Evaluation of Tef [*Eragrostis tef* (Zucc.)Trotter] Genetic Diversity Using Morphological Characters and Microsatellite Markers

By

Mahilet Tadesse Kidane

A Thesis Submitted to Department of Applied Biology

School of Applied Natural Science

Office of Graduate Studies

Adama Science and Technology University

July, 2020

Adama, Ethiopia

Evaluation of Tef [*Eragrostis tef* (Zucc.)Trotter] Genetic Diversity Using Morphological Characters and Microsatellite Markers

By

Mahilet Tadesse

Advisor: Dr. Mulugeta Kebede

A Thesis Submitted to Department of Applied Biology

School of Applied Natural Science

Office of Graduate Studies

Adama Science and Technology University

July, 2020

Adama, Ethiopia

TABLE OF CONTENTS

Contents Pages

APPROVAL PAGE

We, the undersigned, members of the board of Examiners of the final open defense by Mahilet Tadesse Kidane have read and evaluated this thesis entitled "Evaluation of Tef [*Eragrostis tef* (Zucc.)Trotter] Genetic Diversity Using Morphological Characters and Microsatellite Markers" and examined the candidate. This is therefore to certify that the thesis has been accepted in the partial fulfillment of the requirement of the degree of Master of Science in Applied Biology.

Name	Signature	Date
_____	_____	_____
Name of Student		
_____	_____	_____
Advisor		
_____	_____	_____
External Examiner		
_____	_____	_____
Internal Examiner		
_____	_____	_____
Chair person		
_____	_____	_____
Head of Department		
_____	_____	_____
School Dean		
_____	_____	_____
Post graduate Dean		

DECLARATION

I declare that this thesis is my original work and has not been presented for a degree or diploma in this or other University and that all sources of materials used for this thesis have been duly acknowledged. I have followed all ethical and technical principles of scholarship in the prepar - ation, data collection, data analysis and compilation of this Thesis. Any scholarly matter that is included in the Thesis has been given recognition through citation. This Thesis is submitted in partial fulfillment of the requirements for M.Sc. degree at the Adama Science and Technology University. The Thesis is deposited in the Adama Science and Technology University library and is made available to borrowers under the rules of the library.

Mahilet Tadesse

Signature: _____

This thesis has been submitted for examination with my approval as a University advisor and co-advisor.

Advisor: Mulugeta Kebede, PhD

Signature: _____

Co-advisor: Dejene Girma, PhD

Signature: _____

BIOGRAPHICAL SKETCH

The author was born from her father Tadesse Kidane and her mother Bezawerk Tadesse on January 25, 1993 in Awash Melkassa, East Shewa Zone of Oromia Regional State. She attended elementary and junior education at Awash Melkassa Elementary and Junior School from 2000-2007. Then she went to Adama High school from 2008-2009 and she went to Hawas Preparatory School from 2010 -2011, where she had completed her preparatory classes in 2011. After the completion of her preparatory school education, she joined University of Gondar School of Applied and Computational Sciences Atse Tewodros campus and graduated with B.Sc. Degree in Biotechnology on July 05, 2014. Soon after graduation; she was employed by Jimma University College of agriculture and veterinary medicine as technical assistant and served for five months. Then she was employed by Ethiopian Institute of Agricultural Research (EIAR) in April 2015. She was assigned to plant biotechnology research case team especially on plant tissue culture at Debre Zeit Agricultural Research Center (DZARC) as Junior Researcher until she joined the School of Graduate Studies at Adama Science and Technology University to pursue her M.Sc. degree in Biotechnology.

ACKNOWLEDGMENTS

First of all, I would like to thank the Almighty GOD for all the protection and support in any path, and who made all the hopelessness possible. I would like to thank my research main advisor Dr. Mulugeta Kebede for guiding, critical reading and comments throughout my thesis work. It is my pleasure to express great thanks and deep sense of gratitude to my co-advisor Dr. Dejene Girma, without whose excellent guidance, constructive suggestions and critical comments, this work would not have been possible.

My special appreciations go to Dr. Zerihun Tadele and Dr. Tesfaye Disasa for the chemical support during the laboratory work and Dr. Kebebew Assefa for all his support. I am also grateful to the Ethiopian Institute of Agricultural Research (EIAR) and National Agricultural Biotechnology Research Center (DZARC) for giving me opportunity to pursue my study.

