should be included in their respective cluster. From these tables it is possible to understand the direction of gene flow between groups.

As discrimination function of traits revealed, each trait has different discrimination power for each cluster formed for different anchote groupings (Appendix Tables 10-12).

Table 18. Generalized Squared Distance D^2 between clusters based on the standardized mean values of 149 anchote accessions

Cluster	1	2	3	4	5	6	7	8
1	0.00							
2	3.63	0.00						
3	2.77	2.84	0.00					
4	3.28	3.28	1.51	0.00				
5	13.36	10.40	14.08	16.08	0.00			
6	8.72	8.71	5.41	5.62	25.86	0.00		
7	18.67	18.18	14.65	16.17	30.14	20.86	0.00	
8	6.31	5.70	8.41	8.83	22.58	14.05	24.43	0

Cluster 1, 2,…as indicated in table 16

Table 19. Generalized Squared Distance (D^2) between clusters constructed from standardized mean values of anchote accessions grouped by their origin (Woreda)

Cluster	Cluster								
	1	2	3	4	5	6	7	8	9
1	0.00								
2	16.27	0.00							
3	35.84	34.14	0.00						
4	55.00	87.60	66.34	0.00					
5	151.19	194.76	187.31	68.56	0.00				
6	88.96	75.48	132.91	169.94	267.61	0.00			
7	108.90	98.23	133.88	145.59	130.23	147.69	0.00		
8	559.79	569.66	634.67	514.96	380.70	571.96	452.87	0.00	
9	193.39	168.38	205.22	239.48	252.17	279.39	212.74	536.82	0.00

Table 20. Generalized Squared Distance (D^2) between clusters (population) constructed from standardized mean values of anchote accessions grouped into populations based on their close proximity (sharing common markets)

Cluster	1	2	3	4	5
1	0				
2	534,961,541	0			
3	362,873,215	26,865,793	0		
4	204,459,034	93,398,554	26,580,929	0	
5	277,057,362	63,275,906	11,192,734	17,178,324	0

Cluster 1 = pop 16; cluster 2 = pop 5, 6 & 7; cluster 3 = pop 8, 9, 10, 11, 12**;** cluster 4 = pop 14&15; cluster 5 = pop 1,2,3,4.

Table 21. Discriminate analysis for 149 anchote accessions showing their distribution over clusters

Cluster	Number of observations (accessions) represented to clusters								% of accessions correctly clustered (in the row cluster)
	1	2	3	4	5	6	7	8	
1	53	6	11	6	7	2	0	7	57.61
2	0	11	2	1	0	0	0	1	73.33
3	2	2	8	4	0	1	0	1	44.44
4	2	3	0	7	0	2	1	2	41.18
5	0	0	0	0	3	0	0	0	100
6	0	0	0	0	0	2	0	0	100
7	0	0	0	0	0	0	1	0	100
8	0	0	0	0	0	0	0	1	100
Total	57	22	21	18	10	7	2	12	86
% accn. out of the total that should be assigned	38.26	14.77	14.09	12.08	6.71	4.7	1.34	8.05	57.71

Table 22. Number of observations (Woreda) and percent classified into class

Cluster	Clusters into which each observations (accessions) are grouped									% of correctly clustered accessions
	1	2	3	4	5	6	7	8	9	
1	3	5	2	2	0	1	0	0	0	23.08
2	2	4	1	0	0	0	0	0	0	57.14
3	0	0	3	0	0	0	0	0	0	100
4	1	0	0	1	0	0	0	0	0	50
5	0	0	0	1	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0
7	0	1	0	0	0	0	0	0	0	0
8	0	0	0	0	1	0	0	0	0	0
9	0	1	0	0	0	0	0	0	0	0
Total	6	12	6	4	1	1	0	0	0	11
% woreda out of the total that should be assigned	20	40	20	13.33	3.33	3.33	0	0	0	36.67

Table 23. Number of observations (populations) and percent classified into class

Cluster	Clusters to which each observations (accessions) are grouped					% of correctly clustered accessions
	1	2	3	4	5	
1	0	0	0	1	0	0
2	0	1	2	0	0	33.33
3	0	0	2	2	1	40
4	0	0	1	0	1	0
5	0	0	2	1	1	25
Total	0	1	7	4	3	4
% popn. out of the total that should be assigned	0	6.6	46.7	26.7	20	26.67%

5.2. Polymorphism and Pattern of EST-SSR Genetic Diversity

5.2.1. EST-SSR Polymorphism

As the application of PCR is sensitive to quality DNA, gel testing and Nano-Drop measurements were done for each sample of gDNA extracted. Those DNA samples of low quality and quantity were extracted repeatedly until quality DNA was obtained otherwise did not included for PCR (Figure 14).

Among a total of 47 EST-SSR primer pairs, designed and applied on anchote, only 19 primers (40.43%) were able to amplify specific regions of the genome, but only 13 (27.66%) of them amplified consistently the target region (expected size), while six primer pairs (12.77%) amplified larger fragments (>1000 bp) (Figure 15).

The 13 selected and fluorescently labeled EST-SSRs primer pairs produced fragments (loci), containing microsatellite sequences of different repeat types and base compositions, but all are with perfect types of repeats. All of the markers, except WM-32, have thyamine (T) nucleotide (eight out of 13 contained either AT or TA repeats) among which 10 (77%) are di-nucleotide repeats and three (23%) are tri-nucleotides. However, only eight out of thirteen loci showed polymorphism, while the rest four were monomorphic and one was multi-sized bands per sample, the latter five loci had been excluded from further genetic analyses (Table 25). For the bulked sample, only seven loci were polymorphic (Table 26 B).

Overall, from 146 individually sampled plants, 24 alleles and from 33 samples of bulked leaves, 20 alleles were recorded, respectively, across eight loci. Four alleles, including 179 (D) (of locus WM-5), 231(C) (of locus WM- 30), 323(A) (of locus WM-29), and 234(A) (of locus WM- 25) appeared in “individually collected sample data” but missing in bulked data (Table 26 A&B).

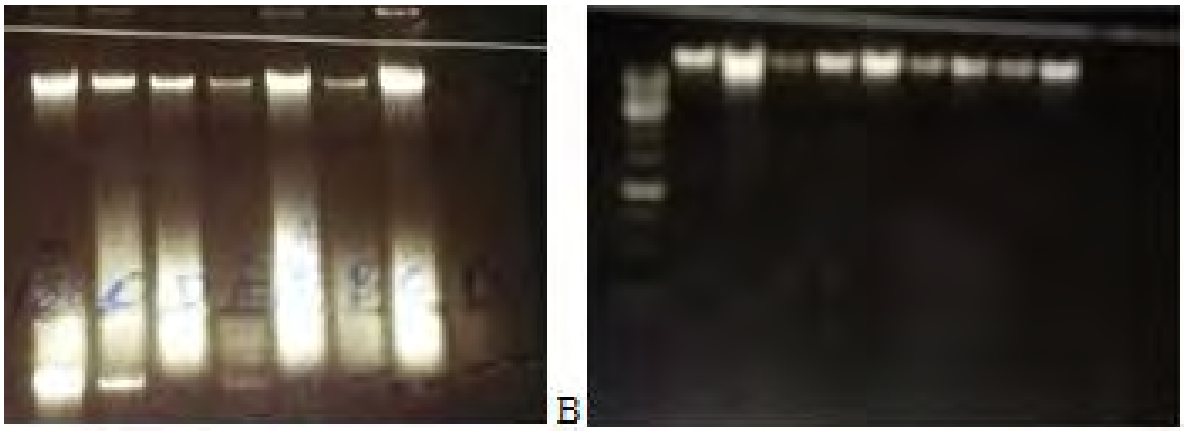
The number of observed alleles per locus (Na) ranges from two to six with an average of three per locus, whereas the effective number of allele (Ne) ranges from 1.06 to 4.8 with an average of 1.93. Alleles that scored highest and lowest weighted overall allelic frequency are WM-25 "C" (0.9296) and WM-5 "D" (0.007), respectively. Largest polymorphic information content (PIC) was recorded by locus WM-24 (0.76), while the least by WM-29 (0.062) with an average of 0.3142. The highest and lowest Shannon's information indices (I) were scored for WM-24 (1.64) and WM-29 (0.15) loci, respectively with overall average value of 0.633 (Table 27 A&B).

Appendix Table 9A & B show the frequency distribution of each allele across populations varying from 0.0 to 1.0, among which nine (41.67%) are rare alleles with overall weighted allele frequency $\leq$ 0.10. One allele is private allele (179 of locus WM-5), occurring only in one accession (ANC10). In other ways, there are many alleles of a locus with frequency $\geq$ 0.95 per population that may not be considered as polymorphic locus.

All the rare alleles in "individual sample" were also occurring rarely in "bulked samples", except allele 227 (of locus WM-30). Since some of the alleles appearing in one data set are missing in another, null allele may occur. Therefore, an estimation of null allele frequency per locus per population was done and the result obtained show that it ranges from 0.0 to 0.95 (Table 28).

Locus WM-24 showed highest relative allelic richness (highly informative) while locus WM-29 was the poorest.

Populations including Arjo-Leka Dulecha, Gimbi-Nejo, and Abay Chomen-Bako Tibe are the three highly ranked populations in allelic richness. The least allelic richness was observed in


B

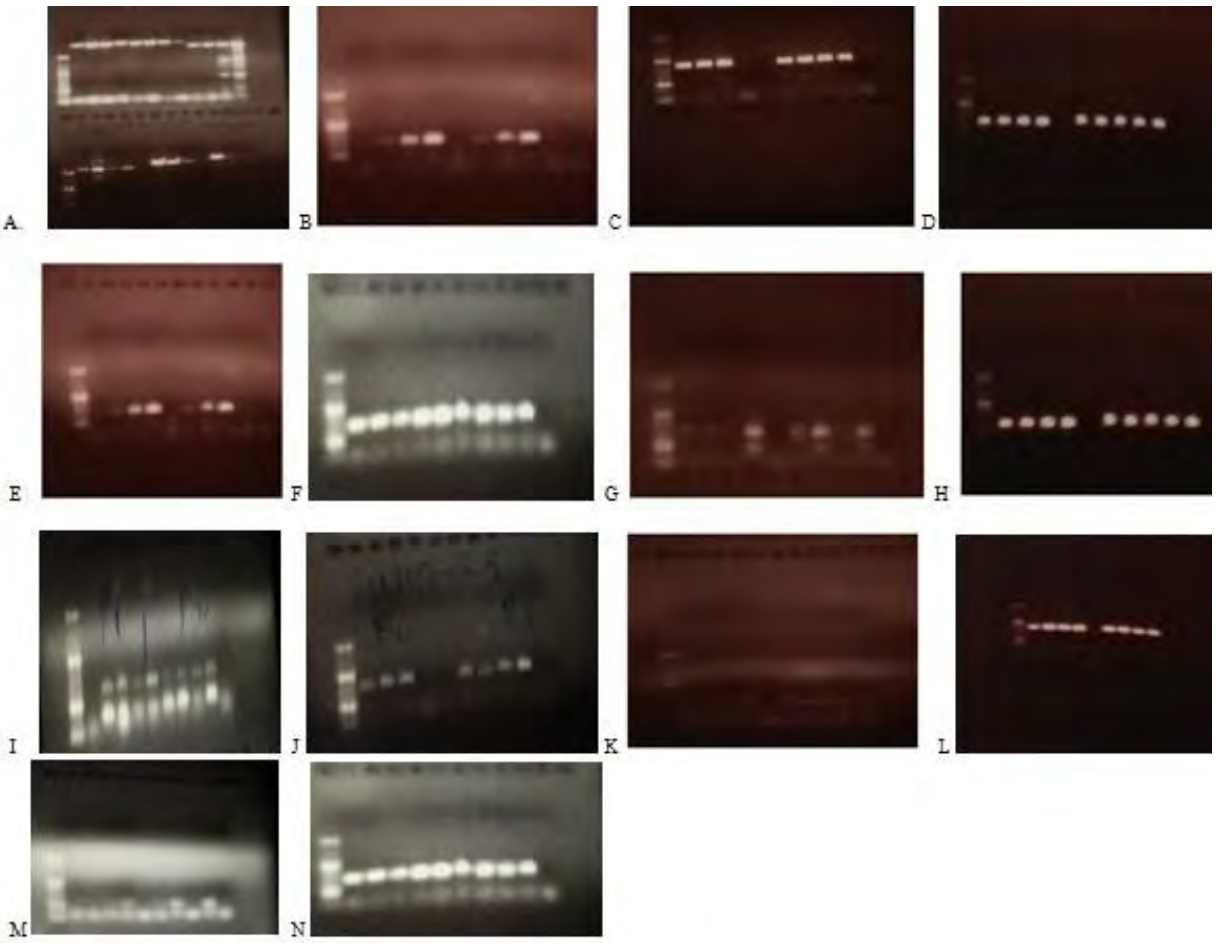

A
B
C
D
E
F
G
H
I
J
K
L
M
N

Table 24. Functional 13 EST-SSR loci of anchote with their respective PCR allele products designed

S. No.	Locus	Expected Allele Size (bp)	Allele sizes observed in anchote	Microsatellite repeat	Max. No. of Allele	Tagged gene
1	WM-5	189	170, 173, 176 & 179	(CTT)8	4	*Citrullus lanatus* sub.spp *Vulgaris* cDNA clone WMLS377 5'
2	WM-8	390	395 & 398	(ATC)6	2	*C. lanatus* cDNA clone SSH0003795
3	WM-11	372	397	(AT)9	1	*C. lanatus* cDNA clone SSH0002986
4	WM-12	284	291	(AT)9	1	*C. lanatus* cDNA clone SSH0002980
5	WM-14	297	316	(AT)8	1	*C. lanatus* cDNA clone SSH0002819
6	WM-15	283	Multisized	(TA)7	11	*C. lanatus* cDNA clone SSH0002809
7	WM-24	240	258, 260, 262, 264, 266 & 268	(TC)6	6	*C. lanatus* cDNA clone SSH0002342
8	WM-25	242	234, 243 & 246	(CCT)7	3	*C. lanatus* cDNA clone SSH0002329
9	WM-29	300	323 &339	(TA)7	2	*C. lanatus* cDNA clone SSH0002159
10	WM-30	236	227, 229 & 231	(TA)11	3	*C. lanatus* cDNA clone SSH0001982
11	WM-32	400	402 & 404	(AG)14	2	*C. lanatus* cDNA clone SSH0001794
12	WM-34	245	225 & 229	(CT)14	2	*C. lanatus* cDNA clone SSH0001734
13	WM-46	300	322 &324	(AT)8	1	*C. lanatus* cDNA clone SSH0000715

Table 25. A Summary of overall allele frequency per locus for 146 individual samples (A) and for 33 bulked samples (B)

A

Allele/Locus	WM-5	WM-34	WM-32	WM-30	WM-8	WM-24	WM-29	WM-25
Allele A	0.0106	0.508	0.2797	0.1	0.2207	0.0479	0.0331	0.0176
Allele B	0.6092	0.492	0.7203	0.8793	0.7793	0.2397	0.9669	0.0528
Allele C	0.3732	-	-	0.0207	-	0.2295	-	0.9296
Allele D	0.007	-	-	-	-	0.1747	-	-
Allele E	-	-	-	-	-	0.2534	-	-
Allele F	-	-	-	-	-	0.0548	-	-

B

Allele/locus	WM-5	WM-34	WM-32	WM-30	WM-8	WM-24	WM-29	WM-25
A	0.0303	0.4808	0.3824	0.2143	0.25	0.0857	-	-
B	0.6667	0.5192	0.6176	0.7857	0.75	0.3286	1	0.0735
C	0.303	-	-	-	-	0.1714	-	0.9265
D	-	-	-	-	-	0.1714	-	-
E	-	-	-	-	-	0.2	-	-
F	-	-	-	-	-	0.0429	-	-

Letter A, B, C… indicate the alleles for each loci in their respective order of allele size mentioned in table 25

Table 26. Summary of Na, Ne, I, and PIC for all loci (A) and for 15 populations (B)

A

Locus	Na*	Ne*	I*	PIC
WM-5	4	1.9587	0.7528	0.3859
WM-34	2	1.9995	0.693	0.3749
WM-32	2	1.6749	0.5927	0.3218
WM-30	3	1.2761	0.4236	0.2003
WM-8	2	1.5243	0.5278	0.2848
WM-24	6	4.7586	1.6376	0.7568
WM-29	2	1.0684	0.1453	0.0619
WM-25	3	1.1531	0.2943	0.1274
Mean	3	1.9267	0.6334	0.3142
St. Dev	1.414	1.1958	0.453	

B

Population	Diversity index	WM-5	WM-34	WM-32	WM-30	WM-8	WM-24	WM-29	WM-25	Average
Ay-La	N_a	4	2	1	2	2	4	2	1	2.250
	N_e	2.60	1.97	1.00	1.11	1.11	2.99	1.92	1.00	1.711
	I	1.09	0.69	0.00	0.20	0.20	1.19	0.67	0.00	0.505
Ds-Dw	N_a	2	2	2	3	2	5	2	1	2.375
	N_e	1.98	1.98	1.98	1.50	1.22	3.85	1.69	1.00	1.899
	I	0.69	0.69	0.69	0.61	0.33	1.45	0.60	0.00	0.631
Sy-An	N_a	2	2	2	2	2	5	1	1	2.125
	N_e	1.80	1.96	1.25	1.67	1.25	4.38	1.00	1.00	1.788
	I	0.64	0.68	0.35	0.59	0.35	1.53	0.00	0.00	0.518
Gm-Nj	N_a	2	2	2	3	2	5	1	3	2.500
	N_e	1.98	1.98	1.92	1.50	1.22	3.77	1.00	1.80	1.897
	I	0.69	0.69	0.67	0.61	0.33	1.43	0.00	0.75	0.645
Mt-Al	N_a	2	2	2	2	2	5	1	2	2.250
	N_e	1.34	2.00	1.60	1.34	1.25	4.26	1.00	1.11	1.736
	I	0.42	0.69	0.56	0.42	0.35	1.50	0.00	0.20	0.519
Hr-Yy	N_a	3	2	2	3	2	4	1	1	2.250

	N_e	1.68	2.00	1.98	1.50	1.11	2.60	1.00	1.00	1.608
	I	0.73	0.69	0.69	0.61	0.20	1.14	0.00	0.00	0.507
Ch-De	N_a	2	2	2	1	2	6	1	2	2.250
	N_e	2.00	1.98	1.12	1.00	1.22	4.08	1.00	1.12	1.689
	I	0.69	0.69	0.22	0.00	0.33	1.58	0.00	0.22	0.465
Gc-Dh	N_a	2	2	2	2	1	2	1	2	1.750
	N_e	1.12	1.97	1.34	1.11	1.00	1.92	1.00	1.47	1.366
	I	0.22	0.69	0.42	0.20	0.00	0.67	0.00	0.50	0.337
Gu-Go	N_a	2	2	2	2	2	3	1	1	1.875
	N_e	1.91	1.96	1.39	1.67	1.91	2.79	1.00	1.00	1.702
	I	0.67	0.68	0.45	0.59	0.67	1.06	0.00	0.00	0.515
Gr-Sh	N_a	2	2	2	2	2	4	1	1	2.000
	N_e	1.92	1.98	1.72	1.25	1.84	3.28	1.00	1.00	1.748
	I	0.67	0.69	0.61	0.35	0.65	1.24	0.00	0.00	0.526
Dd-Dc	N_a	2	2	2	1	2	4	1	1	1.875
	N_e	1.80	2.00	1.91	1.00	1.84	2.04	1.00	1.00	1.573
	I	0.64	0.69	0.67	0.00	0.65	0.93	0.00	0.00	0.446
Gg-Dg	N_a	2	2	2	1	2	4	1	3	2.125
	N_e	1.98	1.96	1.84	1.00	2.00	3.64	1.00	1.59	1.875
	I	0.69	0.68	0.65	0.00	0.69	1.34	0.00	0.68	0.591
Ac-Bt	N_a	2	2	2	2	2	5	1	3	2.375
	N_e	1.88	1.95	1.80	1.39	1.98	3.95	1.00	1.29	1.904
	I	0.66	0.68	0.64	0.45	0.69	1.46	0.00	0.46	0.630
Ss-Wt	N_a	2	2	2	1	2	3	1	1	1.750
	N_e	2.00	1.98	1.91	1.00	1.91	3.00	1.00	1.00	1.723
	I	0.69	0.69	0.67	0.00	0.67	1.10	0.00	0.00	0.477
Ar-Ld	N_a	2	2	2	3	2	5	2	3	2.625
	N_e	2.00	1.97	1.92	1.36	1.22	4.44	1.11	1.23	1.906
	I	0.69	0.69	0.67	0.52	0.33	1.54	0.20	0.39	0.629

* N_a = Observed number of alleles * N_e = Effective number of alleles [Kimura and Crow (1964), $Ne = 1/(\sum pi^2)$] * I = Shannon's Information index [Lewontin (1972)]

Table 27. Locus by population estimated null allele frequencies

	Ay-La	Ds-Dw	Sy-An	Gm-Nj	Mt-Al	Hr-Yy	Ch-De	Gc-Dh	Gu-Go	Gr-Sh	Dd-Dc	Gg-Dg	Ac-Bt	Ss-Wt	Ar-Ld
WM-5	0.24	0.00	0.00	0.13	0.16	0.12	0.20	0.00	0.10	0.00	0.00	0.00	0.15	0.17	0.07
WM-34	0.35	0.33	0.00	0.33	0.00	0.00	0.32	0.35	0.38	0.00	0.47	0.38	0.00	0.00	0.00
WM-32	NI	0.45	0.88	0.63	0.77	0.47	0.94	0.89	0.82	0.63	0.67	0.63	0.75	0.58	0.63
WM-30	0.00	0.00	0.05	0.00	0.16	0.00	NI	0.00	0.05	0.21	NI	NI	0.00	NI	0.00
WM-8	0.95	0.89	0.88	0.89	0.88	0.95	0.89	NI	0.33	0.71	0.71	0.32	0.58	0.58	0.89
WM-24	0.00	0.35	0.00	0.11	0.06	0.17	0.21	0.00	0.00	0.00	0.13	0.14	0.04	0.00	0.00
WM-29	0.77	0.85	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	0.95
WM-25	NI	NI	NI	0.78	0.95	NI	0.94	0.84	NI	NI	NI	0.77	0.87	NI	0.90

NI = No information for the locus for particular population (with <2 alleles)

Table 28. Allelic richness per locus and population

	Population															
Locus	Ay-La	Ds-Dw	Sy-An	Gm-Nj	Mt-Al	Hr-Yy	Ch-De	Gc-Dh	Gu-Go	Gr-Sh	Dd-Dc	Gg-Dg	Ac-Bt	Ss-Wt	Ar-Ld	Average
WM-5	3.26	2.00	2.00	2.00	1.90	2.66	2.00	1.56	2.00	2.00	2.00	2.00	2.00	2.0	2.00	2.16
WM-34	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.0	2.00	2.00
WM-32	1.00	2.00	1.82	2.00	1.98	2.00	1.56	1.90	1.93	2.00	2.00	2.00	2.00	2.0	2.00	1.97
WM-30	1.50	2.40	1.99	2.40	1.90	2.40	1.00	1.50	1.99	1.82	1.00	1.00	1.93	1.0	2.26	1.85
WM-8	1.50	1.76	1.82	1.76	1.82	1.50	1.76	1.00	2.00	2.00	2.00	2.00	2.00	2.0	1.76	1.92
WM-24	3.45	4.28	4.60	4.21	4.39	3.51	4.88	2.00	2.98	3.49	2.98	3.84	4.34	3.0	4.59	4.50
WM-29	2.00	2.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.0	1.50	1.29
WM-25	1.00	1.00	1.00	2.48	1.50	1.00	1.56	1.96	1.00	1.00	1.00	2.63	2.25	1.0	2.00	1.59
Total	15.71	17.44	16.23	17.85	16.49	16.07	15.76	12.92	14.90	15.31	13.98	16.47	17.52	14.0	18.11	17.28

Table 29. Hardy-Weinberg equilibrium test (Chi-square test at ά level) for each locus per population and for overall population

S. No	Population	Locus								For overall population	NPL (%PL)
		WM-5	WM-34	WM-32	WM-30	WM-8	WM-24	WM-29	WM-25		
1	Ay-La	ns	ns	Mn	ns	ns	ns	*	Mn	ns	6 (75)
2	Ds-Dw	ns	*	Ns	Ns	ns	ns	**	Mn	ns	7 (87.5)
3	Sy-An	ns	*	Ns	Ns	ns	ns	Mn	Mn	ns	6 (75)
4	Gm-Nj	ns	*	Ns	Ns	ns	*	Mn	ns	ns	7 (87.5)
5	Mt-Al	ns	**	Ns	Ns	ns	**	Mn	ns	**	7 (87.5)
6	Hr-Yy	*	**	Ns	ns	ns	*	Mn	Mn	*	6 (75)
7	Ch-De	ns	**	ns	Mn	ns	*	Mn	ns	*	6 (75)
8	Gc-Dh	ns	*	Ns	Ns	Mn	ns	Mn	ns	ns	6 (75)
9	Gu-Go	ns	*	Ns	Ns	ns	ns	Mn	Mn	ns	6 (75)
10	Gr-Sh	ns	*	Ns	**	ns	ns	Mn	Mn	ns	6(75)
11	Dd-Dc	ns	ns	Ns	Mn	ns	ns	Mn	Mn	ns	5 (62.5)
12	Gg-Dg	ns	ns	Ns	Mn	ns	ns	Mn	ns	ns	6 (75)
13	Ac-Bt	ns	ns	Ns	Ns	ns	ns	Mn	ns	ns	7(87.5)
14	Ss-Wt	ns	*	Ns	Mn	ns	ns	Mn	Mn	ns	5 (62.5)
15	Ar-Ld	ns	*	Ns	Ns	ns	ns	ns	***	ns	8(100)
Overall mean		ns	***	Ns	Ns	ns	*	***	ns	**	6.267(78.33)

Key: ns = not significant, * P<0.05, ** P<0.01, *** P<0.001, NPL = number of polymorphic loci, %PL = percentage of polymorphic loci, Mn = a locus considered as monomorphic or not considered as polymorphic (not used for p value calculation)

5.2.2. Genetic Diversity

Gene diversity estimated across the populations by each locus showed the highest diversity was contributed, for majority of the populations, by locus WM-24 and the least by locus WM-29 (Table 31 and Table 32). Genetic diversity scores over eight loci also showed that the least observed and estimated Nei's heterozygosity by loci were recorded for locus WM-29 (0.01 and 0.06, respectively) and the highest observed heterozygosity (0.84) for locus WM-34 and Nei's heterozygosity (0.79) for locus WM-24. The average heterozygosity across all the loci was 0.35 (Table 32).