I am also very much grateful to Messele Molla, Tsion Fikre, Solomon Tamiru and Sisay Kidane for their thoughtfulness, encouragement, advice and support during laboratory work, data analysis and for their genuine cooperation in every aspect. I am very much grateful to NABRC laboratory personnel for their technical assistance especially to Kalkidan Tesfu, Yitaseb Nigusie, Lidiya Ashenafi. Dejene, Biruktait, Moges. Doni, Derara and Ejigayehu.

I would like to extend my appreciation and sincere thanks to Nigussu Hussein and DZARC tef greenhouse workers for their assistance in trial sowing, data collection and material support. My appreciation also goes to the staff of DZARC tef breeding research program for facilitating green house materials support.

All my friends, Tewabech Alemu, Tadele Getachew, Dagnachew Genene, Kalkidan Tesfu and Bizuwerk Tafes. Last but not least my special appreciations go to Ethiopian Biodiversity Institute and staff for the tef materials support.

I would like to express my deepest thanks and regards to my dear family Bezawerk Tadesse, Shawel Tadesse, Wudnesh Tilahun, Tamirat Negash, Yordanose Shawel and other close relatives for their continuous love and daily prayers for my happiness and better life. Thank you very much for your love and encouragement. Finally, I like to thank individuals, who have contributed in my thesis work and guide in different ways.

LIST OF TABLES

Tables Pages

LIST OF FIGURES

Figures **Pages**

LIST OF APPENDICES

Appendix **Pages**

ABBREVIATIONS AND ACRONYMS

AFLP	AMPLIFIED FRAGMENT LENGTH POLYMORPHISM
AMOVA	ANALYSIS OF MOLECULAR VARIANCE
ANOVA	ANALYSIS OF VARIANCE
CRD	COMPLETELY RANDOMIZED DESIGN
CSA	CENTRAL STATISTICAL AGENCY
CTAB	CETYL TRIMETHYL AMMONIUM BROMIDE
EBI	ETHIOPIAN BIODIVERSITY INSTITUTE
PIC	POLYMORPHIC INFORMATION CONTENT
PCA	PRINCIPAL COMPONENT ANALYSIS
PCoA	PRINCIPAL COORDINATE ANALYSIS
RFLP	RESTRICTION FRAGMENT LENGTH POLYMORPHISM
SNP	SINGLE NUCLEOTIDE POLYMORPHISM
SAS	STATISTICAL ANALYSIS SYSTEM
SSR	SIMPLE SEQUENCE REPEATS
TAE	TRIS-ACETATE-EDTA

ABSTRACT

Tef (Eragrostis tef) is a cereal crop of great cultural and economic importance in Ethiopia. Even though Ethiopia is the center of origin and diversity; exploiting the genetic potential of the crop has been a challenge because of the limited information available regarding the diversity state of the germplasm accessions in the Gene Bank. Assessment of the genetic diversity of these accessions using morphological and molecular markers is one of the primary tasks in tef breeding. Enhancing its breeding using markers requires a systematic and continuous evaluation of the genetic diversity of the crop species. Therefore, this study was conducted to assess the genetic diversity of tef accessions obtained from the Gene Bank using 12 morphological and 10 SSR markers. Sixty four tef accessions were planted in completely randomized design at DZARC during the 2019 cropping season. Analysis of variance for quantitative traits showed significant variation ($P < 0.05$) among the tested accessions. Cluster analysis of quantitative traits revealed that the accessions were grouped in to eight distinct clusters. The first five principal components explained 62.62% of the total variation. The analysis of molecular variance showed highly significant ($P<0.001$) variances; 56% and 39% of the total variability could be attributed to differences within and among tef accessions, respectively. Lowest Fst value (0.05) was recorded among the studied tef populations. A total of 914 alleles were detected with an average 14.5 alleles per locus. Polymorphic Information Content (PIC) ranged from 0.06 (CNLTs2) to 0.9 (CNLTS6), with a mean of 0.8 suggesting sufficient discrimination power of the markers to discriminate the tested accessions. The mean value of major allele frequency and number of effective alleles were 0.33 and 3.32 respectively. The mean value of gene flow (Nm) and Shannon's information index (I) 4.74 was 1.65 respectively. The highest numbers of private alleles (1.0) were recorded from West Gojjam populations. The Nei's genetic distance value ranged from 0.63 to 0.91.Three major groups were created by cluster analysis by merging accessions from different populations, which may imply the existence of gene flow. The PCoA analysis showed that accessions from different collection sites often clustered together. Therefore, attention should be given for these areas in any improvement and conservation programs. The result from these studies suggests that possibilities for identification of tef accessions superior for yield related traits, indicating the need for the initiation of a planned breeding and conservation programs.