Using two loci, WM-29 and WM-25, it was difficult to estimate genetic diversity and determine null allele frequency determination for most populations, because of the absence of at least two alleles in some populations. For such populations, the genetic diversity per locus became zero and the null allele frequencies were not determined (Table 28 and 32). Three populations (Ac-Bt, Ds-Dw, and Gm-Nj) have highest allelic richness and shared highest genetic diversity (0.42). Population with low allelic richness and as a consequence the least unweighted average of overall genetic diversity (0.23) was Gc-Dh (Table 29 & 31).

5.2.3. Differentiation and Gene Flow among Populations

The highest estimated population differentiation (F_{ST}) was recorded by locus WM-29 whereas the least by locus WM-34. The average population differentiation (F_{ST}) was 0.11, which shows weak population sub-structuring. Negative within population and total genetic differentiations (F_{IS}, F_{IT}) were also observed in some loci. Apparent over all gene flow per locus is also evident from Table 32 &33, at locus WM-34 with an average of 2.06, i.e., more than two individuals migrate to either population per generation. Higher gene flow between each pair of population,

the inverse of differentiation, is the indication that anchote population has wider random mating system (Table 33). Conversely, assuming that the heterozygosity deficit is caused entirely by selfing, the proportion of progeny produced by only selfing ($S = 2F_{IS}/1\text{-}F_{IS}$) is highest for locus WM-29 and least for WM-34, respectively (Table 33). Table 33 shows low-to- moderate pair wise population differentiation (F_{ST}) and higher gene flows (N_m), where the highest differentiation was within moderate range ($F_{ST} = 0.132$).

The AMOVA of anchote populations indicated that the highest variation (83.75%) was observed within individuals, followed by variance among individuals within population (8.35), whereas among populations variance was very low 7.87%. Then the proportion of variance in each hierarchical class was tested by permuting individual genotypes. Within accession, the scored variation was the least (Table 34). Fixation index (F_{IS}) presented in Table 35 indicated that locus WM-34, is heterozygote excess for its observed alleles while others possess deficient or both alleles. The negative F_{IS} indicates heterozygote excess of allele, otherwise it is deficient.

The Nei's original measures of genetic identity (I) show that most populations have more than 90% identity, i.e., minimal genetic distances (D) were observed between populations. The lowest genetic diversity (0.0074) was recorded between Gr-Sh and Ac-Bt populations, whereas the largest distance was observed between Ay-La and Gu-Go (0.145) populations (Table 36), and in the latter case there is limited gene flow (Nm = 1.68).

Table 30. Gene diversity (H_e) per locus and population using eight EST-SSR loci for 15 anchote populations

	Population														
Locus	Ay-La	Ds-Dw	Sy-An	Gm-Nj	Mt-Al	Hr-Yy	Ch-De	Gc-Dh	Gu-Go	Gr-Sh	Dd-Dc	Gg-Dg	Ac-Bt	Ss-Wt	Ar-Ld
WM-5	0.66	0.52	0.47	0.53	0.28	0.43	0.54	0.11	0.51	0.50	0.47	0.52	0.52	0.55	0.53
WM-34	0.52	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.53	0.52	0.50	0.50	0.50
WM-32	0.00	0.51	0.21	0.51	0.40	0.51	0.11	0.28	0.29	0.43	0.51	0.48	0.49	0.50	0.51
WM-30	0.10	0.35	0.43	0.35	0.28	0.35	0.00	0.10	0.43	0.22	0.00	0.00	0.29	0.00	0.28
WM-8	0.10	0.19	0.21	0.19	0.21	0.10	0.19	0.00	0.50	0.49	0.49	0.51	0.53	0.50	0.19
WM-24	0.69	0.78	0.81	0.77	0.80	0.66	0.81	0.50	0.67	0.72	0.54	0.78	0.80	0.69	0.82
WM-29	0.60	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
WM-25	0.00	0.00	0.00	0.48	0.10	0.00	0.11	0.34	0.00	0.00	0.00	0.40	0.24	0.00	0.20
UW	0.33	0.42	0.33	0.42	0.32	0.32	0.28	0.23	0.36	0.36	0.32	0.40	0.42	0.34	0.39

UW = unweighted average

Table 31. Overall Nei's estimation of heterozygosities, differentiation measures, Gene flow and proportion of progenies produced by selfing

	Genetic diversity parameter												
Locus	Ho	He	Ave.Het	Hs	Ht'	D_{ST}	F_{IS}	F_{IT}	F_{ST}	G_{IT}	G_{ST}	Nm	S
WM-5	0.35	0.49	0.45	0.478	0.49	0.01	0.23**	0.30**	0.09	0.28	0.03	2.60	0.85
WM-34	0.84	0.50	0.49	0.504	0.50	0.00	-0.69	-0.67	0.01	-0.65	-0.01	19.71	-0.80
WM-32	0.36	0.40	0.36	0.383	0.41	0.02	0.00	0.10	0.10	0.05	0.05	2.23	0.22
WM-30	0.19	0.22	0.20	0.212	0.22	0.01	0.06	0.15*	0.09	0.12	0.04	2.61	0.34
WM-8	0.29	0.35	0.28	0.292	0.35	0.06	-0.06	0.16*	0.20**	0.00	0.16	0.98	0.37
WM-24	0.71	0.79	0.69	0.723	0.80	0.07	-0.04	0.09*	0.13**	0.01	0.09	1.65	0.21
WM-29	0.01	0.06	0.07	0.074	0.10	0.02	0.90**	0.93**	0.30**	0.91	0.22	0.59	25.97
WM-25	0.08	0.13	0.12	0.125	0.13	0.01	0.26**	0.35**	0.12	0.31	0.06	1.82	1.07
Ave.	0.35	0.37	0.33	0.349	0.37	0.02	-0.08	0.04	0.11	-0.01	0.06	2.06	0.08

Ho = observed heterozygosity; He = expected unbiased Nie's heterozygosity; Hs = gene diversity within populations; Ht' = total gene diversity
D_{ST} = Nei's (1978) unbiased average gene diversity between subpopulations, G_{ST} = analogous to Wright (1951) F_{ST}
Nm = Gene flow estimated from $F_{ST} = 0.25(1- F_{ST})/F_{ST}$ (Nei, 1987); $S = 2*F_{IS}/ (1-F_{IS})$

Table 32. Pair wise population differentiation (F_{ST}) (below diagonal) and gene flow (Nm) (above diagonal) for 15 populations (146 individual samples) (A) and for five populations (bulk leaves sampled) (B)

A

Population	Ay-La	Ds-Dw	Sy-An	Gm-Nj	Mt-Al	Hr-Yy	Ch-De	Gc-Dh	Gu-Go	Gr-Sh	Dd-Dc	Gg-Dg	Ac-Bt	Ss-Wt	Ar-Ld
Ay-La	0.00	***3.233***	***3.122***	***2.437***	***3.200***	***1.946***	***4.742***	***2.308***	***1.648***	***2.279***	***1.721***	***2.024***	***2.222***	***2.088***	***3.196***
Ds-Dw	0.072	0.00	***3.265***	***4.524***	***3.780***	***4.473***	***3.281***	***1.826***	***2.157***	***4.004***	***3.382***	***2.874***	***3.586***	***3.722***	***8.845***
Sy-An	0.074	0.071	0.00	4.646	8.862	5.276	5.946	2.963	3.954	6.486	3.098	2.665	5.387	2.873	7.760
Gm-Nj	0.093	0.052	0.051	0.00	7.176	4.718	4.392	3.365	2.794	4.870	3.222	5.021	6.413	3.718	12.930
Mt-Al	0.072	0.062	0.027	0.034	0.00	***7.909***	***5.102***	***5.701***	***3.396***	***6.268***	***3.967***	***3.124***	***6.116***	***3.725***	***7.454***
Hr-Yy	0.114	0.053	0.045	0.050	0.031	0.00	***2.881***	***2.949***	***2.283***	***5.453***	***5.262***	***2.368***	***4.263***	***3.013***	***8.813***
Ch-De	0.050	0.071	0.040	0.054	0.047	0.080	0.00	***2.816***	***2.515***	***4.158***	***2.936***	***3.325***	***3.468***	***3.778***	***5.908***
Gc-Dh	0.098	**0.120**	0.078	0.069	0.042	0.078	0.082	0.00	***1.788***	***2.493***	***1.836***	***1.841***	***2.562***	***1.748***	***3.194***
Gu-Go	**0.132**	0.104	0.059	0.082	0.069	0.099	0.090	0.123	0.00	***8.639***	***3.262***	***4.150***	***9.959***	***3.408***	***3.628***
Gr-Sh	0.099	0.059	0.037	0.049	0.038	0.044	0.057	0.091	0.028	0.00	***9.849***	***6.835***	***22.164***	***9.797***	***10.190***
Dd-Dc	**0.127**	0.069	0.075	0.072	0.059	0.045	0.078	**0.120**	0.071	0.025	0.00	***4.800***	***6.241***	***7.330***	***6.047***
Ac-Bt	0.110	0.080	0.086	0.047	0.074	0.096	0.070	**0.120**	0.057	0.035	0.050	0.00	***10.740***	***8.735***	***5.742***
Ac-Bt	0.101	0.065	0.044	0.038	0.039	0.055	0.067	0.089	0.024	**0.011**	0.039	0.023	0.00	***8.040***	***9.047***
Ss-Wt	0.107	0.063	0.080	0.063	0.063	0.077	0.062	0.125	0.068	0.025	0.033	0.028	0.030	0.00	***6.605***
Ar-Ld	0.073	0.027	0.031	0.019	0.032	0.028	0.041	0.073	0.064	0.024	0.040	0.042	0.027	0.036	0.00

B

Sub-regional population	WEST WOLLEGA	KELLAM WOLLEGA	EAST COLLEGE	ILLUABABORA	JIMMA
WEST WOLLEGA	0				
KELLAM WOLLEGA	0.07	0			
EAST WOLLEGA	0.02	0.07	0		
ILLUABABORA	0.04	0.04	0	0	
JIMMA	0.07	0.14	0.02	0.02	0

Table 33. Analysis of Molecular Variance (AMOVA) for weighted average over eight loci for 15 populations (above) and 30 accessions (146 individual samples) (below)

Source	D.F	M.S	Variance Component	Percentage Variance	P-value
Among Populations	14	41.51	1789.36	7.87	0.0098
Among Individuals within population	131	153.88	19243.17	8.35	0.000
Within Individuals	146	143	18.81	83.75	0.000
Total	**291**	**338.39**	**21051.34**	**100**	
Among accessions	29	79.29	0.178	15.25	0.001
Among Individuals within accession	116	116.10	0.01	0.92	0.006
Within Individuals	146	143.00	0.98	83.84	0.001
Total	**291**	**338.39**	**1.17**	**100**	

Table 34. Wright's (1978) fixation index (F_{IS}) as a measure of heterozygote deficiency or excess

Allele\locus	WM-5	WM-34	WM-32	WM-30	WM-8	WM-24	WM-29	WM-25
Allele A	0.6631	-0.6804	0.0976	0.1954	0.1579	0.2497	0.8851	-0.0179
Allele B	0.2753	-0.6804	0.0976	0.1227	0.1579	0.2108	0.8851	0.3665
Allele C	0.2775	-	-	-0.0211	-	0.0121	-	0.4621
Allele D	1	-	-	-	-	0.0735	-	-
Allele E	-	-	-	-	-	0.095	-	-
Allele F	-	-	-	-	-	-0.058	-	-
Total	0.295	-0.6804	0.0976	0.1395	0.1579	0.0981	0.8851	0.3636

Table 35. Nei's Original Measures of Genetic Identity (I) (above diagonal) and Genetic distance (Ds) (below diagonal).

population	Ay-La	Ds-Dw	Sy-An	Gm-Nj	Mt-Al	Hr-Yy	Ch-De	Gc-Dh	Gu-Go	Gr-Sh	Dd-Dc	Gg-Dg	Ac-Bt	Ss-Wt	Ar-Ld
Ay-La	****	0.9177	0.9313	0.9105	0.9338	0.876	0.9572	0.9073	0.8648	0.9063	0.8684	0.8903	0.8975	0.9063	0.9204
Ds-Dw	0.0859	****	0.9208	0.9372	0.9277	0.9446	0.9303	0.8685	0.8721	0.9377	0.937	0.9064	0.9192	0.9435	0.9658
Sy-An	0.0712	0.0825	****	0.9466	0.9688	0.9527	0.9674	0.9343	0.9332	0.9661	0.9359	0.918	0.9563	0.928	0.9701
Gm-Nj	0.0937	0.0648	0.0549	****	0.9676	0.9378	0.9565	0.9381	0.9039	0.9467	0.9209	0.95	0.9518	0.9379	0.973
Mt-Al	0.0684	0.075	0.0317	0.033	****	0.9574	0.9595	0.9635	0.9179	0.9548	0.9324	0.9251	0.9553	0.9366	0.956
Hr-Yy	0.1324	0.057	0.0485	0.0642	0.0435	****	0.9133	0.9218	0.8971	0.9606	0.9708	0.8957	0.9405	0.9282	0.968
Ch-De	0.0438	0.0722	0.0331	0.0445	0.0413	0.0907	****	0.9175	0.903	0.9449	0.9219	0.9408	0.9374	0.9479	0.9604
Gc-Dh	0.0972	0.141	0.0679	0.0639	0.0371	0.0815	0.0861	****	0.8924	0.9281	0.8819	0.8805	0.9249	0.8841	0.9229
Gu-Go	0.1453	0.1369	0.0691	0.101	0.0857	0.1086	0.102	0.1138	****	0.9713	0.9321	0.9479	0.9759	0.9314	0.9168
Gr-Sh	0.0984	0.0643	0.0345	0.0548	0.0463	0.0402	0.0566	0.0746	0.0291	****	0.9779	0.9672	0.9926	0.9749	0.9784
Dd-Dc	0.1411	0.0651	0.0662	0.0824	0.0699	0.0297	0.0813	0.1256	0.0703	0.0223	****	0.9459	0.962	0.9619	0.9607
Ac-Bt	0.1162	0.0983	0.0855	0.0513	0.0778	0.1101	0.0611	0.1272	0.0535	0.0333	0.0556	****	0.9792	0.9736	0.9512
Ac-Bt	0.1081	0.0843	0.0447	0.0494	0.0458	0.0613	0.0647	0.078	0.0244	0.0074	0.0388	0.021	****	0.9753	0.9673
Ss-Wt	0.0984	0.0582	0.0747	0.0641	0.0656	0.0745	0.0536	0.1232	0.071	0.0255	0.0388	0.0268	0.025	****	0.964
Ar-Ld	0.083	0.0348	0.0303	0.0273	0.045	0.0326	0.0404	0.0802	0.0869	0.0218	0.0401	0.05	0.0333	0.0366	****

5.2.4. Accessions Clustering

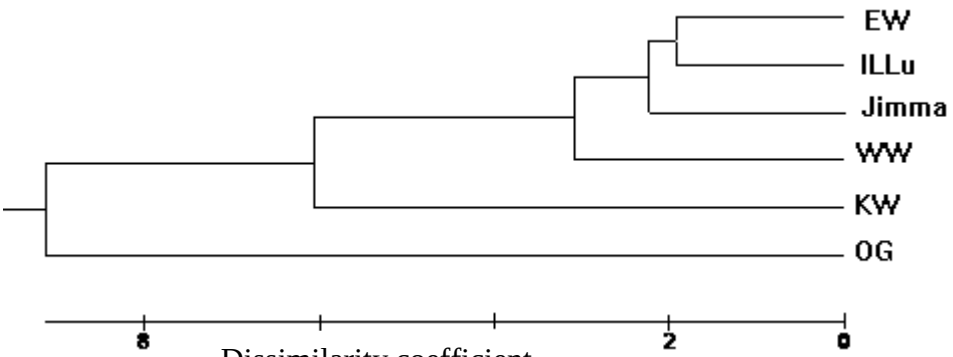
Dendrograms constructed from both bulked samples and separate plant samples showed that the clusters did not follow clear pattern of geographic origins and in all cases, the out-group was a separate cluster. In bulked sample, Kellam Wollega (KW) was separated as a single group while the rest grouped together (Figure 17).

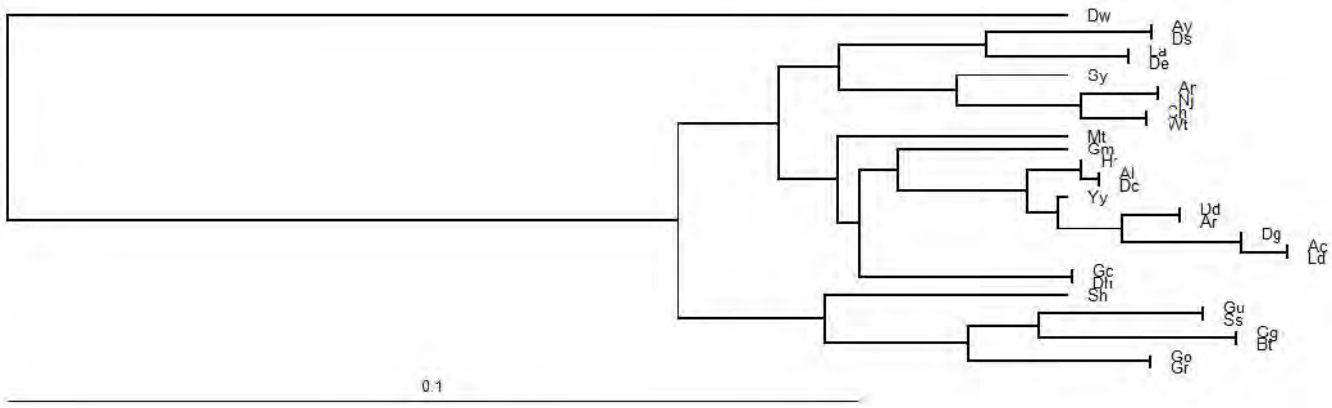
All the thirty woredas were clustered under three sub-groups and exceptionally Dalle Wobera (Dw) formed separate cluster. No two accessions grouped as a single population on the proximity of their woredas of origin, except Gechi (Gc) with Dhidhesa (Dh), Gomma (Go) with Gera (Gr) and Ay with Dalle Sedi (Ds) populations (Figure 18).

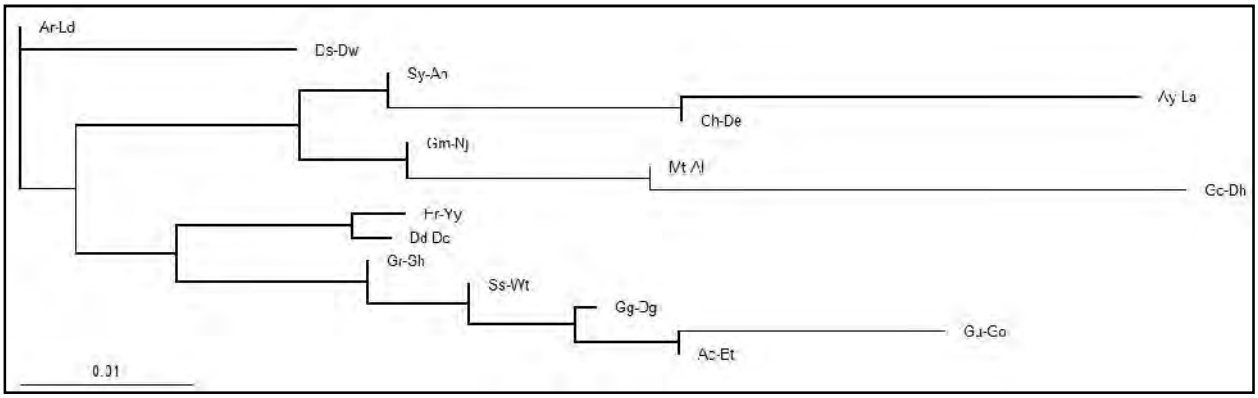
The same pattern of admixture clusters were obtained from population clustering, in which two populations, Arjo-Leka Dulecha (Ar-Ld) and Dalle Sadi-Dalle Wobera (Ds-Dw), formed separate groups (Figure 19).

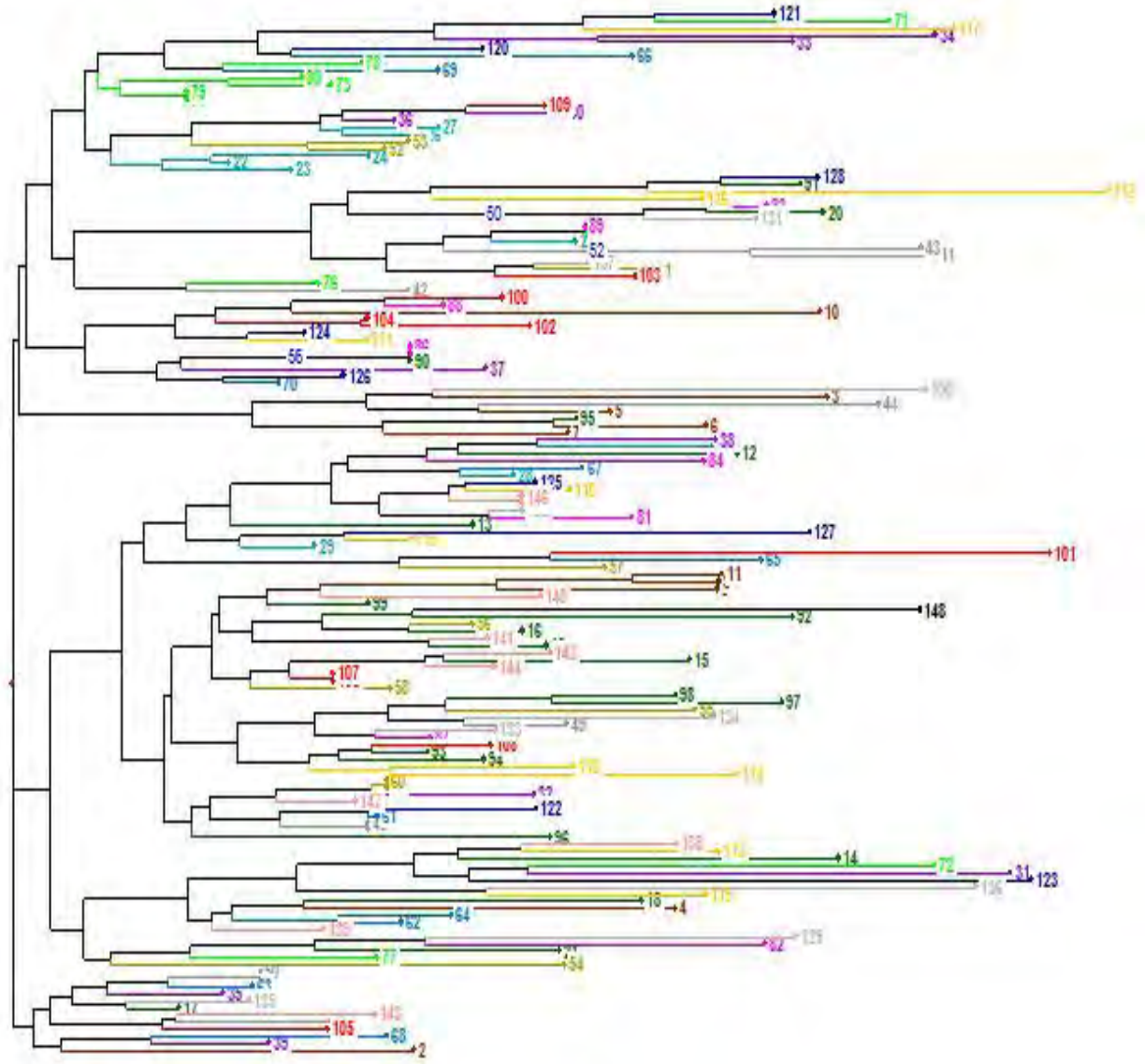
Neighbor joining tree formed from 146 anchote accessions (samples) showed that there are three major groups, which did not follow the geographic pattern (Figure 20).