Keywords: Cluster, Genetic diversity, Morphology, SSR

1. INTRODUCTION

1.1 Background

Tef [*Eragrostis tef* (Zucc.)Trotter] belongs to the grass family Poaceae and genus *Eragrostis*. It is an allotetraploid species with a base chromosome number of 10 (2n=4x=40) with genome size of 672 Mbp (Cannarozzi *et al.,* 2014) and 622 Mbp (VanBuren *et al*., 2018). It is self-pollinated with chasmogamous and hermaphroditic flowers. Tef is believed to have originated and diversified in Ethiopia (Vavilov, 1951). Tef grain is a rich source of protein and nutrients and has additional health benefits including that the seeds are free from gluten (Spaenij-Dekking *et al*., 2005). In general, tef provides quality food and grows under marginal conditions. Tef has remained an important crop to Ethiopian farmers for several reasons namely: the high price of its grain and straw, performs better than other cereals under moisture stress and waterlogged conditions, its grain can be stored for a long period of time without being attacked by weevils, and suffers little or no disease epidemic that has threatened its performance, it fits for various cropping systems, and serves as a catch and low-risk reliable crop (Ketema, 1993). In Ethiopia it is the leading crop accounting for 3.36 million hectares of the total land and 5.47 million tons of the gross grain production of all cereal. The gene bank in Ethiopia hold over five thousand tef accessions collected from diverse geographical regions in terms of climate and elevation (CSA 2014). The productivity of tef is low with the national average standing at 1.64 t-ha (CSA 2016). The major yield limiting factors shattering, low genetic potential of the local varieties, poor pre and post-harvest agronomic management practice and also tef possesses tall, weak stems that easily succumb to lodging caused by wind or rain. A number of scientists have conducted to evaluate the diversity of tef germplasms using morphological markers (Assefa *et al*., 1999, 2000; Adnew *et al*., 2005; Plaza-Wuthrich *et al*., 2013), and others few using genetic markers amplified fragment length polymorphism (AFLP) (Ayele *et al*., 1999); random amplified polymorphic DNA (RAPD) (Bai *et al*., 2000); inter simple sequence repeat (ISSR) (Assefa *et al*., 2003a) and simple sequence repeat (SSR) (Chanyalew, 2007; Zeid *et al*., 2012; Abraha *et al*., 2016). The former two genetic markers showed relatively low level of diversity however, the latter two genetic markers revealed ample type of genetic diversity and point out the value of SSR markers in genetic diversity analysis as well as in establishing genetic relationship among tef genotypes. Genetic markers are useful tools that provide an opportunity to characterize varieties and to measure genetic relationships more precisely than other markers Simple sequence repeat (SSR) markers are particularly useful, since they are co-dominant, high reproducibility, polymorphic, have low operational cost, locus-specificity and revealing of high allelic diversity (Mohan *et al.*, 1997). To that effect, they have been used extensively in tef and other crops for various breeding and diversity studies. Hence, the objective of this study is to estimate the pattern and

level of genetic diversity and relatedness among tef accessions using polymorphic simple sequence repeat (SSR) DNA markers.

1.2 Statement of the Problem

Improvements of tef varieties are strongly depend on exploitation of natural resources. Currently a huge number of accessions are available in the gene pool. However, there are limited diversity related studies on tef. Applying molecular and morphological diversity analysis helps to assess the diversity state of the crop. So characterized different tef accessions which are collections from different geogr-aphical origins by the help of molecular marker and some quantitative traits which can generate base line information.