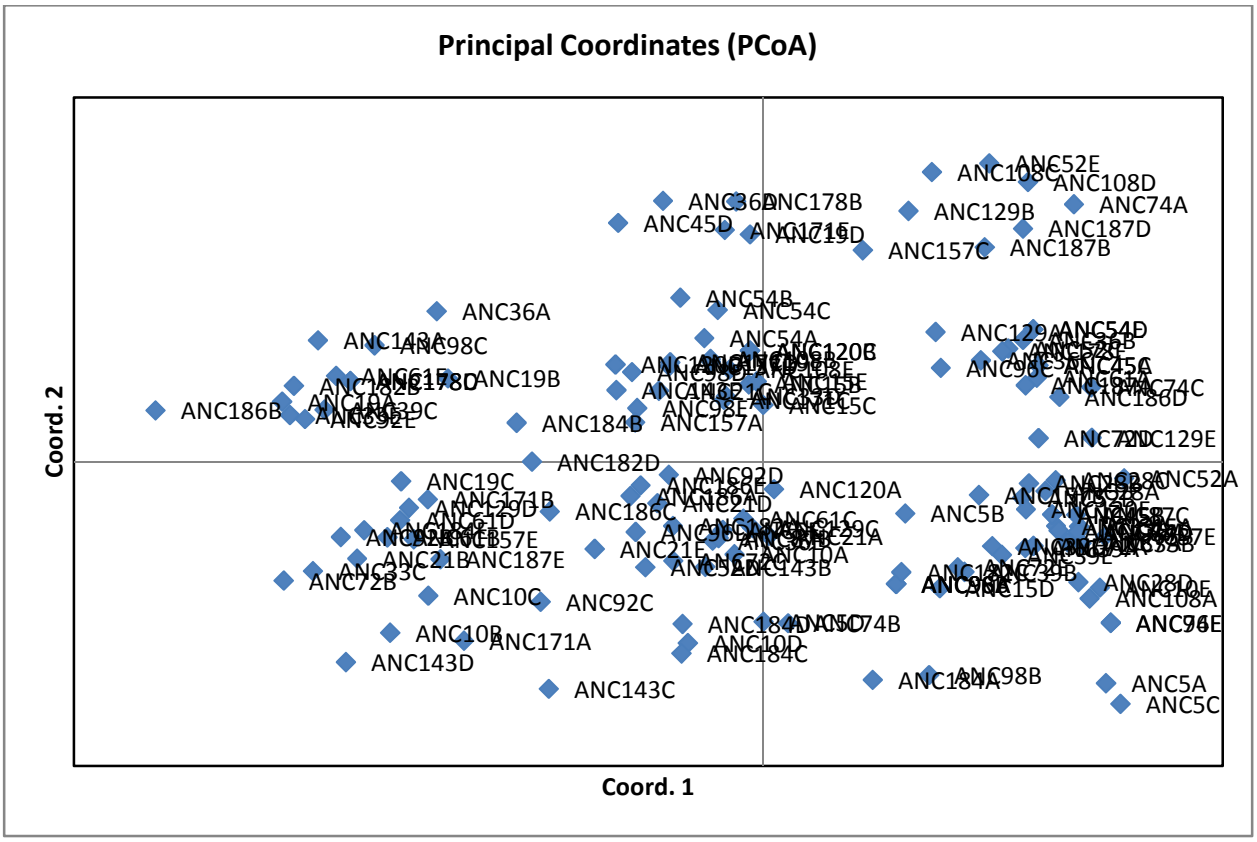
Populations based principal coordinate analysis (PCoA) resulted in some patterns of clustering, regarding their geographic origin, such as Sayo-Anifilo (Sy-An), Gimbi-Nejo (Gm-Nj), and Metu-Alle (Mt-Al) (on positive side of both coordinates) as one group and the rest populations (Gr-Sh, Dd-Dc, Gg-Dg, Ac-Bt, Ss-Wt, and Ar-Ld) (on negative coordinates) as different group, while others were dispersed all over the coordinates. Conversely, for accession-level analysis, the resulting cluster shows admixture of accessions without meaningful pattern of their

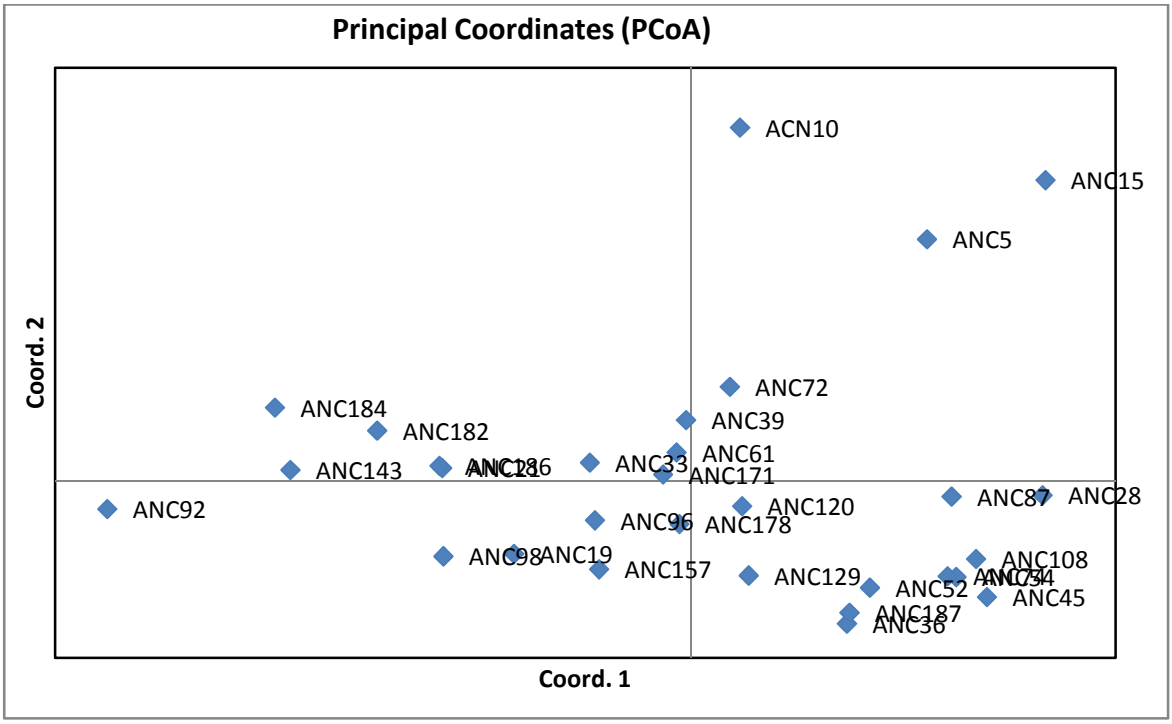

EW
ILLu
Jimma
WW
KW
OG
8
2
0
Dissimilarity coefficient

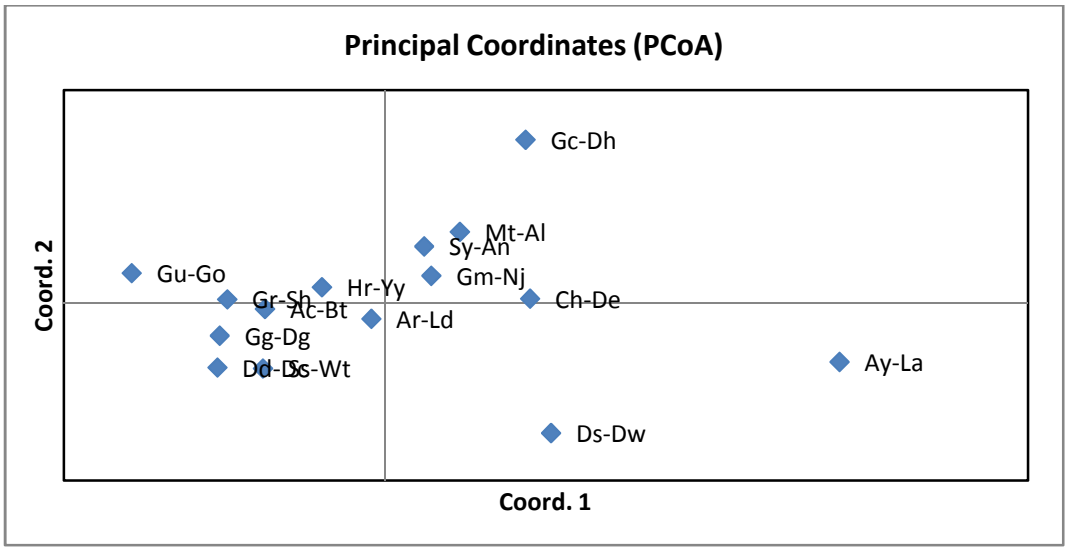

Dw
Ay
Ds
La
De
Sy
Ar
Nj
Ch
Wt
Mt
Gm
H
Al
Dc
Yy
Ud
Ar
Dg
Ac
La
Gc
Dh
Sh
Gu
Ss
Cg
Br
Go
Gr
0.1

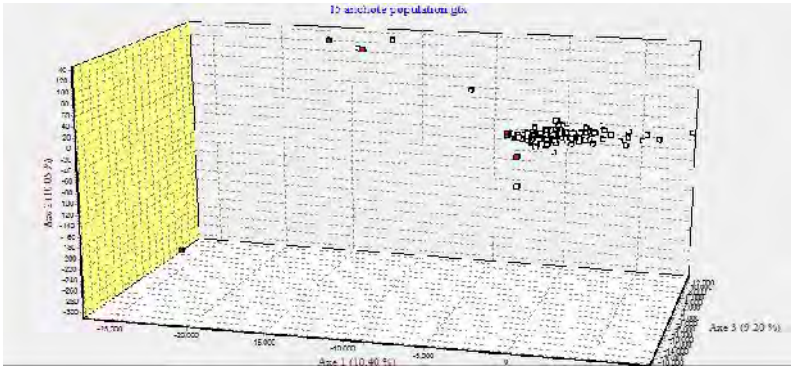

Ar-Ld
Ds-Dw
Sy-An
Ch-De
Ay La
Gm-Nj
Mt Al
Gc-Dh
Fr-Yy
Dd Dc
Gr-Sh
Ss-Wt
Gg-Cg
Gu-Co
Ac-Et
0.01

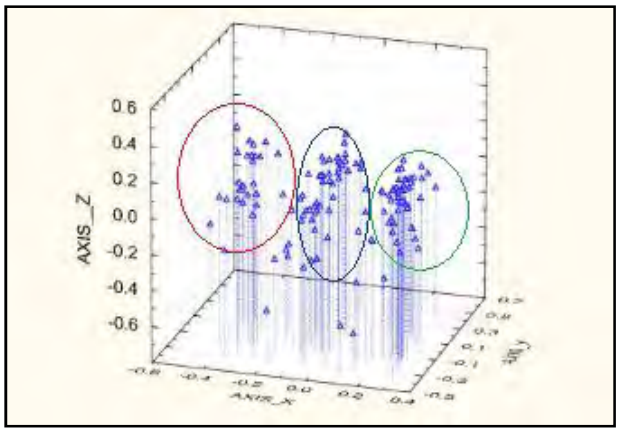

121
71
34
33
120
66
69
73
109
27
24
23
128
20
89
52
43
11
100
10
104
102
124
56
90
37
126
70
95
6
53
84
12
67
135
81
127
101
148
16
15
107
98
97
123
14
72
31
64
62
125
77
105
68
2


Principal Coordinates (PCoA)
Coord. 2
Coord. 1
ANC52E
ANC108C
ANC108D
ANC74A
ANC178B
ANC129B
ANC45D
ANC187D
ANC157C
ANC187B
ANC54B
ANC54C
ANC36A
ANC54A
ANC98C
ANC19B
ANC186B
ANC184B
ANC157A
ANC182D
ANC19C
ANC92D
ANC120A
ANC5B
ANC186C
ANC21D
ANC21E
ANC187E
ANC72B
ANC10C
ANC92C
ANC10B
ANC171A
ANC143D
ANC143C
ANC184A
ANC98B
ANC5A
ANC5C
ANC52A
ANC108A
ANC28D


Principal Coordinates (PCoA)
Coord. 2
Coord. 1
ACN10
ANC15
ANC5
ANC72
ANC184
ANC39
ANC182
ANC61
ANC143
ANC33
ANC171
ANC92
ANC120
ANC87
ANC28
ANC96
ANC178
ANC98
ANC19
ANC157
ANC129
ANC108
ANC52
ANC45
ANC187
ANC36


Principal Coordinates (PCoA)
Gc-Dh
Mt-Al
Sy-An
Gu-Go
Gm-Nj
Hr-Yy
Gr-Sh
Ac-Bt
Ch-De
Ar-Ld
Gg-Dg
Ay-La
Dd-Ds
Wt
Ds-Dw
Coord. 2
Coord. 1

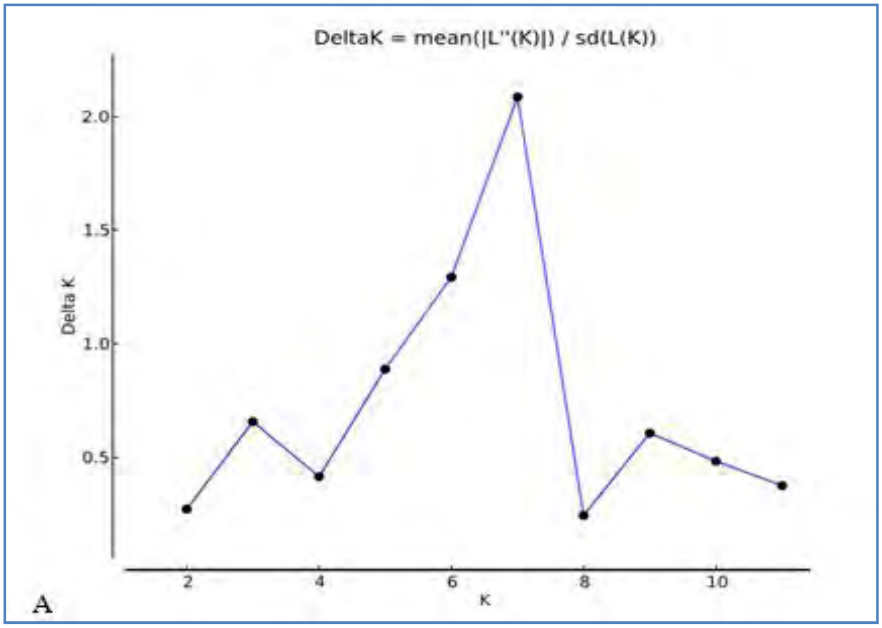


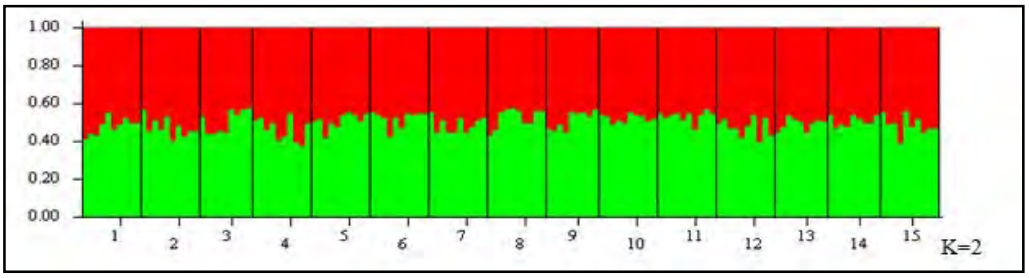

0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
AXIS_Z

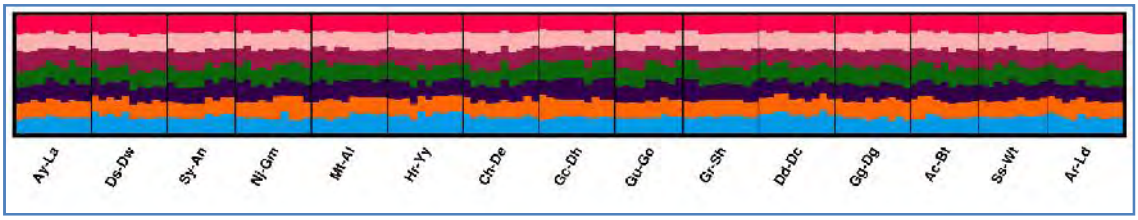
5.2.5. Population Genetic Structure Analysis and Bottleneck

Structure outputs showed that there are different clusters formed by varying the running lengths of STRUCTURE software. However, following the recommendation of Gilbert *et al.* (2012) we obtained K = 7. The result clearly indicated that at any Ks there is no clear sub-structuring of anchote accessions or populations similar to the result obtained by bar plots of K2-to-K7. CLUMPP software aligned the clusters and the plot was visualized by DISTRUCT software, the result of which resembled structure bar plots of all the Ks (Figure 24).

Detection of recent population bottleneck and estimation of effective population size using Sign and Wilcoxon (non-parametric) tests, which are based on heterozygosity excess and mode shift method under either of the three mutational models, we obtained certain significant deviation (indication of bottleneck) for four populations (Appendix Table 13). Two populations were represented from two zones each (Gumay-Gomma and Gera-Shebe Sombo from Jimma zone; Guto Gida-Diga and Sibu Sire-Wayu Tuka from East Wollega). However, many more populations show deviation from normal L-shape or mode shift without significant deviation. Certainly, those populations showing significant deviation might experience a recent bottleneck effect.


DeltaK = mean(|L''(K)|) / sd(L(K))
Delta K
2.0
1.5
1.0
0.5
2
4
6
8
10
K
A


1.00
0.80
0.60
0.40
0.20
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
K=2


Ay-La
Ds-Dw
Sy-An
Nj-Gm
Mt-Al
Hr-Yy
Ch-De
Gc-Dh
Gu-Go
Gr-Sh
Dd-Dc
Gg-Dg
Ac-Bt
Ss-Wt
Ar-Ld

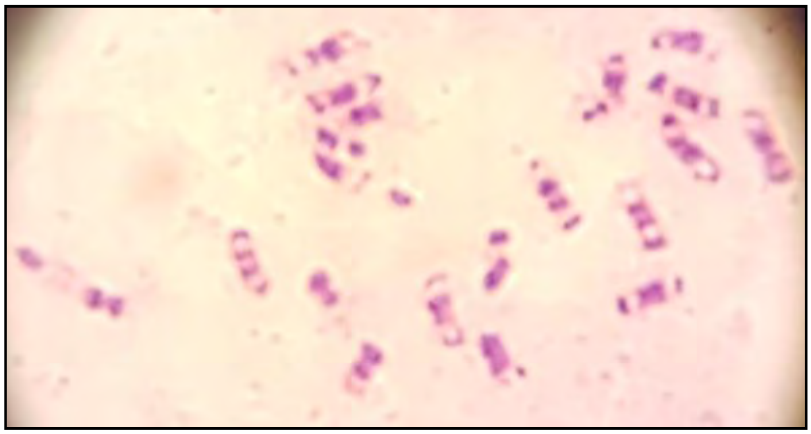
5.3. Karyotype of Anchote

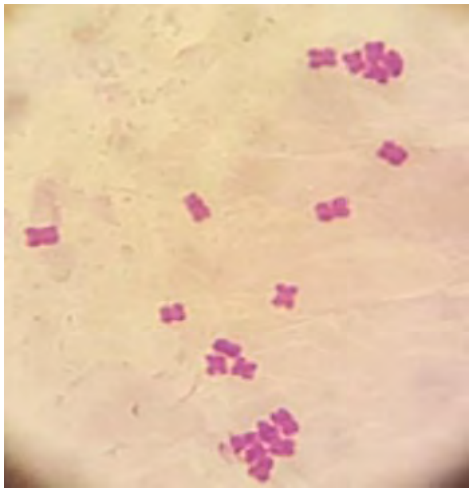
5.3.1. Analysis of Mitotic Chromosome Spread

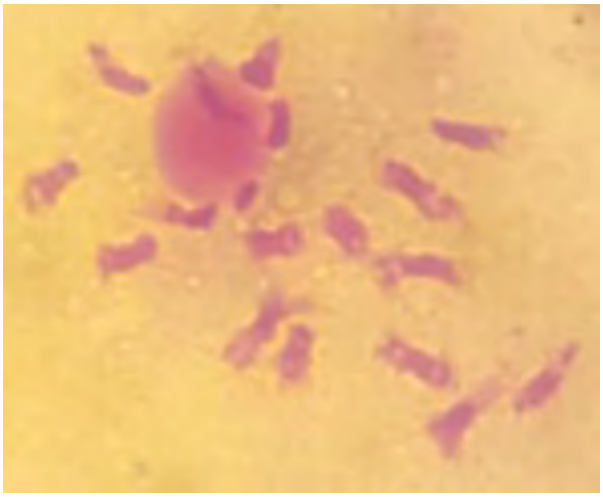
Slides with good quality of mitotic chromosome spread were selected under a microscope and photomicrographs were taken (Figures 16 A-C). In pro-metaphase stage, it was difficult to identify the centromeric regions, whereas in metaphase stage these regions are easily identified and used to characterizing chromosomes. Late prophase stage cells with larger nucleolus were observed in most of the cells (Figure 16C).

Late prophase/prometaphase chromosomes are characterized by differential degree of staining along their length, darkly staining segments alternating with lightly staining segments which apparently look like a deliberately induced banding pattern (e.g. C-banding). This is apparently due to the fact that tightly condensed constitutive chromatin region alternating with loosely condensed euchromatin region. The darkly staining regions are detected in interphase nuclei as numerous dark dots. It seems that as chromosome condensation progress towards the end of prophase, the number of these dots decreases but their size increases due to possibly coalescing of the smaller dots into larger dots. By the late prophase/ prometaphase stage only a few but larger sized blocks of darkly staining regions are separated by weakly staining (euchromatin) regions. It appears that the number, size and location of these darkly staining blocks of chromatin are chromosome characteristics and may be by this stage, the euchromatin used to identify homologous.

As chromosome condensation is progressing towards metaphase the differential staining disappears and only uniformly staining chromosomes are evident. May be the euchromatin







$\overline{Sd}$ $\bar{X}$

Table 36. Karyotypic formulae of anchote using Micro-Measure

Chromosome	Length	% of set	Arm length		Arm ratio(L/S)	Cent. Index	SA/LA	Chromosome nomenclature
			LA	SA				
1	1.649	0.064	0.925	0.724	1.277	0.439	0.783	m
2	1.608	0.062	0.898	0.710	1.265	0.442	0.791	m
3	1.471	0.057	0.796	0.675	1.179	0.459	0.848	m
4	1.411	0.055	0.724	0.687	1.054	0.487	0.949	m
5	1.402	0.054	0.853	0.549	1.553	0.392	0.644	m
6	1.394	0.054	0.710	0.684	1.038	0.491	0.963	m
7	1.356	0.053	0.813	0.543	1.499	0.400	0.667	m
8	1.345	0.052	0.710	0.635	1.118	0.472	0.894	m
9	1.335	0.052	0.710	0.625	1.135	0.468	0.881	m
10	1.329	0.052	0.675	0.654	1.032	0.492	0.969	m
11	1.319	0.051	0.773	0.546	1.414	0.414	0.707	m
12	1.236	0.048	0.724	0.512	1.414	0.414	0.707	m
13	1.230	0.048	0.684	0.546	1.252	0.444	0.799	m
14	1.208	0.047	0.724	0.484	1.497	0.400	0.668	m
15	1.174	0.046	0.599	0.575	1.042	0.490	0.960	m
16	1.168	0.045	0.710	0.458	1.550	0.392	0.645	m
17	1.110	0.043	0.684	0.426	1.606	0.384	0.623	m
18	1.092	0.042	0.648	0.445	1.457	0.407	0.686	m
19	1.062	0.041	0.539	0.524	1.029	0.493	0.972	m
20	0.893	0.035	0.475	0.418	1.138	0.468	0.879	m
Totals for set	25.790	1.000	14.372	11.418	1.259	0.443	0.794	
Average	1.289	0.050	0.719	0.571	0.063	0.022	0.040	
SD	0.183	0.007	0.109	0.098	0.203	0.039	0.126	
A_1	0.960							
TF%	44.274							
A_2	0.142							

ns = not significant (Chi square significant test at $\acute{\alpha} = 0.05$), A_1 = intra-chromosomal asymmetry, A_2 = inter-chromosomal asymmetry

NB. Mag: 1000 Image resolution: 100 pixels per cm

6. DISCUSSION

6.1. Agro-Morphological Traits Variation

6.1.1. Variations in Qualitative Traits

Based on the root color, anchote growers divided the crop into two major groups, white and reddish anchote (Abera Hora *et al.,* 1995). They prefer the white anchote for its taste and easy cooking and the reddish type for folklore medicines (personal communication). These two contrasting economical and cultural interests are very important for conserving both phenotypes "*in situ*".

In this study, certain traits were observed as either lacking or scarce in some accessions or populations, while in others they occur more frequently, so that each population has unique phenotypic richness/deficiency for specific trait. For instance, Chora-Deega and Arjo-Leka Dulecha populations are rich in phenotypic diversity because they possess all the phenotypes considered in this study. Such trait with significant variation between or within groups can be used as descriptors in order to study the heritable (genetic) variation of anchote accessions (and populations).

Therefore, most of the traits used for this study can be taken as descriptors and/or morphological markers since they have higher Shannon diversity indices that range from 0.592 to 0.992. Traits like deep green and purplish stem color are observed to be rare phenotypes (<20% in proportion), which might have been resulted due to being newly emerging or genetically eroded traits as the result of recent population genetic drift as evidenced from populations like Guto Gida-Diga and Sibu Sire-Wayu Tuka (section 5.3.5).

6.1.2. Variations in Quantitative Traits and Components of Variances

The significant variation of mean performances of anchote accessions observed in this study is an indication of higher genetic diversity occurring among the materials that can serve as the genetic resource for its improvement. Although the variations are large between some accessions (or between groups of accessions considered as populations), Duncan grouping method based on each trait showed very limited number of groups. The critical differences are larger among accessions than between Woredas or broader district covering populations, i.e., within accessions variation is more prominent than between groups. From this study, one can think anchote is an outcrossing crop, since its flowering biology is protandry, i.e., the male flowers bloom first on lower nodes and females develop later on upper nodes, at least after two weeks of male flowers already withered (Figure 6, personal observation). This type of mechanism has been developed in many plant species to encourage outbreeding.

Individual plant based root weight measurement showed great variability, i.e., 2 g (ANC 43) to 1,155 g (ANC 28). Some accessions ANC 62, ANC 65, ANC 60 (all from Chora), ANC 39 (Metu), and ANC 157 (Diga) (when averaged over experimental conditions) show higher yield performance. Accessions with higher and lower mean performances can serve as breeding parents for further anchote improvement through activities like selection of segregating lines, pure lines release and maintenance.

Accessions from certain Woredas are high root yielders than others, among which Chora (26.45 t/ha), Yayo (24.56 t/ha), and Dhidhesa (22.57 t/ha) (all from IlluAbabora zone) took the first three consecutive ranks, whereas the higher percentage of dry matter contents were observed in Nunu Kumba (22.93%), Guliso (21.43%), and Diga (21.53%). Conversely, Nunu Kumba (11.33

t/ha) and Guliso (14.82 t/ha) are among the least in root yield, leaf length and diameter, and petiole length. This result certified that dry matter content and yield have contrasting relationship as seen in Simple Pearson's Correlation Coefficient (Table 11).

Populations based mean performance indicated that Hurumu-Yayo (23.4 t/ha), Metu-Alle (21.98 t/ha), and Chora-Deega (22.8 t/ha) (all from IlluAbabBora zone) are the most anchote yielder or better mean performing ones, while the least performing population is Dedo-Decha (14.0 t/ha). Therefore, this zone can be considered as the potential area for anchote improvement effort for high yielding genotypes. Populations with higher dry matter content are Guto Gida-Digga (21.1%) and Guliso-Gida Gebo (21.5%), which are geographically far apart (one located in East Wollega and the other in West Wollega zones). One study aimed to optimize fertilizer rates that give better yield, resulted in a maximum of 15.42 t/ha (Girma and Hailu, 2007), using farmyard manure of 8 t/ha, which is comparable with our result in this study. Recent report from farmers' field also showed that anchote yield was about eight ton per hectare (Anonymous, 2011). Daba Mengash *et al*. (2012) obtained much better anchote root yield (76.45 t/ha), under Ebantu agro-ecological condition, which is by far more than our maximum result. The recent work of Desta Fekadu (2010) also showed the best agronomic performance of anchote was recorded (80 t/ha) on Vertisol soil type in Debre-Ziet Agricultural Research Center, which is not considered as anchote's agro-ecology.

There is a tendency of environmental preference of yield and yield related traits. For instance, mean of dry matter content is higher in experimental environment one (envt-1) where leaf length, petiole length, root length and diameter are lower for this environment as compared to the rest experimental environments. Yield wise, experimental environment two (envt-2) is significantly

better than the other. Overall, experimental comparison showed that envt-1 is more suitable for traits like less DE and larger PEL; in Envt-3, INL and DMC perform better; and all the remaining traits are by far performed better in envt-2 than the two environments.

As anchote is a climbing (prostrate) plant, it is sensitive to touching (thigmotropism) and needs uniform support (stalk) to grow uniformly. Therefore, plant height and related characters are difficult to be considered for diversity evaluation. Plant height affects yield and other traits, because it is directly contributing to root storage processes (photosynthesis). That is why some traits showed higher coefficients of variance (CV) like any other climbers (Gichimu *et al.*, 2009). Mainly, yield (YLD) and average root weight (AVRWT) showed much variation (higher CV) in all grouping variables of anchote accessions and the least were recorded for leaf traits (LL, LD and LLDR).

The existence of significant variations among accessions implies the importance of the used traits as the measure of genetic variability in anchote accessions. All the traits were significantly different among accessions. Traits that are significantly affected by accession-environmental interaction include internodes length (INL) and the petiole to inter-node length ratio (PINLR).

By taking limited morphological traits from root grown anchote we compared with that of seed grown anchote. The results showed that almost all the traits in root-grown anchote were superior than seed grown, indicating the importance of anchote propagation via root for larger foliage and seed production. That was why, we couldn't harvest enough fruit from the first year seed grown anchote, to incorporate into other agro-morphological traits for analyses, at both experimental environments. In this regard, farmers are aware of the problem and usually preserve the selected anchote mother plant ("gubo") for large (commercial) seed multiplication (Abera Hora *et al.*,

1995). The largest average number of seed per fruit (177) were obtained from accessions ANC46 (Ale) and ANC175 (Wayu Tuka), while larger fruits were obtained from accessions ANC40 (Metu) and ANC56 (Yayo).

Combined ANOVA results showed that only INL and LLDR were not significantly differed among accessions for root grown anchote as compared to seed grown anchote, whose all traits were significantly different among accessions and among environments.

In this study, the magnitude of phenotypic coefficient of variations (PCV) were higher than genotypic coefficient of variations (GCV), for all traits, that shows the advantage of using genotypic variance for better estimation of variance between accessions. Similar results (higher PCV) were observed for various crops diversity analyses such as wild and cultivated watermelon (*Citrullus sp*.) accessions (Gichimu *et al.*, 2009) and cassava (Boakye *et al.*, 2013).

High heritability and genetic advances, observed in most of the traits, confirmed the possibility of improvement of those traits though selection. Large differences between the PCV and GCV of some traits such as AVRWT, YLD, and INL are the indication of the trait's complexity, i.e., they are determined by other environmental factors. Relatively, leaf and root dimensions and petiole length are the traits with higher heritability (Table 10), indicating their resistance to environmental influences. In other words, the highest percentage of genetic gain was obtained for YLD trait, which shows the possibility of maximum (21%) improvement of the trait through selection whereas the least trait to be improved through selection is INL (1.4%).

6.1.3. Trait Relationships

Correlation coefficient has been used for similar response analysis of the traits against selection pressures. Therefore, high correlation coefficient may be the result of genetic factors such as pleiotropism or linkage of the traits (genes) (Comuzzie *et al*., 1997). In our study, most traits are significantly correlated ($p<0.05$) including yield and yield related traits, which implies the high co-heritability of many of these related traits as seen in traits of sweet sorghum (*Sorghum bicolor* L. Moench) (Kisua *et al*., 2015). Because yield is the derivative of root weight, average root weight is highly correlated with yield than any other traits. In this study, significant correlation coefficients were obtained for yield with leaf, root, petiole, and inter-node diameter and lengths and with number of primary branching. Dry matter content shows negative correlation with most traits but not statistically significant, except for inter-node length and leaf-diameter ratio. Slight negative yield to dry matter relationship shows the yield constituent is more of moisture content. High relationship between petiole and leaf size indicated that larger petiole supports large sized leaf. The cylindrical root (larger root length to root diameter ratio) and disproportion of PEL to INL may have large negative effect on yield (Table 11). Based on these results, the yield related traits could be used for early stage selection of anchote seedling for better yield.

Most traits do have significant regression coefficients implying that the small change of one of the trait changes the other trait largely, i.e., they are co-heritable traits. Most of the significance tests for correlation and regression show similar results (Table 11).

Character associations become more clearly understandable when genotypic and phenotypic correlations are separated. In most of the traits, the significant phenotypic correlations are more (58/78 = 74.35%) than the significant genotypic correlations (45/78 = 57.69%). For all

significant genotypic correlations, there are also significant phenotypic correlations and have similar sign and magnitude, indicating that those traits showing significance for both correlations are less affected by environmental factors and are suitable for improving them through selection (Table 12). Traits like yield and root weight show the highest significant genotypic and phenotypic correlations. However, for juvenile stage selection, vine number, leaf length and leaf diameter are the appropriate traits (Table 12).

6.1.4. Clustering and Relationship among Populations

Clustering based on qualitative and quantitative traits showed some pattern of groupings but did not strictly follow the geographical origins, except population-based clustering. Altitudinal clustering revealed that only lowland anchote accessions, that exhibiting larger INL trait, were separated from the rest, i.e., anchote's agro-morphological performance is not dependent on altitudes.

Qualitative morphological data grouped the 35 populations (Woredas) of anchote accessions into 13 major clusters, most clusters were separate Woredas, which include: Gw-characterized by reddish root and no lobed leaf; Bt-with heavy stem hairiness, late emerging of seedling, and no leaf lobbing; Gl-having deep green stem; Ss; Hr-the least in bushiness growth type; An, Sh-heavy in stem hair and late flowering; Al-late in flowering; Dc; Gd-least in bushy type; and Ac-absence of hairiness as well as light and purple stem color. Population from Nunu Kumba (Nk) was clustered regardless of its geographical location with Dw and La. Most of the woredas were clustered together, the indication of sharing common phenotypic frequencies. The 30 woredas that we grouped into 15 populations were subjected to cluster analysis. Cluster 1&2 were grouped according to their geographic proximity otherwise the rest grouping did not follow

pattern of geographical location. This implies the possibility of far distance seed exchange as indicated by Tilahun Wondimu (2014), Desta Fekadu (2010), and Abraham Bekele *et al*. (2014).

Population level clustering also shows some separate populations. For instance, populations like Ac-Bt and Gl-Gd are unique in some phenotypes, which separate them from the rest. Ac-Bt group has no deep primary lobbing, least frequency in intermediate stem hairiness, and dominant in purple stem color. The latter population is characterized by least in primary shallow lobbing of leaf, absence of stem hair, least in light green of stem, and flower early.

Qualitative traits based clustering of accessions also indicated that some of them formed groups according to their origin, whereas most of them were scattered over other groups.

Quantitative based cluster analysis showed that, except Gg-Dg, all the populations were clustered following their origin and that is why discriminant analysis assigned this population under cluster 4 or 5. The ungrouped population, Gl-Gd, is different in many of its traits including minimum YLD, AVRWT, and the least among all population in LL, LD, PEL, INL and RD. The second cluster includes three populations (Mt-Al, Hr-Yy, and Ch-De) which are similar in some of their traits. They share high stem primary branching, better in YLD, AVRWT, LL, LD, PEL, and RD but least in INL.

Clustering at Woreda level indicated the significance of some traits to discriminate them into different groups. The better stem branching (VN), larger RD, and better YLD put Ale Woreda as a separate group. Nunu Kumba, with least root yield but highest in dry matter content is another separate group. Dhidhesa and Guliso formed other two separate clusters, with their respective characteristic traits like least RL for the former and least in LL and PEL, but highest in RL, for

the latter. Although association of traits can be elucidated from such analyses, there is no concordance between clusters made by qualitative and quantitative traits. This could be caused by multiple reasons including, the accessions by themselves are not pure lines, which contribute for sampling errors (biases) during data recording or measuring of traits; quantitative traits are more affected by environmental factors than qualitative traits so that their discriminating or clustering power became low.

Discriminate analysis for accessions showed that most of them were grouped under their respective cluster. However, some accessions were distributed in other clusters, explaining the possibility of the seed (genetic) exchange at common markets or even over longer distances by merchants.

6.1.5. Principal Components

The first five PCs are responsible for most of the variances observed between anchote accessions, for quantitative traits whereas nine PCs are required for qualitative traits. This is the indication of higher discrimination powers of each qualitative (phenotypic) trait than metric traits, so that they are distributed over many PCs. The determination of variances among accessions by many PCs is, usually, considered as a guarantee to use the traits for genetic diversity and conservation study of the plant (Asnakech Tekalign, 2014).

For quantitative data, the first PC was mostly of yield and yield related components while the second PC was responsible for internodes and petiole lengths components. PC's loading plot also explained the distributional patterns or relationships of anchote traits and accessions or groups on the first two PCs.

From those loading plots, it is possible to observe the most convenient traits for diversity study since some accessions are uniquely separated from the rest, by using the variances of those traits (PCs). The accessions located near to zero on x-y axis are the ones which show minimum overall variations by the two PCs and those distributed to either of the extreme corner of the axis may be those showing significant variations by the traits highly contributed for the PCs.

6.2. SSR Based Diversity and Population Structure in Anchote

6.2.1. EST-SSR Markers Polymorphism

A total of eight EST- SSR markerswere selected out of the initial 13 screened, i.e., five of them were monomorphic (non-informative markers). All the five monomorphic markers are composed of AT di-nucleotide repeats. Theses markers could be rare alleles so that more samples could be required to elucidate them (Kalinowski, 2004) and the genotypes in our collection may represent similar groups for these markers. The rare alleles are usually related with out-breeding population where they occur in the form of heterozygous (Frankham *et al.,* 2002).

The two commonly occurring types of microsatellites (di-nucleotide = 77% and tri-nucleotide = 13%) observed in anchote agreed with studies on other species such as mango (*Mangifera indic*a L.) plants (Dillon *et al*., 2014), *Jatropha curcas* L. (Wen *et al*., 2010), and Lettuce (*Lactuca sativa* L. (de Simko, 2009). However, since EST-SSRs have potential to cause selective deleterious frame shift mutation, usually they occur in tri-nucleotide repeats (Ellis and Burke, 2007).

The maximum number of alleles found per locus (EST-SSR marker) was six, which is relatively small as compared to genomic SSR but comparable with results of other studies (Reddy, 2009) in Cucurbitaceae spp., Mao *et al*. (2014) in *Cucurbita pepo*, Abel Teshome *et al.* (2014) in

Ethiopian field pea (*Pisum sativum* L.). In contrast to genomic SSR, EST-SSR markers are not completely neutral DNA region but usually have a limited mutation rate that could results in the formation of few alleles per locus (Ellis and Burke, 2007).

Anchote accessions/populations do not have equal allele frequency which is the sign of gene flow between populations or bottleneck effect.

Most of the loci that we investigated were moderately informative while locus WM-24 was highest in allelic richness, PIC, and relatively the least in possessing null alleles. Such marker irregularity in polymorphism may affect the estimation of population differentiation. For instance, overestimation of population differentiation can happen when null alleles are present in a locus (Chapuis and Estoup, 2007), since it reduces the genetic diversity within populations (Paetkau and Strobeck, 1995).

Deviation from H-W equilibrium of the three loci (WM-24, WM-34 and WM-29) and the three populations (Metu-Alle, Hurumu-Yayo and Chora-Deega) may be attributed to the excess or deficient heterozygosity or even to minor SSR genotyping error of those loci and populations, respectively (Morin *et al.,* 2009). Absence of pair of loci with significant linkage disequilibrium is the indication of nonrandom association of alleles of different loci (Slatkin, 2008).

6.2.2. Genetic Diversity and Differentiation in Anchote Populations

This study showed moderate genetic diversity using eight EST-SSR markers inferring the modest selection pressure based on some economically (Wang *et al.*, 2011) or culturally (personal observation) valuable traits. Population with highest genetic diversities such as Ds-Dw, Gm-Nj and Ac-Bt can be considered as anchote diversification locations, whereas Gc-Dh and Ch-De

were the populations scored the least genetic diversity, which might be inferred to recent introduction of anchote to these areas (founder effect) or population bottleneck (Hawks *et al*., 2000) or intensive artificial selection pressure to maximize root yield (section 5.1.2.1). However, there is no clue about bottleneck effect for these populations (section 5.3.5).

Low-to- moderate population differentiations (F_{ST} < 0.15 or G_{ST} = 0.06) observed in our study revealed that, anchote accessions have little sub-structuring into demes, but considered as a wider randomly mating population, with significant gene flow. This premise is supported by high and significant gene flow (average N_m = 2.06) as computed in this study. In other words, fixation indices ($F_{IS,}$ $F_{ST,}$ and G_{ST}), the measure of degree of differentiation, were higher for WM-29 and WM-25 loci. These loci have higher heterozygote deficiency, because for most populations the alleles of these loci were less than two (Table 29). Contrary to these, locus WM-34, WM-32, and WM-24 represented by excess heterozygosity (negative fixation indices), indicating that anchote population is not strictly inbreeding or because of selection favoring of the heterozygote, whereas the positive indices for the rest loci referring the existence of inbreeding system (Koff *et al.,* 2008). Such disparity in fixation indices within a genome provides important insights into the genome regions under evolutionary processes (or selection) beyond providing the demographic history of a population (Holsinger and Weir, 2009). The genetic differentiation measures vary greatly, with mean value of F_{ST} = 0.11 and G_{ST} = 0.06, because they follow different algorism while considering mutation effect. In such cases, more markers should be applied to minimize the biasness introduced by any one locus (Frankham *et al.,* 2002).

AMOVA partitioned the variance between groups and within groups, however in all the cases within group variations (>83%) were much higher than between groups (Table 34). This is also because of high genetic exchange between anchote populations.

Anchote lacks strong geographic based clustering instead showed high level of admixture between subpopulations. This suggests that the geographical limit of the plant is very wider across different agro-ecologies or the distinct anchote population had spread throughout the sample collection areas. There is also high gene flow between the subpopulations as a consequence population differentiation became negligible.

Distance based (D_S) neighbor-joining tree showed that Arjo-Leka Dulecha (Ar-Ld), Dalle Sadi-Dalle Wobera (Ds-Dw), and Gera- Shabe Sombo (Gr-Sh) are the oldest lineages, while Gechi-Dhidhessa (Gc-Dh) and Ayira-Leka Dulecha (Ay-La) seem recent populations. However, the tree was not supported by strong bootstrap values (Figure 18-20), the indication of high genetic exchange between the groups.

6.2.3. Population Genetic Structure

The STRUCTURE result indicated as anchote is sub-structured into 7 groups (K = 7), however the DISTRUCT plot did not support any apparent population sub-division. The uniformity in proportion of each color (genotype's proportion) among all the population, in the DISTRUCT plot, is the indication of every individual/population can have equal probability of being a member of any other group. Structure results are sensitive for marker type, number of loci, number of populations and number of individuals in a sample (Evano *et al.,* 2005).

6.2.4. Population Bottleneck

Evaluation of population bottleneck have been used for conservation strategy, especially for geographically limited (endemic) and indigenous crops. In this study, no any significant test results obtained by both Sign and Wilcoxon methods, simultaneously, using either of mutation models which accompanied with mode shift. In other words, there are populations with mode shift but not supported by significant tests. As microsatellite mutation has been usually defined by S.M model, only Sibu Sire-Wayu Tuka population experienced recent population bottleneck (Appendix Table 13). Since the S.M is strict and one step model, for microsatellite T.P model is preferred (Kuo and Janzen, 2004), under which our result became four populations (Gumay-Gomma, Gera-Shabe Sombo, Guto Gida-Diga, Sibu Sire-Wayu Tuka) which experienced bottleneck. However, simply because of its endemic and indigenous nature, anchote needs conservation whether it passed through bottleneck effect or not.

6.3. Cytogenetic Analysis of Anchote

With the aim of determining somatic chromosome number and the ploidy level of anchote, this study found that anchote has 2n = 20 chromosomes and it is diploid (2n = 2x = 20). The diploid observed in this study agreed with that of most of Cucurbitaceae species. However, some uncommon hexa-, octa-, and deca-ploid cytotypes (2n = 6x = 66, 8x = 88, and 10x = 110) were reported for *Trichosanthes kirilowii* (Waminal and Kim, 2015). The present result is the first report on anchote chromosome number along with its karyotypic description. Our result matched with the inferential conclusion of Holstein, (2012) from her work of nuclear and plastid DNA sequences of phylogenetic analysis of 30 *Coccinia* species. The Holstein's diploid number (2n =

20) was inferred based on the two investigated species (*C. trilobata* and *C. rehmannii*), the group which included *C. abyssinica* in her phylogeny.

The size comparison of all the twenty chromosomes of *C. abyssinica* showed that they constitute somatic chromosomes, i.e., there is no heteromorphic or sex chromosomes, unlike its sister species *C. grandis* (2n = 22+XX/XY) (Holstein, 2012) in which both sexes occur on separate plants (Gautam, 2013). The absence of differentiated sex chromosomes is consistent with the monoecious sex of *C. abyssinica*.

Under condensed chromosomes of prophase and pro-metaphase chromosomes showed a sort of banding patterns, indicating the presence of regions that are in relaxed (euchromatin) and condensed (heterochromatin) states in a chromosome. This pattern was observed in all chromosomes and the number, size, and location appear chromosome specific such that homologous chromosomes can easily be distinguished pattern.

In metaphase, the whole regions of the chromosomes are fully condensed. All chromosomes have similar arm ratio, which is an indication of symmetrical type of karyotype (Kalvandi *et al*., 2012). However, the reduced chromosome number, in relation to other *Coccinia* species, makes anchote an advanced or evolved karyotype among its phylogenetic group (Stebbins, 1974; Sutar *et al*., 2013).

Karyotypic parameter analyses show that there was no significant variation between anchote chromosomes within relative size and arm ratio. This uniformity of chromosome parameters as described by Chen *et al*., (1997) proves the absence of strict self-pollination of anchote crop. In other way, chromosome symmetric parameters: A_1 = 0.02 and A_2 = 0.3, investigated in this

study, characterize anchote crop. The larger nucleolus observed in anchote root cells may indicate that the tissue is highly engaged in protein synthesis (Cooper and Hausman, 2007). This result is in agreement with earlier works on anchote's high protein content (Habtamu Fufa and Kelbessa Urga, 1997).

The information generated from our findings will provide guideline for breeders and biotechnologists for further crossing (hybridization) or mapping genes of anchote on their respective chromosomes. Further works are expected to identify possible existence of anchote cytotypes.

7. CONCLUSION

This study is the first comprehensive report on anchote regarding its agro-morphological evaluation, over multiple experimental environments by growing from seed and root, cytogenetic preparation and chromosome count, and EST-SSR markers development as well as genetic diversity and population structuring analyses using these markers.

Anchote can grow in wider agro-ecologies and perform similarly but needs good cultivation practices and soil fertility.

Based on the morphological descriptors used in this study, anchote accessions have significant variability, indicating the broad genetic diversity. Both common and rare morphological traits were also identified.

Mean based comparative analyses of anchote accessions, using different methods, showed that some are superior with certain traits while others by another trait(s).

Anchote plant grown from “gubo” (root mother plant) produces more fruits (seeds) than the one raised from the seed.

Clustering analyses of anchote showed admixtures of accessions or groups rather than with traceable patterns of geographical origin. This is because of the existence of high gene flow (seed exchange) between groups.

The chromosome study of anchote indicated that it is diploid (2n = 20), with all chromosomes being metacentric forming a symmetrical karyotype.

In this study a total of 13 EST-SSR markers were successfully developed for anchote from watermelon DNA sequence. This is the first anchote crop focused molecular marker developed, which is used for its genetic diversity study and can also be used beyond, since the marker is highly transferable.

Moderate mean genetic diversity and low genetic differentiations identified in this study show that anchote did not form populations separated by geographic barriers. Moreover, highest within population variance also shows that the crop involved cross-pollination so that the genetic diversity will be maintained.