1.3 Significance of the Study

The Ethiopian Biodiversity Institute holds about 5,000 tef accessions which could be used for developing new tef varieties. The genetic diversity of these accessions has been probed using different molecular marker systems. The body of knowledge present in the literature shows that the large portion of the tef accessions in the gene bank has remained un-characterized. This provides the rational to conduct molecular characterization of the genetic resources available at the Gene Bank. Therefore, this study is expected to:

- Enhance the knowledge of tef diversity using molecular and morphological markers.
- Evaluate genetic relationships among tef accessions.
- Provides documentary information about relationship between and within the tef accessions.

1.4 Objectives

1.4.1 General Objective

- To assess the genetic diversity of tef [*Eragrostis tef* (Zucc.)Trotter] accessions using morphological and microsatellite markers

1.4.2 Specific Objectives

- ✓ To determine the molecular diversity of tef using SSR marker system
- ✓ To characterize the accessions according to their morphological properties
- ✓ To evaluate the genetic variability among and within populations

2. LITERATURE REVIEW

2.1 Origin, Taxonomy and Distribution of Tef

Tef *(Eragrostis tef* (Zucc.)Trotter) is a domestic wild grass belonging to the Poaceae family, sub-family *Chloridoideae* (*Eragrostoideae*), tribe *Eragrostidae*, sub-tribe *Eragrostae*, and genus *Eragrostis* (Assefa *et al.*, 2011). The genus *Eragrostis* comprises about 350 species from which only tef is cultivated for human consumption. Tef is autogamous small cereal crop widely cultivated in the Horn of Africa (Eritrea and Ethiopia) for over 2000 years (Ponti, 1978). It is entirely cultivated only in Ethiopia as food crop and distributed to several other countries in the 19th century. This wild relative is the closest relative of the cultivated tef. *Eragrostis pilosa* (*E. pilosa*) is also an allotetraploid and has a karyotype similar to tef (Tavassoli, 1986). These two species are morphologically similar. The only known consistent morphological distinction between *E. pilosa* and tef is spikelet shattering of *E. pilosa*. The multi-floret and spikelets of *E. pilosa* readily break apart at maturity as a means of natural seed dispersal (Phillips, 1995). Tef is a C4, self-pollinated annual grass, 40 – 80cm tall. It has a shallow fibrous root system with mostly erect stems, although some cultivars are bending or elbowing types. Its sheaths are smooth, glabrous, open and distinctly shorter than the internodes. According to Hailu *et al*. (1990) has a panicle type of inflorescence showing different forms from loose to compact, the latter appearing like a spike. The flowers of tef are hermaphroditic with both the stamens and pistils being found in the same floret. Florets in each spikelet consist of three anthers, two stigmas and two lodicules that assist in flower opening. Its grain is tiny with 0.9 – 1.7 mm long and 0.7 – 1mm wide and its color varies from white to dark brown (Tadesse, 1975).

2.2 Use and Socio-Economic Values

A food with good nutritional value refers to that composes of most essential nutrients in a fairly balanced manner. The nutrient composition variations may be derived from either of genetic or environmental factors. White tef is preferred for making injera. However, in the rural areas and elderly believe that red tef is better in nutritive values than the white varieties. Tef grain has very good nutritive value, with grain protein content, 10-12% (Gamboa *et al.,* 2008).Though others reported up to 14% (Sadik *et al*., 2013). Besides containing protein and energy, it is a good source of minerals, particularly iron and zinc, which are the most limiting minerals in developing countries. It has very high Calcium, Phosphorus, Copper, Aluminum, Barium and thiamine (Yigzaw *et al.*, 2001, Tadesse, 2017). The proximate values of tef were reported to be 9.4% - 13.3% protein, 73.0% carbohydrate, 1.98% - 3.5% crude fiber, 2.7% - 3.0% ash, and 2.0% - 3.1% fat (Bultosa *et al.*, 2004). It also contains excellent amino acid composition Histidine, Phenylanin- e, Leucine, Isoleucine, Methionine, Valine,