8. RECOMMENDATION

- Anchote's biogeographical distribution should be clearly known, together with its domestication history, since it is limited to southern and southwestern parts of Ethiopia.
- Although anchote has moderate genetic diversity, the occurrence of rare traits may be the indication of genetic erosion, hence it needs conservation priority.
- For anchote improvement, through selection, one can use some traits of high genetic correlation with yield such as leaf length, leaf diameter, petiole length, and root diameter.
- For the effort of seed multiplication and evaluation, "gubo"- anchote mother plant is preferable than seed grown anchote, since the later method produced limited number of fruits or seeds.
- The full scale karyotypic studies on many anchote accessions/ populations and linkage mapping may reveal different cytotypes of anchote.
- EST-SSR markers can be used for several applications such as taxonomic status determination of the genus (*Coccinia*) and its subspecies, phylogenetics, and linkage mapping or QTL analyses.
- Exhaustive and all round methods including some additional EST-SSR and/or other molecular markers may give better detailed diversity and population structuring of anchote crop.

9. REFERENCES

Abd El- Mohsen, A.A. and Abo-Hegazy, S.R. (2013). Comparing the relative efficiency of two experimental designs in wheat field trials. *Scientific Research and Review Journal* **1(3):** 101–109.

Abel Teshome, Bryngelsson, T., Kifle Dagne, and Mulatu Geleta (2015). Assessment of genetic diversity in Ethiopian field pea (*Pisum sativum* L.) accessions with newly developed EST-SSR markers. *BMC Genetics* **16**:102. DOI 10.1186/s12863-015-0261-5.

Abera Hora, Edwards, S., Mirutse Giday, and Yilma Tesfaye (1995). *Anchote -An Endemic Tuber Crop.* Artistic Printing Press Enterprise, Addis Ababa, pp.1–75.

Abraham Bekele, Tileye Feyissa, and Kassahun Tesfaye (2014). Genetic diversity of anchote (*Coccinia abyssinica* (Lam.) Cogn.) from Ethiopia as revealed by ISSR markers. *Genet Resour Crop Evol* **61**: 707–719.

Acquaah, G. (2007). *Principles of Plant Genetics and Breeding*. Blackwell Publishing, Malden, pp.178.

Agrawal, B.L. (1996). *Basic Statistics,* 3rd edn. New Age International Limited, New Delhi, pp.1–713.

Akhtar, M.S., Oki,Y., Adachi,T. and Khan., Md. H. R. (2007). Analyses of the Genetic Parameters (Variability, Heritability, Genetic Advance, Relationship of Yield and Yield Contributing Characters) for some Plant Traits among Brassica Cultivars under Phosphorus Starved Environmental Cues. *Journal of the Faculty of Environmental Science and Technology* **1(12)**: 91–98.

Almeda, F. (2013). Systematic and phylogenetic significance of chromosome number diversity in some Neotropical Melastomataceae. *Mem. New York Bot. Garden* **108**: 167–189.

Amare Getahun (1973). Developmental Anatomy of "tuber"s of Anchote: A potential Dry land Crop. **In**: *Acta Horticulturae, Technical Communication of ISHS.*51–63.

Anderson, J., Churchill, G., Autrique, J., Tanksley, S., Sorrells, M. (1993). Optimizing parental selection for genetic linkage maps. *Genome* **36**: 181–186.

Anonymous (2011). *West Wallaga Zone Agriculture and Rural Development Office.* Annual Report pp. 1–22.

Asnakech Tekalign Beyene (2014). *Genetic analysis and characterization of faba bean (Vicia faba) for resistance to chocolate spot (Botrytis fabae) disease and yield in the Ethiopian highlands.* Ph.D. Dissertation, University of KwaZulu-Natal, Pietermaritzburg.

Balkerm-boomstra A.G., Zijlstra, S., verstappen, F.W., Inggamer, H., Mercke, P.E. Jongsman, M.A. Bouwmeester, H.J. (2003). Role of Cucurbitacin C in Resistance to Spider Mite (Tetranychus urticae) in Cucumbr (*Cucumis sativus* L.). *J. Chem Ecol.* **29(1)**: 225–235.

Ballard L, Adams P, Bao Y, Bartley D, Bintzler D, Kasch L. *et al*. (2002). Strategies for genotyping: Effectiveness of tailing primers to increase accuracy in short tandem repeat determinations. *Biomol Tech*. **3**: 20–9.

Berndt, R. (2007). A global survey of Puccinia-rust on Cucurbitaceae. *Mycol. Progress* **6**: 151–178.

Bioversity International (2007). Guidelines for the development of crop descriptor lists. Bioversity Technical Bulletin Series. *Bioversity International* Rome, Italy, pp.1–72.

Blanca, J., Canizares, J., Roig, C., Ziarsolo, P., Nuez, F., Pico, B. (2011). Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genomics*, **12:** 104–119.

Boakye, B., Kwadwo, O., Isaac, K.A., and Parkes, E.Y. (2013). Genetic variability of three cassava traits across three locations in Ghana. *Afr. J. Plant Sci***. 7(7):** 265–267.

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

Botstein, D., White, R.L., Skolnick, M. et al., (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* **32**: 314–331.

Bula Sirika Wayessa (2016). *Toward a history of the Oromo of Wallaga in southwestern Ethiopia: an ethnoarchaeological study of ceramic technological style and tuber crop domestication*. Ph.D. dissertation, University of Calgary, Alberta.

Cavalli-Sforza, L.L., Edwards, A.W.F. (1967). Phylogenetic Analysis: Models and Estimation Procedures. *Am. J. Hum. Genet*. **19**: 233–257.

Chapuis, M.P., Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Mol Bio Evol* **24**: 621–631.

Chen, J.F., Staub, J.E., Tashiro, Y., Issiki, S., Myiazaki, S. (1997). Successful interspecific hybridization between *Cucumis sativus* L. and *C. hystrix* Chakr. *Euphytica*, **96**: 413–419.

Comuzzie, A.G., Mahaney, M.C., Almasy, L., Dyer, T.D. and Blangero, J. (1997). Exploiting pleiotropy to map genes for oligogenic phenotypes using extended pedigree data. *Genetic Epidemiology* **14**: 975–980.

Cooper, G.M. and Hausman, R.E. (2007). *The Cell: A Molecular Approach*, 4th edn. ASM, Press, Washington D.C. pp. 375–385.

Cornuet, J.M. and Luikart, G. (1996). Description and power analysis for two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.

Crow, J. and Kimura, M. (1970). *An introduction to population genetics theory*. Burgess Publishing Company, Minneapolis, United States of America.

Da Costa, C. P. and Jones, C.M. (1971). Cucumber Beetle Resistance and Mite Susceptibility Controlled by the Bitter Gene in *Cucumis sativus* L. *Science* **172**(**3988**): 1145–1146.

Daba Mengash, Derbew Belew, Wosene Gebresilasie, and Waktole Sori (2012). Growth and yield performance of Anchote [*Coccinia abyssinica* (Lam.) Cogn.] in response to Contrasting Environment. *Asian J. Plant Sci*. **11(4):** 172–181.

Dandena Gelmesa (2010). Shifting to alternative food source: Potential to overcome Ethiopians' malnutrition and poverty problems. *Innovation and Sustainable Development in Agriculture and Food* (ISDA). Montpellier-France, June 28–July 1st.

DARWIN (Dissimilarity Analysis and Representations for Windows) (2014). *http://Darwin.Cirad.fr*.

Dawit A, Estifanos H. (1991). Plants as a primary source of drugs in the traditional health practices of Ethiopia. **In**: *Plant genetic resources of Ethiopia*. pp.101–112, (Engles, J.M.M., Hawkes, J.G, and Worede M., eds). Cambridge University Press, Cambridge.

de Simko, I. (2009). Development of EST-SSR Markers for the Study of Population Structure in Lettuce (*Lactuca sativ*a L.). *Journal of Heredity* **100(2):** 256–262.

de Vienne, D., Santoni, S.and FaZque, M. (2003). Principal Sources of Molecular Markers. **In**: *Molecular Markers in Plant Genetics and Biotechnology*. pp. 3–46, (de Vienne, ed), Science Publishers, Inc. Enfield, USA.

Desta Fekadu (2011). *Phenotypic and nutritional characterization of Anchote (Coccinia abyssinica (Lam.) Cogn.) accessions of Ethiopia*. MSc thesis, Jimma Uiversity, Jimma.

Di Reinzo, A., Peterson, A.C. Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B. (1994). Microsatellite Mutation and Interferences about Human Population. *Proc Natal Acad Sci USA* **91(8)**: 3166–3170.

Dillon, N.L., Innes, D.J., Bally, I.S.E., Wright, C. L., Devitt, L.C., and Dietzgen, R.G. (2014). Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR) Marker Resources for Diversity Analysis of Mango (*Mangifera indica* L.). *Diversity* **6**: 72–87.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., Mattick, J.S., (1991). Touchdown PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, **19:**4008.

Edwards, K.J., Barker, J.H.A., Daly, A., Jones, C., Karp, A. (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques* **20:** 758–760.

Edwards, S., Mesfin Tadesse and Hedberg, I. (1995). *Flora of Ethiopia and Eritrea*. Vol. 2(2). National Herbarium, Addis Ababa University and Uppsala University, Sweden. pp 456.

Ellis, J.R. and Burke, J.M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity* **99:** 125–132.

Esquinas-Alcazar, J.T. and Gulick, P.J. (1983). *Genetic resources of Cucurbitaceae*: *a global report*. IBPGR Secretariat, Rome, Italy.

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

Excoffier, L., Laval, G. and Schneider, S. (2006). *Arlequin ver 3.01 User Manual: An Integrated Software Package for Population Genetics Data Analysis.* Computational and Molecular Population Genetics Lab (CMPG) Institute of Zoology University of Berne, Switzerland. http: p://cm//cmpp.unibe.ch/sog.unibe.ch/softwarwaree/ar/arlequiinn3.

Falconer, T.S. and Mackay, D.F.C. (1996). Introduction to Quantitative Genetics, 4th edn, Longman Green, UK. pp 464.

Ferra˜o, L.F.V., Caixeta, E.T., Pena, G., Zambolim, E.M., Cruz, C.D., Zambolim, L. *et al.* (2014). New EST–SSR markers of Coffea arabica: transferability and application to studies of molecular characterization and genetic mapping. *Mol Breeding* **35**: 31.

Frankham, R., Ballou, J.D., Briscoe, D.A. and Mclnness, K.H. (2002). *Introduction to conservation genetics.* Cambridge University Press, New York, pp. 61–617.

Gautam, G.A. (2013). Characterization of *Coccinia grandis* as a model system to study sex determination in plants. M.Sc. Thesis, Indian Institute of Science Education and Research (IISER), Pune.

Geleta, M., Herrera, I., Monz, A., Bryngelsson, T. (2012). Genetic Diversity of Arabica Coffee (*Coffea arabica* L) in Nicaragua as Estimated by Simple Sequence Repeat Markers. *Sci World J.* **2012**: 1–11.

Gibbs, R.D. (1974). *Chemotaxonomy of flowering plants*. McGill-Queens University Press. Montreal and London **2:** 829–843.

Gichimu, B.M., Owuor, B.O., Mwai, G.N. and Dida, M.M. (2009). Morphological characterization of some wild and cultivated watermelon (*Citrullus sp*.) accessions in Kenya. *Journal of Agricultural and Biological Science* **4(2):** 10–18.

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

Gilbert, K.J., Andrew, R.L., Bock,D.G., Franklin, M.T., Kane, N.C. *et al*. (2012). Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology* 1–6.

Girma Abera and Hailu Gudeta (2007). Response of Anchote (*Coccinia abyssinica*) to Organic and Inorganic Fertilizers Rates and Plant Population Density in Western Oromia, Ethiopia. *East African Journal of Sciences* **1(2):** 120–126.

Glaubitz, J. (2005). A user Friendly program to format diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes* **4**: 309–310.

Gomez, K.A. and Gomez. A.A. (1984). *Statistical Procedures for Agricultural Research*, 2nd edn., pp 680. John Wiley & Sons. New York.

Goudet, J., Raymond, M., de Meeiis, T. and Roussett, F. (1996). Testing Differentiation in Diploid Population. *Genetics* **144(4)**: 1933–1940.

Habtamu Fufa Aga and Kelbessa Urga Badada (1997). Nutritional and antinutritional characteristics of Anchote (*Coccinia abyssinica*). *Ethiop. J. Health Dev*. **11(2):** 163–168.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser*. **41**: 95–98.

Haussmann, B.I.G., Parzies, H. K., Presterl1, T., Susic, Z. and Miedaner, T. (2004). Plant genetic resources in crop improvement. *Plant Genetic Resources* **2(1):** 3–21.

Hawks, J., Hunley, K., Lee, S-H. and Wolpoff, M. (2000). Population Bottlenecks and Pleistocene Human Evolution. *Mol. Biol. Evol.* **17(1):** 2–22.

Holland, J.B. (2006). Estimating Genotypic Correlations and Their Standard Errors Using Multivariate Restricted Maximum Likelihood Estimation with SAS Proc MIXED. *CropSci*. **46**: 642–654.

Holsinger, E. and Weir, B.S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting $F_{ST.}$ *Nature Review* **10**: 639–650.

Holstein, N. (2012). Evolution, Biogeography, and Monographic Treatment of *Coccinia* (Cucurbitaceae). Ph.D. Dissertation, Munich University, Munich.

Holstein, N. and Renner S.S. (2011). A dated phylogeny and collection records reveal repeated biome shifts in the African genus Coccinia (Cucurbitaceae). *BMC Evolutionary Biology* **11**:28 http://www.biomedcentral.com/1471-2148/11/28.

Huang *et al*., (2009).The genome of the cucumber, Cucumis sativus L. *Nature Genetics* **41(12):** 1275–1283.

INRA (1999). *Bottleneck Version 1.2.02*. (16.II.99): Laboratoire de Modélisation et Biologie Evolutive.

Jakobsson, M. and Rosenberg, N.A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**(18): 1–6.

Kalinowski, S.T. (2005). Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity* **94**: 33–36.

Kalvandi, R., Hejazi, S.M.H., Atri, M., Mirza,M., Jamzad, Z. and Safikhan, K. (2012). Karyotype Analysis among 10 populations of *Thymuserio calyx* (Ronniger) Jalas species in Iran. *Annals of Biological Research* **3(8):** 3916–3925.

Karin, Z. (2002). Enset (Enset ventricosum (Welw.) Cheesm.) in subsistance farming systems in Ethiopia. Conference on International Agricultural Research for Development. Deutscher Tropentag, October 9–11.

Kazem, Y., Houshmand, S., Madani, B. and Martínez–Gómez, P. (2010). Karyotypic studies in Iranian wild almond species. *Caryologia* **63(2):** 117–123.

Kifle Dagne, and Heneen, W.K. (1992). The karyotype and nucleoli of *Guizotia abyssinica* (Compositae). *Hereditus* **117:** 73–83.

Kim, K.S., Ratcliffe, S.T., French, B.W., Liu, L., Sappington, T.W. (2008). Utility of EST-derived SSRs as population genetics markers in a beetle. *J Hered.* **99:** 112–124.

Kimura, M and Crow, J.F. (1964). The numbers of alleles that can be maintained in a finite population. *Genetics* **49**: 725–738.

Kisua, J., Mwikamba, K., Makobe, M. and Muigai, A. (2015). Genetic diversity of sweet and grain sorghum populations using phenotypic markers. *Int. J. Biosci* **6(9)**: 34–46.

Koffi, K.K., Gbotto, A.A., Malice, M., Dje, Y., Bertin, P., Baudoin, J-P., Bi, I.A.Z. (2008). Morphological and allozyme variation in a collection of Cucumeropsis mannii Naudin (Cucurbitaceae) from Coˆte d'Ivoire. *Biochemical Systematics and Ecology* **36:** 777–789.

Kuo, C-H. and Janzen, F.J. (2004). Genetic effects of a persistent bottleneck on a natural population of ornate box turtles (Terrapene ornata). *Conservation Genetics* **5:** 425–437.

Kuriyan, R., Rajendran, R., Bantwal, G., and Kurpad, A.V. (2008). Effect of Supplementation of *C. cordifolia* Extract on Newly Detected Diabetic Patients. *Diabetes Care* **31**: 216–220.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.

Lasalita-Zapico, F.C., Namocatcat, J.A., and Cariño-Turner, J.L. (2010). Genetic Diversity Analysis of Traditional Upland Rice Cultivars in Kihan, Malapatan, Sarangani Province, Philippines Using Morphometric Markers. *Philippine Journal of Science* **139 (2):** 177–180.

Levan, A., Fredga, K. and Sandberg, A. V. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* **52**: 201–220.

Litt, M. and Lutty, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet* **44**: 397–401.

Liu, K. and Muse, S.V. (2005). PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics* **21**:2128–2129.

Lu, J., Hu, X., Liu, J. and Wang, H. (2011). Genetic diversity and population structure of 151 *Cymbidium sinense* cultivars. *Journal of Horticulture and Forestry* **3(4):** 104–114.

Ma, Y., Islam-Faridi, M.N., Crane, C.F., Stelly, D.M., Price, H.J. and Byrne, D.H. (1996). A New Procedure to Prepare Slides of Metaphase Chromosomes of Roses. *Hortscience* **31(5):** 855–857.

Mahalanobis, P.C. (1936). Generalized distance in statistics, *Proceedings National Institute of Science, India*, **12**: 49–55.

Mao, W., Xu, S., Ye, S., Wang, H., Gao, X., Mo, X. and Gong, Y. (2014). Development of 19 transferable *Cucurbita pepo* EST-SSR markers for the study of population structure and genetic diversity in pumpkin (*Cucurbita Moschata*). *POJ* **7(5):** 345–352.

Martins, W.S., Lucas, D.C.S., Neves, K.F.S., and Bertioli, D.J. (2009). WebSat - A Web Software for MicroSatellite Marker Development. *Bioinformation* **3(6):** 282–283.

Meirmans, P.G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* **24:** 3223–3231.

Minitab, (1998). *Minitab Statistical Package for Windows*. Minitab Inc., Sales (USA).

Mohammedi, S.A. and Prasanna, B.M. (2003). Analysis of genetic diversity in crop plants-Salient statistical tools and considerations. *Cropsci*. **43**: 1235–1248.

Morin, P.A., Leduc, R.G., Archer, F.I., Martien, K.K., Huebinger, R., Bickham, J.W. and Taylor, B.L. (2009). *Molecular Ecology Resources* **9**: 498–504.

Mullis, K. B. and Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. *Methods in Enzymology* **155**: 335–350.

Munasinghe, M. A. A. K., Abeysena, C., Yaddehige, I. S., Vidanapathirana, T. and Piyumal, K. P. B. (2011). Blood Sugar Lowering Effect of *Coccinia grandis* (L.) J. Voigt: Path for a New Drug for Diabetes Mellitus. *Diabetes Research.* doi:10.1155/2011/978762.

Nei, M. (1973). The theory and estimation of genetic distance, pp. 45-54. **In**: *Genetic Structure of Populations*, (Morton N.E, ed). University Press of Hawaii, Honolulu.

Ngampongsai, S., Watanasit, A., Srisombun, S. and Srinives, P. (2009). Current Status of Mungbean and the Use of Mutation Breeding in Thailand. **In**: *Induced Plant Mutations in the Genomics Era*. pp.355-357 (Shu, Q.Y, ed.). Food and Agriculture Organization of the United Nations, Rome.

Novelli, V.M., Cristofani-Yaly, M., Bastianel, M., Palmieri, D.A. and Machado, M.A. (2013). Screening of genomic libraries. **In**: *Methods in Molecular Biology: Microsatellites Methods and protocols*, pp.17–24, (Kantartzi, S.K. ed). Human Press, Carbondale.

Oliveira, E. J., Padua, J.G., Zucchi, M.I., Vencovsky, R. and Vieira, M.L.C. (2006). Origin, evolution and distribution of microsatellite.*Genet. Mol. Biol.* **29(2):** 294–307.

Ostrander, E.A., Jong, P.M., Rine, J., Duyk, G. (1992). Construction of small- insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proc Natl Acad Sci USA.* **89:** 3419–3423.

Paetkau, D. and Strobeck, C. (1995). The molecular basis and evolutionary history of a microsatellite null allele in bears. *Mol Ecol*. **4**: 519–520.

Patterson, H.D. and Williams, E.R. (1976). A new class of resolvable incomplete block designs. *Biometrika* **63**: 83–92.

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

Peakall, R., Gilmore, S., Keys, W., Morgante, M. and Rafalski, A. (1998). Cross-species amplification of soybean (Glycine max) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology and Evolution* **15(10):** 1275–1287.

Purseglove, J.W. (1966). *Tropical crops, Dicotylendons I*. Longmans Green and Co. Ltd., Lond & Harlow, pp.100–138.

Ranade, S.S., Lin, Y-C., Zuccolo, A., de Peer, Y.V. and García-Gil, M.R. (2014). Comparative *in silico* analysis of EST-SSRs in angiosperm and gymnosperm tree genera. *BMC Plant Biology* **14**: 2–20.

Reddy, B.U. (2009). Cladistic analyses of a few members of Cucurbitaceae using *rbcL* nucleotide and amino acid sequences. *International Journal of Bioinformatics Research* **1(2):** 58–64.

Reinhard, F. and Admasu A. (1994). *Honeybee Flora of Ethiopia*. OED. Weikersheim: Margraf Verlag, Deutscher Entwicklungsdienst. The National Hetrbarium, Addis Ababa, pp.510.

Robinson, R.W. and Decker-Walters, D.S. (1999). *Cucurbits*. CAB International, Wallingford, Oxford, UK, pp.1–320.

Romero-Zarco, C. (1986). A new method for estimating karyotype asymmetry. *Taxon*, **35**:526–530.

Rosenberg, N.A. (2004). Distruct: a program for the graphical display of population structure. *Mol Ecol Notes*. **4**: 137–8.

Rousset, F. (2007). GENPOP v.4.0.10 for windows and Linus http://WWW.Cecill. Info/index.en.html.

Rozen, S., Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. **In**: *Bioinformatics methods and protocols*, pp. 365–86, (Misener, S. and Krawetz, S. eds). Humana Press, New Jersey.

Samuel, R., Balasubramaniam, S. and Morawetz, W. (1995). The Karyology of Some Cultivated Cucurbitaceae of Sri Lanka. *Cey. J. Sci. (Bio. Sci.)* **24(1):** 17–22.

Sargent, R.D. and Otto, S.P. (2004). A phylogenetic analysis of pollination mode and the evolution of dichogamy in angiosperms. *Evolutionary Ecology Research* **6**: 1183–1199.

SAS Institute (2002-2004). SAS Institute Inc. NC.

Schaefer, H. and Renner, S.S. (2011). Cucurbitaceae. **In**: *The families and genera of vascular plants*, Vol. 10. (Kubitzki, ed), Springer Heidelberg Dordrecht London New York.

Schaefer, H., Heibl, C., Renner, S.S., (2009). Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc. R. Soc. B* **276**: 843–851.

Sebastian, P.M. (2011). *Phylogenetics and biogeography of two clades of Cucurbitaceae: one New World, one Old World, and both with economically important species*. Ph.D. Dissertation, Munich University, Munich.

Sessions, S.K. (1996). Chromosomes: Molecular Cytogenetics. **In**: *Molecular Systematics*, 2nd edn., pp. 121–168, (Hillis, D.M., Moritz, C. and Mable, B.K. eds). Sinauer Association, Inc., Sunderland, Massachusetts, USA.

Singh, R.K. and Chaudhary, B.D. (1985). *Biometrical Methods in Quantitative Genetic Analysis*.Kalyani Publishers, New Delhi, pp.13-55.

Singh, S. and Parab, M. (2015). Fingerprinting intra-specific diversity among *coccinia grandis* landraces. *International Journal of Recent Scientific Researc*h **6(3):** 3025–3032.

Sivaraj, A., Jenifa, B.P., Kavitha, M., Inbasekar, P., Senthilkumar, B. and Panneerselvam, A. (2011). Antibacterial activity of *Coccinia grandis* leaf extract on selective bacterial strains. *Journal of Applied Pharmaceutical Science* **1(7):** 120–123.

Slatkin, M. (1995). Hitchhiking and associative overdominance at a microsatellite locus. *Molecular Biology and Evolution* **12(3):** 473–480.

Slatkin, M. (2008). Linkage disequilibrium-understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*. **9(6):** 477–485.

Stebbins, G.L. (1971). *Chromosomal evolution in higher plants*. Edward Arnold Publishers Ltd, London. UK, pp.10–216.

Sutar, S.P., Khedasana, K., Rao, S.R., Bhat, K.V. and Yadav, S.R. (2013). Chromosome counts in the genus *Cucumis* L. [Cucurbitaceae] from India. *Plant Sciences Feed* **3(7)**: 80–82.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Sudhir Kumar (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Method. *Molecular Biology and Evolution* **28**: 2731–2739.

Tautz, D. and Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.***12**: 4127–4138.

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res* **11**: 1441–1452.

Tesfaye Messele, Abebe Demissie, (1988). Root and tuber crops collecting expedition in southwestern Ethiopia. *PGRC/E ILCA Germplasm Newsletter* **18**: 17-23

Thompson, S.A., Weber, C.W., Berry, J.W. and Benis, W.P. (1978). Protein quality of buffalo gourd seed and seed fraction. *Nutrition Reports International* **18:** 515–520.

Tilahun Wondimu, Sentayehu Alamerew, Amsalu Ayana, and Weyessa Garedew (2014). Genetic diversity analysis among anchote (*Coccinia abyssinica*) accessions in western Ethiopia. *Int. J. Agric. Res*. **9(3):** 149–157.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., et al. (1995). AFLP-A new technique for DNA-fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.

Waminal, N.E. and Kim, H.H. (2015). FISH Karyotype Analysis of Four Wild Cucurbitaceae Species Using 5S and 45S rDNA Probes and the Emergence of New Polyploids in Trichosanthes kirilowii Maxim. *Korean J. Hortic. Sci. Technol*. **33(6):** 869–876.

Wang, J. H., Kang, J., Son, B., Kim, K. and Park, Y. (2011). Genetic Diversity in Watermelon Cultivars and Related Species Based on AFLPs and EST-SSRs. *Not Bot Horti Agrobo* **39(2):** 285–292.

Wang, Z., Yu, G., Shi, B., Wang, X., Qiang, H. and Gao, H. (2014). Development and Characterization of Simple Sequence Repeat (SSR) Markers Based on RNA-Sequencing of *Medicago sativa* and *In silico* Mapping onto the *M. truncatula* Genome. *PLoS ONE* **9(3):** e92029. doi:10.1371/journal.pone.0092029.

Ward, J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of American Statistical Association* **58**: 236–244.

Weir, B.S. and Cockerham, C.C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). *DNA Fingerprinting in Plants Principles, Methods, and Applications*, 2nd edn. Taylor & Francis Group, CRC Press is an imprint of Taylor & Francis Group, Boca Raton London New York Singapore, pp. 444.

Wen, W., Wang, H., Xia, Z., Zou, M., Lu, C. and Wang, W. (2010). Development of EST-SSR and genomic-SSR markers to assess genetic diversity in *Jatropha curcas* L. *BMC Research Notes* **3**: 42.

Westphal, E. (1974). *Pulses in Ethiopia, their taxonomy and agricultural significance*. Doctoral thesis, Centre for Agricultural Publishing and Documentation, Wageningen.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V., (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers, Nucleic *Acids Res.* **18**: 6231–6235.

Wolde Michael (1987). *A Glossary of Ethiopian Plant Names*, 4th edition. Addis Ababa, Ethiopia.

Yeh, F.C., Yang, R. and Boyle, T. (1999). *Popgene version* 1.31.

Zane, L., Bargelloni, L, Patarnello, T (2002): Strategies for microsatellite isolation: a review. *Mol Ecol* **11(1):** 1–16.

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

/

10. APPENDICES

Appendix Table 1.Mean performance of morphological data of anchote accessions from different woredas

Woreda	VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINR	RL	RD	RLDR
Al	4.00	89.35	22.34	18.56	7.75	7.44	1.05	8.53	3.33	2.85	9.44	5.46	2.01
An	1.79	73.49	18.06	19.43	7.38	7.10	1.04	9.18	3.19	3.14	9.75	4.22	2.32
Ar	1.69	54.44	13.38	20.05	7.24	7.20	1.01	9.68	3.16	3.79	9.03	3.74	2.46
Ay	2.17	75.41	18.98	18.10	6.96	6.86	1.02	8.66	3.60	2.47	11.95	3.77	3.22
Ch	2.78	106.52	**26.45**	18.88	8.19	8.37	0.98	11.30	3.43	3.66	11.31	4.50	2.66
Dd	1.54	56.22	14.05	19.90	7.82	7.85	1.00	10.21	3.06	4.29	8.69	4.13	2.13
De	2.54	87.80	19.95	20.48	7.52	7.67	0.98	10.77	3.14	3.85	9.86	4.23	2.51
Dg	3.00	80.29	19.12	**21.53**	7.51	7.54	1.00	10.13	3.66	3.50	9.45	4.19	2.31
Dh	2.33	90.29	**22.57**	16.43	8.01	8.50	0.94	10.26	3.42	3.61	7.19	5.02	1.50
Ds	1.92	73.07	17.97	18.56	7.78	7.91	0.99	9.06	3.84	2.56	12.20	4.18	3.18
Dw	2.75	89.74	22.35	20.01	7.83	7.82	1.01	9.89	3.35	3.21	10.58	4.17	2.56
Gc	1.95	70.43	17.50	19.79	7.83	7.94	0.99	10.68	3.08	4.11	8.94	4.27	2.20
Gg	2.34	71.27	18.24	20.64	7.35	7.50	0.99	10.02	3.20	3.98	8.88	4.23	2.15
Gl	2.13	59.27	14.82	**21.43**	6.19	6.27	0.99	7.61	3.15	2.56	12.23	3.31	3.80
Gm	2.42	68.04	17.01	20.90	7.55	7.48	1.01	9.92	3.93	2.75	9.91	3.93	2.54
Go	2.18	79.38	17.59	16.99	7.99	8.04	1.00	10.57	3.08	4.05	9.07	4.58	2.05
Gr	2.07	80.77	20.19	18.10	7.96	8.06	0.99	10.92	2.76	4.47	7.82	4.81	1.64
Gu	1.96	75.28	18.64	18.72	7.57	7.77	0.98	10.32	3.69	3.27	9.68	4.44	2.37
Hr	2.50	89.68	22.29	18.57	8.22	8.40	0.98	10.39	3.08	3.54	10.42	4.54	2.38
La	2.19	61.55	15.11	20.40	7.62	7.61	1.00	9.81	3.80	2.81	9.80	3.82	2.75
Ld	2.08	71.68	18.19	16.99	7.48	7.64	0.99	9.98	3.12	4.00	9.09	4.28	2.14
Mt	2.50	87.90	21.98	18.36	7.81	8.00	0.98	9.81	3.39	3.10	9.99	4.13	2.60
Nj	1.75	71.49	18.52	19.60	7.74	7.57	1.03	9.56	3.56	2.93	10.10	4.07	2.73
Nk	2.00	45.57	11.33	**22.93**	6.35	6.33	1.01	8.99	2.63	3.87	8.60	3.67	2.41
Sc	2.33	66.45	17.22	16.92	8.19	8.16	1.01	10.35	3.63	3.28	9.73	4.25	2.41
Sh	1.50	44.34	13.52	12.61	7.26	7.18	1.01	10.87	2.69	4.54	9.59	3.82	2.87
Ss	2.00	48.85	12.04	18.22	7.74	7.78	1.00	10.11	2.92	4.07	9.26	3.91	2.36
Sy	2.50	68.66	16.30	19.98	7.37	7.49	0.98	9.31	3.16	3.15	9.98	3.81	2.74
Wt	1.91	64.63	15.35	19.83	7.14	6.99	1.02	9.66	3.18	3.65	10.28	3.68	2.92
Yy	2.50	99.10	**24.56**	18.24	**8.55**	8.77	0.98	11.10	3.15	3.90	10.11	4.66	2.26
Average	2.24	73.37	18.19	19.04	7.60	7.64	1.00	9.92	3.28	3.50	9.76	4.19	2.47
Stdev	0.49	15.32	3.64	1.93	0.50	0.57	0.02	0.82	0.33	0.58	1.13	0.44	0.46
Max	4.00	106.52	26.5	22.93	8.55	8.77	1.05	11.30	3.93	4.54	12.23	5.46	3.80
Min	1.50	44.34	11.3	12.61	6.19	6.27	0.94	7.61	2.63	2.47	7.19	3.31	1.50
CV	24	39	51	4	27	14	15	8	19	40	45	23	24
CD	0.99	30.03	6.46	4.60	0.88	0.898	0.066	1.64	0.82	0.894	1.49	0.782	0.57
Group	3	3	3	3	4	4	2	3	2	3	5	3	5

Woreda's code is given in table 1, DE = days to 50% emergence, VN = vine number, AvRwt = Average root weight, YLD = yield, DMC= dry matter content, LL = leaf length, LD = leaf diameter, LLDR = leaf length to its diameter ratio, PEL = petiole length, INL = inter-node length, PINLR = petiole to inter-node lengths ratio, max = maximum, min=minimum, CV = coefficient of variation, CD = pair group critical difference (Duncan's Multiple Range Test), D.G = Duncan's group.

Appendix Table 2. Mean performance of morphological traits grown from seed by population

Population	VN	AVWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINR	RL	RD	RLDR
Ay-La	2.20	68.10	16.80	19.40	7.30	7.30	1.01	9.28	3.70	2.70	10.80	3.80	3.00
Ds-Dw	2.30	79.70	19.70	19.10	7.80	7.90	0.99	9.39	3.60	2.80	11.60	4.20	2.90
Sy-An	2.30	68.90	16.40	19.90	7.40	7.40	0.99	9.24	3.20	3.10	9.91	3.90	2.60
Gm-Nj	2.10	70.80	18.00	20.20	7.70	7.50	1.03	9.71	3.70	2.90	9.97	4.00	2.60
Mt-Al	2.90	88.40	**22.10**	18.40	7.80	7.90	1.00	9.48	3.40	3.00	9.81	4.50	2.40
Hr-Yy	2.50	94.40	**23.40**	18.40	8.40	8.60	0.98	10.80	3.10	3.70	10.30	4.60	2.30
Ch-De	2.70	96.00	**22.80**	19.80	7.80	8.00	0.98	11.00	3.30	3.70	10.60	4.40	2.60
Gc-Dh	2.00	71.20	17.70	19.50	7.90	8.00	0.99	10.60	3.10	4.10	8.77	4.30	2.10
Gu,Go, Mn	2.10	77.60	18.00	17.80	7.80	7.90	0.98	10.40	3.40	3.70	9.34	4.50	2.20
Gr,Sc,Sh	2.10	72.30	18.60	17.20	8.00	8.10	1.00	10.70	3.10	4.00	8.70	4.50	2.00
Dd-Dc	1.60	56.20	14.00	20.00	7.80	7.80	0.99	10.20	3.10	4.30	8.72	4.10	2.10
Ac-Bt	2.60	74.90	18.40	21.10	7.40	7.50	0.98	10.10	3.40	3.70	9.14	4.20	2.20
Ac-Bt	1.90	60.30	14.40	20.20	7.10	7.00	1.01	9.62	3.10	3.70	9.90	3.70	2.80
Ss-Wt	1.80	58.20	14.30	19.40	7.30	7.30	1.00	9.76	3.20	3.80	9.02	3.90	2.40
Ar-Ld	2.10	59.30	14.80	21.50	6.20	6.30	0.99	7.62	3.10	2.60	12.30	3.30	3.80
Mean	2.20	71.90	17.60	19.40	7.60	7.60	1.01	10.00	3.30	3.60	9.60	4.10	2.40
St.Dev	0.9	38.7	8.66	5.44	1.1	1.2	0.1	2.02	1.3	1.7	2.4	1	0.8
Max	2.9	96	23.4	21.5	8.4	8.6	1.03	11	3.7	4.3	12.3	4.6	3.8
Min	1.6	56.2	14	17.2	6.2	6.3	0.98	7.62	3.1	2.6	8.7	3.3	2
CV	40	48	41	27	13	13	7	18	28	28	18	21	26
CD	0.59	18.59	3.95	2.85	0.529	0.54	0.04	1.0	0.50	0.55	0.92	0.48	0.35
Group	3	3	4	2	4	5	2	3	2	3	5	4	5

DE=days to 50% emergence, VN=vine number, AvRwt =Average root weight, YLD= yield, DMC=dry matter content, LL=leaf length, LD=leaf diameter, LLDR=leaf length to its diameter ratio, PEL=petiole length, INL=inter-node length, PINLR=petiole to inter-node lengths ratio, max=maximum, min=minimum, CV=Coefficient of variation, CD= pair group critical difference (Duncan's Multiple Range Test), D.G= Duncan's group.

Appendix Table 3. Combined mean performances of 13 agro-morphological traits of cultivated 149 anchote accessions grown from seed at three different environments with two replications by accession

Accn	VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINLR	RL	RD	RLDR
1	2.75	67.82	16.95	21.40	7.62	7.75	0.98	10.03	3.61	2.96	10.04	4.72	2.34
2	2.00	70.53	17.63	21.18	7.66	7.68	1.00	9.38	3.71	2.78	9.72	3.70	2.73
4	1.75	57.59	13.48	17.90	7.96	7.68	1.04	9.74	3.61	3.00	9.26	3.56	2.86
5	2.25	50.25	12.38	21.12	7.24	7.34	0.99	10.07	4.28	2.48	10.19	3.29	3.08
6	2.00	59.64	14.91	20.86	5.91	6.05	0.98	7.40	3.23	2.46	11.91	3.46	3.50
7	2.25	58.90	14.73	22.00	6.47	6.49	1.00	7.82	3.07	2.66	12.55	3.16	4.10
8	2.25	72.49	17.65	17.36	6.94	6.62	1.05	8.55	3.61	2.56	12.11	4.03	3.12
9	1.67	86.38	23.53	18.30	6.79	6.89	0.99	7.58	3.59	2.34	13.08	3.74	3.50
10	2.25	64.12	15.71	18.26	7.02	7.09	1.00	8.79	3.55	2.63	10.87	3.50	3.16
11	2.50	78.63	19.03	18.46	7.09	6.85	1.04	9.73	3.66	2.36	11.72	3.79	3.11
12	2.00	93.81	22.84	17.83	8.59	8.78	0.98	10.95	4.53	2.54	12.86	3.77	3.41
13	2.00	49.15	12.29	19.89	7.07	6.97	1.02	7.95	2.98	3.03	11.82	4.35	3.36
14	1.75	76.26	18.78	17.97	7.68	7.99	0.97	8.27	4.00	2.10	11.93	4.41	2.77
17	3.00	89.45	22.19	20.92	7.86	7.99	0.99	9.32	3.39	2.94	10.95	4.18	2.66
19	2.50	90.02	22.51	19.09	7.80	7.65	1.02	10.46	3.31	3.47	10.20	4.16	2.45
21	3.25	89.03	16.23	19.06	7.32	7.31	1.00	10.47	3.20	3.46	9.81	3.83	2.68
22	2.00	92.25	23.06	18.12	7.59	7.59	1.01	10.87	2.72	4.11	11.53	4.12	3.06
23	1.75	60.96	15.24	18.86	7.22	7.39	0.98	8.74	2.89	3.32	9.05	4.08	2.21
24	2.50	50.30	12.58	18.42	7.77	7.86	0.99	8.89	3.82	2.45	9.23	3.14	3.01
26	3.00	71.18	17.79	23.64	7.99	7.93	1.01	8.77	3.14	2.86	10.16	4.09	2.56
27	2.25	83.85	20.47	19.17	7.60	7.21	1.06	9.95	3.11	3.36	9.93	4.35	2.29
28	1.33	63.13	15.64	19.68	7.15	6.98	1.02	8.40	3.27	2.92	9.56	4.08	2.35
29	2.25	66.83	16.71	21.17	7.35	7.41	0.99	9.49	3.04	3.37	9.83	3.96	2.70
30	2.75	50.07	12.52	20.60	6.33	6.95	0.91	7.92	3.31	2.47	10.28	3.48	2.95
31	2.75	76.75	19.19	20.56	7.72	7.72	1.00	11.57	4.63	2.76	9.70	4.05	2.45
32	2.25	48.12	12.03	20.53	7.17	6.95	1.03	8.79	3.53	2.73	9.43	3.73	2.51
33	2.25	79.26	19.81	21.60	7.76	7.78	1.00	9.40	3.62	2.77	10.61	4.00	2.66
35	1.75	63.15	15.79	20.07	7.78	7.31	1.07	9.54	2.89	3.39	10.81	3.49	3.16
36	1.75	73.84	18.29	20.89	7.45	7.49	1.00	8.89	4.05	2.50	9.78	4.63	2.15
37	2.00	79.75	22.68	18.95	7.30	7.35	1.00	8.88	3.49	2.78	9.47	4.42	2.76
38	1.50	69.22	17.30	18.50	8.42	8.14	1.05	10.92	3.81	3.03	10.34	3.73	2.85
39	2.50	114.8	28.70	19.28	8.84	9.12	0.97	12.22	3.95	3.40	9.31	4.58	2.05
41	2.50	61.00	15.25	17.44	6.78	6.88	0.99	7.39	2.83	2.80	10.66	3.68	3.14
45	4.00	89.35	22.34	18.56	7.75	7.44	1.05	8.53	3.33	2.85	9.44	5.46	2.01
50	2.50	99.54	24.88	19.78	8.63	8.79	0.98	10.47	2.88	3.68	10.20	4.82	2.11
51	2.50	102.2	25.55	18.94	8.12	8.56	0.95	10.38	3.36	3.51	11.23	4.45	2.51
52	2.25	82.70	20.39	17.83	8.34	8.21	1.02	10.54	3.30	3.05	10.80	4.10	2.91
53	2.75	74.25	18.35	17.74	7.79	8.05	0.96	10.18	2.76	3.93	9.43	4.79	1.99
54	2.50	101.1	25.28	18.00	8.68	8.29	1.05	10.72	3.37	3.71	10.79	4.50	2.46
55	2.75	86.08	20.86	19.71	8.38	8.75	0.96	10.92	3.19	3.37	9.97	4.18	2.48
56	2.00	111.7	27.92	15.20	8.23	9.03	0.92	10.40	3.09	3.81	9.94	5.40	1.88
57	2.75	97.48	24.19	20.06	8.90	9.02	0.99	12.34	2.93	4.71	9.72	4.55	2.20
58	2.50	66.47	15.73	20.89	7.45	7.46	1.00	10.29	3.39	3.44	12.77	4.01	3.04
59	2.25	99.74	24.93	16.76	9.18	9.43	0.97	12.10	4.49	2.97	10.98	4.29	2.68
60	3.25	110.7	**27.51**	18.22	7.72	7.76	1.00	11.14	3.78	3.16	9.91	5.60	1.96
61	3.00	107.1	26.78	17.83	7.32	7.24	1.01	10.48	2.01	4.77	10.26	4.16	2.57
62	2.67	129.5	**32.37**	19.41	8.34	8.87	0.95	11.18	3.50	3.82	12.33	3.93	3.12
65	3.00	125.6	**31.40**	20.15	9.12	9.47	0.96	12.63	3.40	3.77	11.59	4.98	2.61
67	2.75	100.4	25.10	20.19	8.02	8.09	0.99	12.01	3.18	4.11	11.36	4.66	2.46
68	3.50	114.6	17.28	16.92	6.81	7.16	0.95	8.57	3.43	2.83	10.98	4.38	2.66
69	2.25	85.82	20.81	18.94	7.32	7.61	0.96	11.24	2.84	4.37	9.01	4.97	1.83
70	2.00	70.39	17.60	20.09	7.96	7.88	1.01	10.86	3.18	3.77	8.98	3.84	2.38
71	1.75	66.04	16.51	26.43	7.38	7.47	0.99	10.89	2.90	4.55	9.79	3.16	3.53
72	3.00	89.51	22.38	20.28	7.60	7.82	0.97	11.05	3.29	3.44	9.06	4.37	2.17

Accn	VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINLR	RL	RD	RLDR
73	2.25	80.58	20.14	19.44	7.25	7.73	0.94	10.97	2.56	4.96	8.16	4.51	1.87
74	2.33	50.95	12.74	23.52	7.26	6.98	1.05	10.81	2.20	5.54	8.14	3.91	2.17
75	1.50	65.50	16.38	17.93	8.61	8.29	1.05	10.82	3.96	3.26	9.20	4.55	2.02
76	1.75	59.89	14.49	17.42	8.29	8.02	1.04	9.92	3.00	4.16	8.88	3.85	2.55
77	2.00	108.6	26.83	15.47	8.43	8.93	0.95	11.69	3.40	3.75	10.12	4.86	2.22
79	2.00	63.21	15.71	20.51	7.12	7.92	0.90	10.42	3.24	3.82	9.04	4.11	2.29
80	2.25	65.47	16.37	22.46	7.67	7.71	1.00	10.51	3.15	4.04	8.23	4.02	2.15
81	1.50	69.25	17.31	21.60	8.02	7.95	1.02	10.30	3.11	3.31	9.77	4.38	2.31
87	2.33	90.29	22.57	16.43	8.01	8.50	0.94	10.26	3.42	3.61	7.19	5.02	1.50
88	1.75	71.70	17.66	15.70	7.82	8.23	0.95	10.64	2.71	4.51	10.24	4.67	2.47
89	2.00	81.58	19.81	16.47	7.72	7.76	1.00	9.72	3.91	2.76	8.39	4.66	1.81
90	2.00	71.88	17.87	16.21	7.43	7.84	0.95	12.10	4.03	3.68	9.81	4.26	2.37
91	2.00	55.28	13.82	18.53	6.87	6.59	1.04	7.03	3.51	2.15	8.94	4.23	2.17
92	2.00	94.89	23.72	22.20	7.72	8.14	0.95	10.43	4.14	2.87	10.71	4.10	2.69
93	2.00	76.35	18.97	23.19	7.87	8.08	0.98	12.00	3.85	3.66	9.97	4.69	2.68
94	2.25	58.44	14.61	17.84	7.78	8.10	0.96	10.11	3.27	3.63	9.05	3.96	2.46
95	2.75	66.85	15.93	13.59	7.68	7.65	1.00	11.03	2.38	4.95	9.29	4.32	2.16
96	2.25	115.4	22.04	17.04	9.45	9.44	1.00	11.42	3.52	4.20	10.27	4.64	2.33
97	2.00	63.05	15.56	18.56	7.72	7.59	1.02	10.87	2.98	4.13	9.76	4.23	2.31
98	2.33	104.0	26.01	19.83	8.46	8.65	0.98	10.96	3.00	3.85	8.40	4.16	1.91
99	2.33	60.80	15.20	15.29	7.51	7.32	1.02	9.36	2.46	4.30	7.95	4.46	1.81
100	2.00	73.08	18.27	20.90	7.52	7.81	0.96	10.63	2.60	4.63	8.11	4.76	1.71
101	2.00	64.74	16.18	17.03	7.09	7.40	0.96	10.59	2.53	5.02	6.40	4.90	1.35
102	2.00	86.43	21.61	18.84	7.70	7.53	1.03	11.23	2.59	4.52	8.78	5.14	1.67
103	1.75	95.52	23.88	16.72	9.45	9.64	0.99	12.77	3.39	4.47	7.28	5.46	1.38
104	2.00	72.02	18.01	17.47	6.92	6.88	1.01	8.25	2.73	3.60	7.55	4.74	1.59
106	2.00	104.2	18.08	19.13	8.37	8.42	1.00	11.93	3.02	4.19	8.70	5.52	1.60
107	2.00	75.71	18.93	15.33	8.03	8.17	0.99	10.36	3.63	3.63	8.89	4.63	1.90
108	1.50	70.32	15.55	18.42	7.14	7.12	1.01	10.36	2.82	4.46	10.80	4.41	2.48
109	1.67	57.40	14.23	19.37	7.95	8.15	0.97	10.94	3.08	4.54	7.45	4.74	1.62
110	1.00	57.52	14.62	17.52	7.72	7.94	0.98	10.04	2.83	4.91	8.62	3.61	2.39
111	1.33	51.48	14.18	23.58	8.12	8.12	1.00	11.05	3.26	4.09	8.68	4.14	2.08
112	1.75	45.85	12.02	24.20	7.99	8.08	0.99	8.02	3.15	3.46	8.48	3.85	2.21
113	1.75	54.80	13.70	17.66	7.58	7.41	1.03	9.52	3.17	4.42	8.36	4.24	1.98
114	1.33	67.28	16.82	18.36	8.56	8.56	1.01	11.22	3.47	4.15	8.76	4.03	2.16
115	2.00	45.12	11.28	20.09	7.52	7.45	1.01	10.55	2.69	4.32	8.38	4.04	2.08
116	2.00	61.50	17.82	13.50	8.46	8.45	1.01	11.25	3.59	3.98	9.63	3.51	2.73
117	2.33	67.84	16.96	16.89	7.53	7.26	1.03	10.39	4.00	2.60	8.78	4.46	1.93
118	2.50	59.93	14.98	18.94	7.93	8.19	0.97	9.80	3.12	3.41	9.27	5.17	1.98
119	2.50	76.51	19.13	18.33	8.82	8.72	1.01	9.96	3.82	3.14	11.22	3.84	2.98
120	1.50	44.34	13.52	12.61	7.26	7.18	1.01	10.87	2.69	4.54	9.59	3.82	2.87
123	1.25	41.64	10.91	18.50	6.70	6.64	1.01	8.75	3.00	3.34	9.47	3.39	2.95
124	2.50	53.65	13.41	20.15	7.30	7.32	1.00	9.09	2.85	4.18	9.02	3.48	2.56
125	1.75	43.41	9.73	21.42	7.18	6.98	1.04	9.12	2.78	4.07	9.64	3.51	2.80
126	1.67	41.58	10.39	19.63	6.56	6.54	1.01	9.28	3.15	3.58	9.13	3.49	2.67
127	1.00	49.96	12.49	18.43	6.65	6.84	0.98	9.97	3.22	3.97	9.43	3.65	2.69
128	2.25	53.69	13.16	24.47	7.54	7.56	1.00	10.10	3.28	4.30	8.37	3.86	2.17
129	2.25	44.26	11.01	20.59	7.06	6.71	1.05	9.12	3.66	3.13	9.00	3.56	2.55
130	1.50	60.55	15.14	17.81	6.99	7.33	0.97	9.17	2.67	4.04	8.94	4.08	2.23
131	1.50	45.75	11.06	18.36	8.59	7.95	1.07	9.32	3.13	3.96	9.50	3.38	2.82
132	1.60	50.26	12.57	19.60	7.22	7.33	0.98	10.49	3.29	3.91	8.14	3.79	2.21
133	1.75	68.80	18.16	19.06	7.55	7.60	0.99	10.97	3.56	3.94	8.83	4.13	2.13
134	1.50	39.88	9.21	20.39	6.83	6.64	1.02	9.03	3.33	3.23	8.52	3.35	2.55
135	1.80	70.10	17.53	19.06	7.61	7.44	1.03	9.99	2.86	3.88	9.27	4.01	2.34
136	1.33	47.44	12.85	20.74	7.61	7.66	0.99	9.88	3.09	3.85	8.89	3.72	2.39
137	1.50	66.87	16.84	21.26	7.25	7.13	1.02	9.13	3.20	3.63	9.77	3.83	2.58
138	1.50	85.67	17.69	21.37	7.63	7.67	1.00	9.82	3.25	3.69	9.28	3.92	2.44

Accn	VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINLR	RL	RD	RLDR
139	1.50	62.05	15.51	20.97	6.88	7.00	0.98	9.22	3.18	3.04	8.04	4.20	1.92
140	1.25	52.59	15.52	20.86	6.83	6.46	1.06	9.56	2.95	4.06	10.34	3.81	2.70
141	2.67	66.97	13.48	20.24	8.01	8.06	0.99	11.17	3.50	4.08	8.45	3.99	2.12
142	1.75	43.58	10.89	18.12	6.84	7.04	0.98	10.36	3.22	3.95	8.59	3.60	2.43
143	1.75	69.27	17.32	20.02	7.89	7.68	1.06	10.30	2.70	4.53	7.98	4.09	1.96
144	2.25	79.70	19.93	19.15	7.10	7.06	1.01	10.19	3.55	4.14	10.04	4.38	2.31
145	2.00	50.31	12.46	21.44	7.22	7.67	0.95	10.01	3.17	3.94	8.90	3.79	2.38
146	2.33	69.88	22.19	22.69	7.42	7.56	0.98	8.58	3.68	3.17	8.61	4.16	2.23
147	2.75	86.45	21.61	19.96	7.87	7.90	1.00	10.35	2.89	4.34	9.83	4.90	2.06
148	2.50	84.82	20.44	17.76	7.31	7.55	0.97	11.00	3.30	3.79	8.06	4.06	2.04
149	2.00	77.09	19.27	21.71	7.42	7.86	0.95	10.72	3.05	4.56	8.02	4.25	1.90
150	2.75	61.73	13.63	23.14	7.20	7.25	0.99	9.57	3.51	3.39	8.68	4.33	2.07
151	2.75	62.19	17.27	19.87	6.74	6.98	0.97	9.43	2.96	4.00	9.79	4.13	2.43
152	2.50	72.54	18.40	19.21	7.27	7.45	0.98	9.72	4.70	2.51	9.92	3.71	2.74
156	4.00	113.9	28.46	19.82	8.22	8.03	1.03	10.87	3.10	4.17	9.15	4.56	2.04
157	2.75	74.43	18.61	28.12	6.97	7.06	1.00	8.40	3.13	3.36	8.63	4.17	2.11
158	2.50	69.60	17.40	23.15	7.99	8.37	0.96	11.45	3.28	4.07	9.85	4.49	2.22
159	3.25	66.17	16.54	19.59	6.65	6.54	1.02	9.25	2.86	3.86	9.57	4.17	2.39
160	3.50	110.7	19.79	19.60	7.91	8.01	0.99	11.62	3.57	3.93	9.81	4.76	2.06
161	3.00	82.32	20.58	21.51	7.65	7.59	1.01	10.05	4.83	3.04	9.64	4.18	2.29
162	2.50	52.72	13.18	21.22	7.43	7.26	1.02	9.65	3.83	3.04	9.03	3.51	2.66
164	2.75	74.81	18.70	16.70	7.10	7.53	0.94	10.16	2.49	4.56	9.87	4.75	2.09
167	2.00	95.17	23.79	13.73	8.41	8.15	1.04	10.29	3.96	3.52	9.90	5.07	1.97
168	2.00	56.20	14.05	17.76	7.75	7.43	1.05	9.66	2.57	4.74	8.64	3.91	2.21
169	1.75	65.49	16.37	21.33	7.34	7.46	0.99	9.80	3.55	3.41	8.68	4.32	2.00
170	2.25	50.03	12.28	15.84	7.06	7.46	0.95	10.44	3.33	3.81	8.70	3.77	2.30
171	1.75	88.40	23.95	16.56	7.21	7.82	0.94	9.54	2.82	3.97	8.72	3.83	2.25
172	1.50	55.30	13.82	18.20	6.21	6.18	1.01	8.90	2.43	4.14	10.27	3.67	2.79
173	2.00	59.25	14.32	21.53	7.24	7.13	1.02	9.92	2.99	3.50	9.34	3.64	2.60
174	1.75	68.58	11.87	18.59	7.68	6.98	1.10	9.89	2.86	4.29	8.97	3.75	2.42
175	2.00	67.62	14.49	16.32	7.30	6.95	1.05	10.47	3.66	3.61	9.81	3.96	2.93
176	2.75	46.86	13.41	21.79	7.02	7.08	0.99	9.00	3.70	2.75	10.62	3.20	3.39
177	1.75	65.18	16.30	26.06	7.33	7.03	1.04	9.19	3.11	3.50	10.63	3.33	3.43
178	1.50	69.58	17.40	18.86	7.21	7.68	0.94	10.53	3.57	3.74	12.73	3.65	3.48
179	2.00	84.69	21.17	17.27	7.14	6.91	1.04	9.36	3.14	3.63	9.87	4.23	2.35
180	1.75	42.38	10.59	19.61	6.21	6.28	0.99	8.95	2.47	4.20	8.49	3.22	2.69
181	2.25	48.76	12.06	26.24	6.49	6.37	1.02	9.03	2.79	3.53	8.70	4.11	2.13
182	2.00	48.85	12.04	18.22	7.74	7.78	1.00	10.11	2.92	4.07	9.26	3.91	2.36
Mean	2.16	72.04	17.7	19.39	7.59	7.634	0.99	10.03	3.266	3.626	9.61	4.142	2.434
max	**4.00**	**129.5**	**32.37**	**28.12**	**9.45**	**9.64**	**1.08**	**12.77**	**4.83**	**5.54**	**13.1**	**5.62**	**4.10**
min	**1.00**	**39.88**	**9.21**	**12.62**	**5.90**	**6.06**	**0.90**	**7.03**	**2.00**	**2.10**	**6.40**	**3.16**	**1.36**
Std	0.54	19.64	4.67	2.47	0.64	0.69	0.03	1.11	0.49	0.69	1.20	0.52	0.48
CV	40	48	41	27	13	13	7	18	28	28	18	21	26
CD	1.65	52.9	11.27	88.12	1.51	1.544	0.11	2.84	1.42	1.55	2.61	1.36	0.99
D.G	2	2	3	2	3	3	2	3	3	3	3	2	2

ACCN= accession number, DE= days to 50% emergence, VN=vine number, AvRwt = Average root weight, YLD= yield, DMC= dry matter content, LL= leaf length, LD= leaf diameter, LLDR= leaf length to its diameter ratio, PEL= petiole length, INL=inter-node length, PINLR= petiole to inter-node lengths ratio, max= maximum, min=minimum, CV=Coefficient of variation, CD= pair group critical difference (Duncan's Multiple Range Test), D.G= Duncan's group.

Appendix Table 4. Separated and combined descriptive statistics of 13 agro-morphological traits of cultivated anchote accessions grown from seed at three different environments with two replications at each environment

Environment*		VN	AVRWT	YLD	DMC	LL	LD	LLDR	PEL	INL	PINR	RL	RD	RLDR
1	Mean	2.128a	66.937a	16.461a	18.115a	7.215a	7.491a	0.967a	10.26a	2.138a	5.243a	9.834a	4.112a	2.47a
	Range	5.00	249.46	41.99	55.58	6.38	7.23	0.52	10.32	3.90	9.72	10.16	5.97	4.55
	Max.	6.00	263.50	46.00	58.76	10.63	11.63	1.14	15.20	4.90	10.00	16.10	8.17	5.52
	Min.	1.00	14.04	4.00	3.18	4.25	4.40	0.62	4.88	1.00	0.28	5.94	2.20	0.97
	Std	1.04	30.19	7.0	5.29	1.12	1.23	0.08	2.05	0.71	1.68	1.94	0.83	0.64
2	Mean	2.202a	83.650b	20.513b	19.650b	7.851b	7.797b	1.009b	10.09a	3.779b	2.865b	11.027b	4.501b	2.57a
	Range	4.00	228.21	46.99	64.32	5.65	5.52	0.48	9.46	5.58	4.88	12.66	8.52	4.64
	Max.	5.00	251.67	52.86	70.39	10.70	10.70	1.28	14.50	7.32	6.15	18.18	10.90	5.86
	Min.	1.00	23.46	5.87	6.07	5.05	5.18	0.80	5.04	1.74	1.27	5.52	2.38	1.21
	Std	0.79	38.74	9.17	5.36	0.91	0.92	0.07	1.66	1.01	0.90	2.04	1.0	0.74
3	Mean	-	65.048a	15.785a	20.490b	7.663c	7.611b	1.010b	9.746b	3.931c	2.719b	7.924c	3.797c	2.27b
	Range	-	382.21	48.21	59.31	6.48	7.20	.66	10.40	9.13	9.25	9.94	8.66	4.95
	Max.	-	387.58	52.05	59.51	11.38	12.00	1.40	15.00	10.13	10.00	14.70	10.70	5.60
	Min.	-	5.37	3.83	0.20	4.90	4.80	0.74	4.60	1.00	0.75	4.76	2.04	0.65
	std	-	43.45	8.92	5.42	1.23	1.26	0.07	2.29	1.38	1.03	2.16	1.11	0.88
Total	Mean	2.1667	71.8523	17.5880	19.4164	7.5870	7.6370	0.9965	10.032	3.2786	3.6125	9.6047	4.1390	2.438
	Std.	0.037	8.360	2.088	0.983	0.267	0.126	0.020	0.214	0.812	1.157	1.278	0.288	0.122
	Range	5.00	382.21	49.03	70.19	7.13	7.60	0.78	10.60	9.13	9.72	13.42	8.86	5.20
	Max	6.00	387.58	52.86	70.39	11.38	12.00	1.40	15.20	10.13	10.00	18.18	10.90	5.86
	Min	1.00	5.37	3.83	0.20	4.25	4.40	0.62	4.60	1.00	0.28	4.76	2.04	0.65
	CV	40	48	41	27	13	13	7	18	28	28	18	21	26
	CD	0.143	5.88	1.252	0.903	0.168	0.171	0.013	0.315	0.157	0.172	0.290	0.151	0.111
	D.G	1	2	2	2	3	2	2	2	3	2	3	3	2

*Envt1= APPRC year1, Envt2 = HARC year 1, Envt3 = HARC year2
CD= pair group critical difference (Duncan's Multiple Range Test), CV=Coefficient of variation, D.G= Duncan's group.

Appendix Table 5. Combined mean table for agro-morphological traits of cultivated anchote planted from root (irrigation & main season) at Holeta Agricultural research center

ACCN	LL	LD	LLDR	PEL	INL	PINLR	FRL	FRD	FRLDR	PEDL	FRAVWT	NSPFR
1	9.4	9.72	0.97	13.37	4.42	3.38	6	5	1	4	73	129
2	9.05	9.35	0.97	14.2	4.37	3.49	6	4	2	3	86	141
4	9.05	9.2	0.98	16.15	4.67	3.84	6	5	1	3	65	-
5	8.83	9.16	0.97	14.32	4.25	3.67	7	4	2	5	72	137
6	7.78	8.49	0.92	14.02	3.62	4.27	-	-	-	-	-	-
7	8.72	9.41	0.93	14.8	4.52	3.61	7	4	2	4	76	90
8	7.64	7.94	0.96	11.09	2.84	4.33	6	4	2	4	53	45
9	9.16	9.68	0.95	15.93	4.48	3.96	7	4	2	5	73	-
10	10.01	10.3	0.97	16.16	5.02	3.82	6	4	1	5	69	109
11	7.74	8.26	0.94	11.2	4.15	3.58	7	4	2	6	80	85
12	7.82	8.49	0.92	12.93	3.66	4.07	6	3	2	4	73	86
13	8.81	9.13	0.97	15.21	4.25	4.34	6	4	2	6	64	100
14	10.33	10.3	1.01	15.8	5.22	3.14	6	4	1	5	76	122
15	9.69	10.42	0.94	20.1	3.82	5.86	6	4	2	4	64	111
17	9.49	10.8	0.88	14.36	3.9	4.06	6	3	2	5	52	98
18	7.88	8.43	0.93	12.2	3.5	3.49	5	2	2	5	37	89
19	8.28	8.8	0.94	14.92	3.46	4.41	6	4	2	5	59	106
20	9.6	9.65	1	13.25	6.35	2.46	-	-	-	-	-	-
21	8.42	9.06	0.93	14.66	4.06	4.04	6	4	2	6	57	122
22	9.3	9.2	1.01	15.05	3.85	4.02	5	-	-	-	-	-
23	10.12	10.24	0.99	16.38	3.51	5	6	4	1	3	62	98
24	8.36	8.68	0.97	13.76	3.16	4.52	5	4	1	3	53	57
25	8.55	10.31	0.83	17.65	4.6	3.9	6	4	2	5	63	109
26	9.06	9.93	0.92	16.15	4.08	4.44	-	4	-	3	46	80
27	8.88	9.42	0.96	14.86	4.84	3.28	7	4	2	3	90	118

ACCN	LL	LD	LLDR	PEL	INL	PINLR	FRL	FRD	FRLDR	PEDL	FRAVWT	NSPFR
28	8.72	9.31	0.93	13.05	4.06	3.79	7	3	2	5	75	120
29	9.39	9.48	0.99	15.74	4.04	4.21	6	3	2	4	62	129
30	7.77	8.4	0.92	13.85	3.46	4.2	-	-	-	-	-	-
31	8.32	9.24	0.9	13.55	4.74	3.2	7	3	2	5	72	114
32	8.79	8.78	1.01	12.45	3.91	3.3	6	2	2	5	49	-
33	7.84	8.14	0.97	9.93	3.64	2.81	-	-	-	-	-	-
34	8.84	9.49	0.93	14.75	3.13	4.89	7	4	2	3	109	128
35	8.93	9.41	0.95	13.2	3.89	3.88	6	3	2	4	73	85
36	9.63	10.29	0.94	15.07	5.89	3.58	6	3	2	5	80	95
37	9.62	9.69	1	14.55	4.2	4.32	-	-	-	-	-	-
38	8.12	8.12	1	14.74	3.94	4.21	-	-	-	-	-	-
39	8.86	10.17	0.88	15.31	4.54	3.91	8	4	2	5	90	123
40	6.82	8.4	0.8	12.09	2.14	5.71	7	4	2	3	148	-
41	8.06	8.62	0.93	11.34	4.44	3.02	-	-	-	-	-	-
42	9.7	11	0.88	23.5	4.52	5.2	-	-	-	-	-	-
43	7.94	9.55	0.84	14.7	3.39	4.37	7	4	2	4	85	105
45	7.81	8.53	0.92	13.18	4.88	2.98	7	3	2	4	73	110
46	10.9	12.07	0.92	15.78	4.9	3.29	8	4	2	2	118	177
47	7.62	8.12	0.94	12.8	3.1	4.15	-	-	-	-	-	-
48	8.54	9.27	0.93	13.52	2.1	6.36	6	3	2	5	103	121
49	12.21	12.22	1	21.96	3.87	5.74	7	4	2	4	97	136
50	11	11.45	0.96	14.55	4.76	4.08	7	4	2	3	112	102
51	9.75	10.78	0.91	14.2	4.5	3.2	-	-	-	-	-	-
52	8.86	9.05	0.99	14.65	3.29	4.53	7	3	2	4	88	130
53	9.29	9.65	0.96	18.28	5.55	4.29	7	3	2	3	85	162
54	9.91	10.33	0.96	15.3	3.59	4.37	6	4	2	4	85	112
55	8.65	9.02	0.96	12.01	4.68	2.91	7	3	2	3	85	-
56	9.21	9.47	0.99	14.13	4.69	3.42	9	4	2	3	137	108
57	9.13	10.06	0.9	13	4	3.44	6	3	2	3	81	116
58	9.27	9.18	1.01	14.76	5.23	3.11	6	4	2	4	92	-
59	9.74	10.01	0.99	14.86	4.83	3.47	6	3	2	3	68	158
60	8.14	8.11	1	13.56	4.64	3.71	-	-	-	-	-	-
61	8.77	9.19	0.94	13.82	4.21	4.17	5	2	2	4	41	42
62	8.88	9.96	0.9	16.31	3.75	4.79	6	4	2	4	86	99
63	9.97	10.81	0.92	14.95	4.3	3.63	7	3	2	3	93	115
65	9.88	10.67	0.94	15.91	4.58	4.47	7	3	2	4	95	100
66	9.08	9.87	0.92	20.34	2.93	7.11	-	3	-	3	72	-
67	8.88	9.32	0.97	14.31	3.75	3.88	7	3	2	2	90	150
68	7.69	8.02	0.96	11.76	3.9	3.05	6	3	2	5	75	115
69	9.54	10.63	0.9	17.68	5.33	3.84	6	3	2	4	62	94
70	9.25	9.74	0.95	15.3	3.78	4.18	6	3	2	4	65	111
71	9.92	9.92	1	17.76	5.18	4.26	6	3	2	5	75	-
72	8.69	8.67	1	14.74	3.81	4.23	7	3	2	4	84	103
73	10.36	11.1	0.93	19.25	4.21	5.31	6	4	2	4	90	117
74	9.63	9.4	1.03	17.54	4.19	5.56	6	3	2	3	68	136
75	10.06	9.06	1.12	15.02	4.32	3.83	6	3	2	3	64	93
76	9.65	9.97	0.99	15.91	4.13	5.57	-	-	-	-	-	-
77	9.17	10.17	0.91	16.1	4.23	4.88	8	3	2	4	98	-
78	8.65	9.01	0.96	16.6	4.13	3.96	-	-	-	-	-	-
79	8.21	9.19	0.9	16.18	3.49	5.53	-	-	-	-	-	-
80	9.75	10.26	0.94	16.06	4.33	4.69	7	4	2	4	97	97
81	9.66	9.73	0.99	16.04	4.79	3.61	-	-	-	-	-	-
82	9.72	10.9	0.9	13.29	4.29	3.43	6	3	2	5	68	86
84	10.36	10.9	0.95	23.5	4	5.88	7	3	2	3	84	-
85	10.46	11.59	0.91	20.29	3.15	6.43	7	4	2	4	112	-
86	9.12	10.33	0.88	17.13	3.49	5.27	8	4	2	6	110	94
87	9.15	10.22	0.89	13.01	4.44	3.44	-	-	-	-	-	-
88	8.19	9.11	0.9	13.2	3.22	4.13	-	-	-	-	-	-
89	9.23	9.37	0.99	14.05	5.39	2.85	7	4	2	3	120	105
90	9.1	9.81	0.93	15.07	4.65	3.25	-	-	-	-	-	-
91	8.73	8.62	1.01	13.53	4.59	3.29	7	4	2	4	90	131

ACCN	LL	LD	LLDR	PEL	INL	PINLR	FRL	FRD	FRLDR	PEDL	FRAVWT	NSPFR
92	9.45	10.11	0.93	14.32	3.86	3.85	6	3	2	3	71	93
93	9.41	9.58	0.98	14.35	3.48	4.24	7	-	-	-	-	-
94	8.23	8.52	0.97	13.03	3.6	3.8	6	4	2	2	100	127
95	9.6	10.21	0.94	21.44	3.76	6.6	7	4	2	7	134	78
96	10.6	10.32	1.03	15.93	4.79	3.57	7	4	2	3	90	-
97	8.7	9.45	0.93	15.88	3.63	4.32	6	3	2	3	75	119
98	10.31	10.66	0.97	17.68	5.17	4.07	7	3	2	3	92	159
99	9.17	9.23	0.99	15.54	4.02	4.2	-	-	-	-	-	-
100	9.49	10.29	0.92	15.94	5.52	3.57	7	3	2	4	76	156
101	8.9	9.22	0.97	17.11	4.69	4.63	5	3	2	3	55	-
102	10.62	11.07	0.96	18.87	4.21	5.07	6	3	2	3	57	108
103	9.14	9.45	0.97	15.72	4.4	4.37	7	3	2	8	74	129
104	10.4	9.44	1.04	15	5.85	2.69	5	3	2	5	41	112
105	9.58	10.69	0.9	20.85	3.7	5.64	-	-	-	-	-	-
106	10.43	10.81	0.96	17.03	4.15	4.63	7	3	2	3	88	162
107	9.79	10.78	0.91	15.79	5.01	3.47	6	4	2	2	86	-
108	9.64	10.53	0.91	17.37	3.88	4.65	7	3	2	4	88	120
109	9.06	9.41	0.97	14.66	3.9	3.74	-	-	-	-	-	-
110	10.08	11	0.92	19.56	4.85	5.11	7	3	2	4	90	151
111	9.21	9.94	0.94	15.22	4.84	3.47	-	-	-	-	-	-
112	9.38	9.09	1.04	14.86	4.32	4.35	7	3	3	6	69	87
113	9.95	11.03	0.91	14.77	5.2	3.03	-	-	-	-	-	-
114	9.54	10.04	0.95	18.95	3.88	5.04	7	3	2	3	70	-
115	10.25	10.92	0.95	16.37	4.78	3.81	6	3	2	4	71	121
116	8.97	9.29	0.92	17.63	4.6	3.81	-	-	-	-	-	-
117	10.27	10.4	1	17	5.19	3.53	6	4	2	6	85	-
118	9.74	10.78	0.9	14.1	5.2	2.78	7	3	2	5	69	110
119	9.25	8.75	1.13	10.96	6.83	1.79	-	-	-	-	-	-
120	9.67	10.38	0.93	17.53	3.45	5.46	-	-	-	-	-	-
123	8.53	9.08	0.94	15.81	4.75	3.68	6	3	2	4	62	136
124	9.48	10.06	0.96	14.79	5.5	3.06	-	-	-	-	-	-
125	8.33	9.13	0.91	11.97	4.72	2.57	7	3	2	4	88	-
126	9.27	9.31	0.99	16.38	4.84	4.04	5	3	2	3	58	-
127	9.38	9.67	0.97	12.78	5.24	2.68	-	-	-	-	-	-
128	8.85	9.47	0.92	12.8	4.6	2.78	-	-	-	-	-	-
129	7.65	8.58	0.89	12.85	5.19	2.98	-	-	-	-	-	-
130	8.58	9.18	0.94	13.98	4.22	3.42	7	3	2	7	90	111
131	8.52	9.05	0.94	14.58	4.75	3.15	6	3	2	5	86	-
132	9.49	9.99	0.95	14.06	5.31	3.03	6	3	2	5	73	116
133	8.47	8.99	0.94	12.35	3.54	3.58	7	3	2	6	89	-
134	8.24	8.96	0.92	14.61	4.29	3.49	-	-	-	-	-	-
135	9.46	8.93	1.06	12.66	4.23	3.25	-	-	-	-	-	-
136	8.25	8.17	1.02	12.99	3.99	4.01	-	-	-	-	-	-
137	8.35	8.58	0.98	12.77	3.64	3.6	6	3	2	4	70	70
138	8.69	8.37	1.03	14.23	3.37	4.45	6	3	2	4	60	125
139	8.67	8.98	0.97	12.98	5.32	2.48	7	3	2	5	70	92
140	8.08	8.45	0.97	12.01	5.25	2.63	-	-	-	-	-	-
141	8.05	8.24	0.97	11.55	4.52	2.85	-	-	-	-	-	-
142	8.23	8.86	0.93	12.07	4.53	2.88	-	-	-	-	-	-
143	8.32	8.58	0.97	10.23	3.56	3	-	-	-	-	-	-
144	8.97	9.15	0.98	12.48	4.13	3.47	-	-	-	-	-	-
145	9.31	9.21	1	13.65	6.37	2.28	-	-	-	-	-	-
146	8.32	8.62	0.96	14.14	4.58	3.1	-	-	-	-	-	-
147	8.21	9.19	0.89	10.94	4.92	2.25	7	3	2	6	85	-
148	8.63	9.09	0.95	13.99	5.31	3.12	-	-	-	-	-	-
149	8.28	9.04	0.91	13.2	4.26	3.13	8	3	3	3	53	-
150	8.44	8.83	0.95	12.87	5.05	2.75	7	3	2	5	92	147
151	7.82	8.51	0.92	16.63	2.99	6.65	5	3	2	5	62	114
152	8.52	9.64	0.89	13.49	5.56	2.51	6	3	2	4	57	94
154	6.48	6.28	1.04	8.14	4.14	2.06	-	-	-	-	-	-
155	10.25	10.45	0.98	13.75	5.45	2.58	-	-	-	-	-	-

ACCN	LL	LD	LLDR	PEL	INL	PINLR	FRL	FRD	FRLDR	PEDL	FRAVWT	NSPFR
156	8.29	8.5	0.97	12.25	4.88	2.64	-	-	-	-	-	-
157	8.28	9.09	0.91	12.58	4.73	2.99	6	3	2	4	58	107
158	9.16	9.72	0.95	13.69	4.81	2.99	6	3	2	9	63	-
159	8.47	8.6	0.98	12.16	5.44	2.73	6	3	2	5	54	-
160	9.77	9.79	0.98	15.69	5.56	2.84	-	-	-	-	-	-
161	9.51	10.7	0.9	15.51	5.29	2.97	6	3	2	4	71	96
162	8.63	8.62	1	11.47	5.41	2.46	6	3	2	3	68	-
163	9.3	9.2	1.01	13.9	6.1	2.28	-	-	-	-	-	-
164	8.88	9.61	0.92	17.03	5.25	3.84	6	3	2	3	79	105
167	7.52	7.82	0.97	11.59	4.57	2.59	6	3	2	1	70	-
168	8.18	8.82	0.93	12.99	4.73	2.94	6	-	-	-	-	-
169	8.92	9.31	0.95	14.28	4.78	3.53	6	3	2	2	75	122
170	8.34	8.88	0.94	14.8	4.26	3.97	6	3	2	2	68	97
171	9.05	8.38	0.99	12.94	5.82	2.55	-	-	-	-	-	-
172	8.5	8.64	0.99	18.9	3.99	5.06	6	3	2	5	64	107
173	8.75	9.23	0.96	12.94	6.28	2.09	6	3	2	7	64	83
174	8.23	8.5	0.92	11.85	4.92	2.45	9	3	3	5	95	147
175	8.56	8.84	0.97	13.48	5.28	2.72	7	3	2	4	75	169
176	8.3	8.35	1.01	12.71	4.38	2.83	7	3	2	3	75	113
177	9.85	9.83	1	15.8	5.83	2.9	-	-	-	-	-	-
178	9.58	9.97	0.96	12.17	5.65	2.37	6	2	2	2	47	114
179	8.29	8.58	0.97	12.46	4.71	2.94	6	3	2	3	53	98
180	9.29	9.37	0.99	16.93	4.79	3.7	-	-	-	-	-	-
181	9.99	10.28	0.98	14.64	5.98	2.99	6	3	2	6	54	-
182	9.41	9.49	0.99	15.44	5.59	2.81	6	3	2	3	53	-
183	12.1	13.6	0.89	21.2	6.2	3.42	6	3	2	3	68	-
184	10.63	11.88	0.89	19.88	7.38	2.7	6	3	2	4	75	-
186	8.93	10.04	0.89	15.55	4.1	4.2	7	3	2	7	68	-
187	8	8.7	0.93	12.65	4.05	3.11	6	3	2	4	70	-
188	6.75	9	0.75	9	3.4	2.65	7	3	2	2	75	123
192	8.72	9.12	0.96	17.25	4.91	3.49	7	3	2	4	92	-
193	10.81	12.12	0.92	20.4	8.09	2.96	6	3	2	4	67	106
194	10.05	10.43	0.97	18.85	6.5	3.22	7	4	2	5	100	-
201	9.96	11.52	0.86	17.38	5	3.77	7	4	2	4	89	84
202	8.62	9.17	0.94	16.62	3.81	4.57	6	3	2	5	60	68
Mean	9.05	9.54	0.95	14.89	4.50	3.76	6.44	3.32	1.98	4.10	76.73	111.74
Max	12.21	13.6	1.13	23.5	8.09	7.11	9	5	3	9	148	177
Min	6.48	6.28	0.75	8.14	2.1	1.79	5	2	1	1	37	42
CV	13.44	12.12	8.2	21.31	29.23	31.63	10.21	15.18	18.34	32.06	22.12	23.44
Range	5.73	7.32	0.38	15.36	5.99	5.32	4.00	3.00	2.00	8.00	111.00	135.00
Std	0.885	0.979	0.048	2.566	0.884	0.980	0.735	0.557	0.265	1.299	18.988	25.038

ACCN= accession number, LL= leaf length, LD= leaf diameter, LLDR= leaf length to its diameter ratio, PEL= petiole length, INL=inter-node length, PINLR= petiole to inter-node lengths ratio, FrL= fruit length, FrD= fruit diameter, FrLLDR= fruit length to diameter ratio, PedL= Peduncle length, FrAvWt = Fruit average weight, NSPFr = number of seed per fruit. All measurements are metric (cm), except AvFrWt (in gm) and NSPFr (count), CV = coefficient of variation

Appendix Table 6. The first 5-15% of accessions with highest and lowest means for each morphological trait's performance (seed grown anchote)

Accn	**VN**	**Accn**	**AVRWT**	**Accn**	**DMC**	**Accn**	**LD**	**Accn**	**INL**	**Accn**	**RL**
110	1	134	39.88	120	13	6	6.1	61	2	101	6.4
127	1	126	41.58	167	14	172	6.2	74	2.2	87	7.19
28	1.3	123	41.64	95	14	180	6.3	172	2.4	103	7.28
111	1.3	180	42.38	116	14	181	6.4	95	2.4	109	7.45
114	1.3	125	43.41	112	24	7	6.5	164	2.5	104	7.55
123	1.3	142	43.58	128	25	140	6.5	180	2.5	99	7.95
136	1.3	129	44.26	71	26	77	8.9	99	2.5	143	7.98
140	1.3	120	44.34	177	26	62	8.9	89	3.9	7	12.6
21	3.3	156	113.86	181	26	57	9	5	4.3	178	12.7
60	3.3	68	114.63	156	28	56	9	59	4.5	58	12.8
159	3.3	39	114.8	**Accn**	**LL**	39	9.1	12	4.5	12	12.9
160	3.5	96	115.42	6	5.9	59	9.4	31	4.6	9	13.1
68	3.5	65	125.61	180	6.2	96	9.4	152	4.7	**Accn**	**RD**
156	4	62	129.48	172	6.2	65	9.5	161	4.8	24	3.1
45	4	**Accn**	**YLD**	30	6.3	103	9.6	**Accn**	**PINLR**	7	3.2
		134	9.21	181	6.5	**Accn**	**PEL**	14	2.1	180	3.2
		125	9.73	7	6.5	91	7	91	2.15	71	3.2
		126	10.4	39	8.8	6	7.4	9	2.34	176	3.2
		180	10.6	119	8.8	41	7.4	11	2.36	5	3.3
		123	10.9	57	8.9	9	7.6	24	2.45	177	3.3
		142	10.9	65	9.1	7	7.8	6	2.46	134	3.4
		129	11	59	9.2	30	7.9	30	2.47	87	5
		131	11.1	96	9.5	13	8	5	2.48	167	5.1
		98	26	103	9.5	77	12	61	4.77	102	5.1
		61	26.8	**Accn**	**RLDR**	57	12	110	4.91	118	5.2
		77	26.8	101	1.35	39	12	95	4.95	56	5.4
		60	27.5	103	1.38	59	12	73	4.96	103	5.5
		56	27.9	87	1.5	67	12	101	5.02	106	5.5
		156	28.5	104	1.59	106	12	74	5.54	45	5.5
		39	28.7	106	1.6	160	12			60	5.6
		65	31.4	109	1.62	31	12				
		62	32.4	102	1.67	90	12				
				100	1.71	93	12				
				9	3.5	65	13				
				6	3.5	103	13				
				71	3.53						
				7	4.1						

Note: Anchote's fresh root weight ranges from 2gm (ANC43) -1155gm (ANC 28)

Appendix Table 7. Means and coefficients of variations of anchote populations by altitudinal group

Trait	Population											
	1		2		3		4		5		6	
	Mean	CV	Mean	CV	Mean	CV	Mean	CV	Mean	CV	Mean	CV
VN	2.1	38	2.3	34.5	2.18	41	2.1	46.75	2.76	39.16	1.69	44
AVRWT	76	42	72	49.2	67.2	56.7	76.76	52.9	82.6	66	54.67	54.44
YLD	18.9	42	18.1	48.6	16.5	52.65	18.65	48	19.15	60.36	13.4	51.2
DMC	19.4	29	19	21.9	19.56	28.3	18.25	27	21.1	37.3	20	18.94
LL	7.6	13.5	7.8	14.1	7.4	16.3	7.88	18	7.5	14	7.2	13.73
LD	7.7	14.2	7.8	14	7.35	15.9	7.98	14.96	7.57	14	7.2	13.49
LLDR	0.99	7.9	0.99	5.8	1	8.5	0.99	7.98	0.98	7.6	1	6.85
PEL	10	20.4	10	21.5	9.5	21.86	10.55	18.15	10.37	20	9.69	17.17
INL	3.34	41	3.5	33.3	3.5	35.76	3.1	41.4	3.4	49.4	3.16	44.47
PINR	3.5	48	3.2	38.6	3.26	44.2	4.1	46.77	3.6	43.16	3.8	50.3
RL	9.86	26	9.98	21.8	9.97	24.5	9.14	26.77	9.65	21.8	9	23.4
RD	4.3	26.4	4	23.79	3.9	26.8	4.4	21.5	4.19	23.57	3.74	20.3
RLDR	2.4	31.3	2.57	28.8	2.73	33.3	2.1	31.2	2.42	29	2.5	24.6

Population 1=1497–1666, 2=1667–1836, 3= 1837–2005, 4= 2006–2174, 5=2175–2343, 6 = >2344

Appendix Table 8. 186 anchote accessions and two wild “relatives” used for bulk DNA extraction

Accession Code (ANC-)													
1	15	31	47	62	79	95	109	120	134	147	162	178	193
2	17	32	48	63	80	96	110	123	135	148	164	179	194
3	18	33	49	65	81	97	110	124	136	149	167	180	195
4	19	34	50	67	82	98	111	125	137	150	168	181	200
5	21	35	51	68	84	99	112	126	138	151	169	182	201
6	22	36	52	69	85	100	113	127	139	152	170	183	202
7	23	37	53	70	86	101	114	128	140	154	171	184	204
8	24	38	54	71	88	102	115	129	141	156	172	186	216
9	25	39	55	72	89	103	116	130	142	157	173	187	227
10	26	40	56	73	90	104	117	131	143	158	174	188	232
11	27	41	57	74	91	105	118	132	144	159	175	189	A-1
12	28	43	58	75	92	106	119	133	145	160	176	191	A-2
13	29	45	59	76	93	107	120	133	146	161	177	192	-
14	30	46	61	77	94	108	-	-	-	-	-	-	-

Appendix Table 9. The frequencies of each alleles across eight EST-SSRs loci per population: from 146 individual samples (24 alleles) (A) and from 33 bulked samples (20 alleles) (B).

A

Locus	Allele Size	Population															All-W
		Ay-La	DS-Dw	Sy-An	Gm-Nj	Mt-Al	Hr-Yy	Ch-De	Gc-Dh	Gu-Go	Gr-Sh	Dd-Dc	Gg-Dg	Ac-Bt	Ss-Wt	Ar-Ld	
	170	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	173	0.5	0.5	0.7	0.6	0.9	0.8	0.5	0.9	0.6	0.6	0.7	0.5	0.6	0.5	0.5	0.6
	176	0.4	0.6	0.3	0.5	0.2	0.2	0.5	0.1	0.4	0.4	0.3	0.6	0.4	0.5	0.5	0.4
WM-5	179	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	225	0.6	0.4	0.6	0.4	0.5	0.5	0.5	0.4	0.4	0.6	0.5	0.6	0.6	0.6	0.6	0.5
WM-34	229	0.4	0.6	0.4	0.6	0.5	0.5	0.6	0.6	0.6	0.4	0.5	0.4	0.4	0.4	0.4	0.5
	402	0.0	0.5	0.1	0.4	0.3	0.4	0.1	0.2	0.2	0.3	0.4	0.4	0.3	0.4	0.4	0.3
WM-32	404	1.0	0.6	0.9	0.6	0.8	0.6	0.9	0.9	0.8	0.7	0.6	0.7	0.7	0.6	0.6	0.7
	227	0.1	0.1	0.3	0.2	0.2	0.2	0.0	0.1	0.3	0.1	0.0	0.0	0.2	0.0	0.1	0.1
	229	1.0	0.8	0.7	0.8	0.9	0.8	1.0	1.0	0.7	0.9	1.0	1.0	0.8	1.0	0.9	0.9
WM-30	231	0.0	0.2	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
	395	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.6	0.4	0.4	0.5	0.4	0.4	0.1	0.2
WM-8	398	1.0	0.9	0.9	0.9	0.9	1.0	0.9	1.0	0.4	0.7	0.7	0.5	0.6	0.6	0.9	0.8
	258	0.0	0.0	0.3	0.0	0.0	0.0	0.2	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.2	0.1
	260	0.1	0.4	0.2	0.1	0.1	0.6	0.1	0.0	0.2	0.4	0.7	0.2	0.2	0.3	0.3	0.2
	262	0.3	0.2	0.1	0.2	0.2	0.1	0.1	0.6	0.3	0.3	0.0	0.2	0.3	0.3	0.2	0.2
	264	0.5	0.2	0.1	0.3	0.3	0.0	0.4	0.0	0.0	0.0	0.1	0.3	0.1	0.3	0.1	0.2
	266	0.2	0.1	0.3	0.4	0.3	0.3	0.2	0.4	0.4	0.3	0.3	0.4	0.3	0.0	0.3	0.3
WM-24	268	0.0	0.3	0.0	0.1	0.2	0.1	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1
	323	0.4	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
WM-29	339	0.6	0.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	234	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.0
	243	0.0	0.0	0.0	0.3	0.1	0.0	0.1	0.2	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.1
WM-25	246	1.0	1.0	1.0	0.7	1.0	1.0	0.9	0.8	1.0	1.0	1.0	0.8	0.9	1.0	0.9	0.9

B

Locus	Allele Size	Pop1	Pop2	Pop3	Pop4	Pop5	All-W
WM-5	170	0	0	0.0556	0.0714	0	0.0303
	173	0.5833	1	0.5556	0.7857	0.6667	0.6667
	176	0.4167	0	0.3889	0.1429	0.3333	0.303
WM-34	225	0.5	0.5	0.3571	0.5	0.625	0.4808
	229	0.5	0.5	0.6429	0.5	0.375	0.5192
WM-32	402	0.1	0.1667	0.3333	0.375	0.5714	0.3824
	404	0.9	0.8333	0.6667	0.625	0.4286	0.6176
WM-30	227	0.0833	0.5	0.1667	0.3125	0.0714	0.2143
	229	0.9167	0.5	0.8333	0.6875	0.9286	0.7857
WM-8	395	0.3333	0.5	0.25	0.125	0.2857	0.25
	398	0.6667	0.5	0.75	0.875	0.7143	0.75
WM-24	258	0.0833	0	0	0.125	0.2143	0.0857
	260	0.1667	0.3333	0.5556	0.25	0.2143	0.3286
	262	0	0.5	0.1111	0.25	0.2143	0.1714
	264	0.25	0	0.1667	0.0625	0.2857	0.1714
	266	0.4167	0.1667	0.1667	0.1875	0.0714	0.2
	268	0.0833	0	0	0.125	0	0.0429
WM-29	339	1	1	1	1	1	1
WM-25	243	0.1667	0.3333	0.0556	0	0	0.0735
	246	0.8333	0.6667	0.9444	1	1	0.9265

Pop1 = WESTERN WOLLEGA, Pop2 = KELLAM WOLLEGA, pop3 = EASTERN WOLLEGA, Pop4 = ILLUABABORA, Pop5 = JIMMA,
All-W= overall average allele frequency

Appendix Table 10. Linear Discriminate Function of 13 traits for eight classes

trait	1	2	3	4	5	6	7	8
VN	-0.03556	-0.05277	-0.06381	0.75697	-0.29511	-0.22392	-0.07933	0.87037
AVWT	0.26638	-0.5308	0.33195	1.40274	-1.76042	2.06671	0.18914	-1.8645
YLD	-0.15865	-0.70214	0.42229	-0.80815	0.74845	0.70382	-1.20515	-1.35587
DMC	-0.13862	-0.12812	0.40179	-0.02536	1.33367	-0.80626	3.18965	-0.87059
LL	**-2.90856**	**-1.44398**	**-4.28479**	-5.99873	9.69014	-11.6404	-8.40417	-5.86885
LD	**3.36754**	**1.69785**	**3.42895**	5.23793	-9.84534	9.30524	6.31873	7.97685
LLDR	**1.09162**	0.60054	**1.3465**	2.34071	-4.10255	4.27801	3.14187	2.23862
PEL	0.27834	-0.25381	-0.69803	-0.78871	0.08373	-0.82203	2.45977	0.0588
INL	-0.57814	0.57618	**1.39519**	**1.55983**	0.78619	0.71658	-2.09229	-0.93876
PINLR	-0.79451	0.96042	**1.64913**	**2.10363**	1.2801	1.68003	-2.57214	-0.05981
RL	-0.18808	0.07202	0.36504	0.72334	-1.27971	-0.1995	2.71201	2.51434
RD	0.34859	-0.25984	**-1.1642**	-1.16523	2.2446	-2.45614	-0.98121	-1.44745
RLDR	0.36512	0.0209	**-1.55815**	-1.682	5.33022	-1.67356	-5.03737	-2.8309

Appendix Table 11. Linear Discriminate Function for classes (by woreda)

Trait	1	2	3	4	5	6	7	8	9
Constant	-4.51	-5.97	-13.10	-23.03	-63.28	-38.76	-37.21	-248.78	-81.18
VN	-2.87	-1.72	-4.84	-4.21	5.95	-3.59	4.37	**47.03**	12.18
AVWT	6.60	1.06	0.02	3.43	**-26.35**	**21.11**	-18.06	**-74.09**	**-64.13**
YLD	-0.91	1.61	4.48	-5.78	8.64	-12.50	-0.65	23.17	**45.22**
DMC	-1.89	-2.90	2.08	1.97	9.25	-0.54	14.51	7.52	-12.77
LL	**47.06**	**23.41**	-10.01	0.12	-9.31	**-23.30**	-74.70	**34.17**	27.44
LD	**-48.98**	**-23.42**	**15.99**	0.06	0.52	**32.64**	**70.96**	**-54.64**	**-50.00**
LLDR	-13.09	-6.96	2.24	-2.02	-1.49	4.24	18.76	1.63	-7.59
PEL	0.84	-0.08	8.15	-7.95	-9.84	-7.38	12.72	**-37.32**	19.35
INL	2.60	2.94	-3.71	0.92	-10.77	2.10	-20.62	-4.44	-8.97
PINLR	-4.10	4.77	-2.85	-1.06	-6.69	-4.59	-19.78	22.92	2.05
RL	-8.37	-2.17	**20.40**	**17.62**	-0.42	**-24.96**	6.24	**-49.09**	-19.37
RD	-9.43	-9.76	-14.59	8.87	**48.43**	-3.92	23.69	**139.78**	**33.18**
RLDR	0.93	-7.79	-21.34	0.28	**43.88**	3.38	0.41	122.03	**46.21**

Appendix Table 12. Linear Discriminate Function for five clusters (populations)

Trait	Cluster				
	1	2	3	4	5
VN	-428,120,695.00	54,907,469.00	-38,731,348.00	-148,392,169.00	-86,093,590.00
AVWT	59,069,263.00	-51,744,883.00	11,983,267.00	70,901,385.00	5,120,712.00
YLD	368,203,439.00	-15,073,042.00	-664,081.00	87,111,640.00	14,459,482.00
DMC	144,474,479.00	-17,004,562.00	35,204,999.00	57,169,735.00	70,270,584.00
LL	-74,065,715.00	37,147,893.00	46,318,147.00	19,755,014.00	15,732,963.00
LD	-89,355,528.00	-23,531,687.00	-19,960,459.00	-30,709,108.00	-7,305,806.00
LLDR	-351,831,443.00	16,782,671.00	-80,773,425.00	-153,833,046.00	-110,321,300.00
PEL	14,417,406.00	8,256,711.00	20,400,919.00	-1,909,581.00	47,285,871.00
INL	-225,409,364.00	23,929,680.00	-28,920,802.00	-64,727,797.00	-79,487,740.00
PINLR	-456,488,320.00	48,426,775.00	-87,321,634.00	-161,227,724.00	-186,563,780.00
RL	-71,432,726.00	-3,974,899.00	-44,517,904.00	-57,857,747.00	-43,225,894.00
RD	-50,506,376.00	26,099,388.00	9,096,369.00	-37,776,552.00	29,677,910.00
RLDR	-48,534,335.00	25,384,150.00	37,741,853.00	12,582,110.00	27,403,204.00

Appendix Table 13. Bottleneck significant tests:

Sign, wilcoxon, and mode shift tests for existence of bottleneck in anchote populations by three mutational models (Ratios represent the number of EST_SSR loci exhibiting heterozygosity deficiency to excess).

Population code	Mutation model	Significant test			
		R	Sign test	Wilcoxon Test	Mode shift
Ay-La	I.A.M	2:4	ns	ns	Normal L shaped distribution
	T.P.M	2:4	ns	ns	
	S.M.M	3:3	ns	ns	
Ds-Dw	I.A.M	2:5	ns	*	Shifted
	T.P.M	2:5	ns	ns	
	S.M.M	2:5	ns	ns	
Sy-An	I.A.M	2:4	ns	ns	Shifted
	T.P.M	2:4	ns	ns	
	S.M.M	2:4	ns	ns	
Gm-Nj	I.A.M	2:5	ns	ns	Normal L shaped distribution
	T.P.M	3:4	ns	ns	
	S.M.M	3:4	ns	ns	
Mt-Al	I.A.M	3:4	ns	ns	Shifted
	T.P.M	4:3	ns	ns	
	S.M.M	4:3	ns	ns	
Table 13 continued					
Hr-Yy	I.A.M	3:3	ns	ns	Shifted

	T.P.M	3:3	ns	ns	
	S.M.M	4:2	ns	ns	
Ch-De	I.A.M	3:3	ns	ns	
	T.P.M	3:3	ns	ns	
	S.M.M	4:2	ns	ns	Shifted
Gc-Dh	I.A.M	3:3	ns	ns	
	T.P.M	3:3	ns	ns	Normal L shaped distribution
	S.M.M	3:3	ns	ns	
Gu-Go	**I.A.M**	**0:6**	*	*	
	T.P.M	**1:5**	**ns**	*	
	S.M.M	**1:5**	**ns**	*	Shifted
Gr-Sh	**I.A.M**	**1:5**	**ns**	*	
	T.P.M	**1:5**	**ns**	*	
	S.M.M	**1:5**	**ns**	*	Shifted
Dd-Dc	I.A.M	1:4	ns	ns	
	T.P.M	1:4	ns	ns	
	S.M.M	1:4	ns	ns	Shifted
Gg-Dg	**I.A.M**	**1:5**	**ns**	*	
	T.P.M	1:5	ns	*	Shifted
	S.M.M	1:5	ns	ns	
Ac-Bt	I.A.M	1:6	ns	ns	
	T.P.M	2:5	ns	ns	Normal L shaped distribution
	S.M.M	2:5	ns	ns	
Ss-Wt	**I.A.M**	**0:5**	*	*	
	T.P.M	**0:5**	*	*	Shifted
	S.M.M	**0:5**	*	*	
Ar-Ld	I.A.M	4:4	ns	ns	
	T.P.M	4:4	ns	ns	Normal L shaped distribution
	S.M.M	4:4	ns	ns	

Note: R=ratio between Heterozygosity deficient: excess

**=significant at 0.01% level of significance *=significant at 0.05% level of significance

Pop 1=Ay-La; 2=Ds-Dw; 3=Sy-An; 4=Gm-Nj 5=Mt-Al; 6= Hr-Yy; 7= Ch-De; 8= Gc-Dh; 9= Gu-Go; 10= Gr-Sh; 11= Dd-Dc; 12=Gg-Dg; 13=Ac-Bt; 14=Ss-Wt; 15=Ar-Ld

I.A.M= infinite allele model; T.P.M= two-phase model; S.M.M= stepwise mutation model