Threonine, Thiamine (B1) and Riboflavin (B2) and higher lysine contents than wheat or barley (Gamboa *et al*., 2008). Some scientists think that the high iron content of tef is due to contamination from iron rich soil ground on the external surface of the grains (Bultosa *et al*., 2004). Tef has high socio-cultural values and high price among cereal crops considered as a formal and acceptable family food in the cities, central, and northern part of the country. The color of injera broadly shows the type of tef varieties as aforementioned. The whitest varieties are most appreciated irrespective of their nutritional compositions (Yu *et al*., 2007). Tef grains are low in the glycemic index, which makes them suitable for people with Type 2 diabetes. This, in particular, attracts individuals who suffer from gluten intolerance or celiac disease (Davison and Laca, 2009). Other utilizations of tef include local alcoholic beverages called *tela* and *katikala*, and porridge (Abraham, 2015). Additionally, tef plant residues could be used as fodder for livestock, and often incorporated as construction materials (Cheng *et al*., 2017). Tef is an economically superior commodity in Ethiopia. It often commands a market price two to three times higher than maize, the commodity with the largest production volume in the country (FAO 2015) thus making tef an important cash crop for producers (Abraham, 2015). Outside Ethiopia, global consumers following the super-food wave are willing to pay premiums for tef (Zhu, 2018). Tef-based products are developed to capture the premium market in the form of bread, porridge, muffin, biscuit, cake, casserole and pudding. Concurrently, interests in tef cultivation are spreading to other parts of the world. Those countries include Australia, Cameroon, Canada, China, India, Netherlands, South Africa, the UK, Uganda and the USA (Abraham, 2015). However, comprehensive statistics on the tef production, utilization and trade are little available in those countries. Tef was first introduced to the USA in the 1980s (Crymes, 2015). Their main purpose of production is forage for horses, cattle and other livestock. Another purpose is for the large Ethiopian Diasporas communities in the USA (Laca, 2009).

2.3 Production and Research Constraints

Ethiopia is the largest tef producer in the world. In 2017, tef accounted for 24% of the grain area, followed by maize 17% and sorghum 15%. Amhara and Oromia are the two major regions, and collectively, the two regions account for 85.5% of the tef area and 87.8% of the tef production. Tef is popular because it is a low risk crop (Minten, 2016). It can be harvested two to five months after sowing. It is relatively resistant to many biotic and abiotic stresses (FAO, 2015). Hence, it adapts to a range of growing conditions where major crops may fail (Zhu, 2018). However, the overall tef production in Ethiopia is at a undeveloped stage. Shattering, low genetic potential of the local varieties, poor pre and post-harvest agronomic practices are major yield limiting factors on the other hand,

Adoption of improved tef seeds is low, farm plots are fragmented, mechanization is almost absent, harvest loss is substantial, and public investment in research is limited (Minten, 2016).

2.4 Genetic Diversity Study

Genetic diversity is the variation of heritable characteristics present in a population of the same species (Smith and Smith, 1989). It includes all the variability occurring among different genotypes. Diversity is the base for survival of plants and provides opportunity for breeders to develop new and improved cultivars (Bhandari *et al.*, 2017). Assessment of genetic diversity provides critical information about substantial genetic divergence and serves as a platform for specific breeding objectives. Measuring diversity exposes the genetic variability in diverse populations and provides justification for introgression and ideotype breeding programmes to enhance crop performance (Mustafa *et al.*, 2011). According to Bhandari *et al.* (2017) diversity analysis provides the understanding of genetic relationships among populations and can be carried out using morphological, cytological, biochemical and molecular characterization. Systematic genetic characterization and well-defined genetic relationships among plant genetic resources are fundamental for effective selection and conservation. This would allow for the identification of genetically unrelated and agronomically complementary genotypes for designed crosses and improved selection for important traits.

Molecular markers are more efficient in germplasm characterization than phenotypic or biochemical markers (Aremu, 2011; Jonah *et al*., 2011; Ranade and Yadav, 2014; Jingura and Kamusoko, 2015). Various molecular marker systems were used effectively to assess the genetic relationships and patterns of association among plant genetic resources. These included restricted fragment length polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Sequence Characterized Amplification Regions (SCARS), simple sequence repeat (SSR) or microsatellites and single-nucleotide polymorphism (SNP) (Datta *et al.*, 2011; Govindaraj *et al*., 2015). The simple sequence repeat (SSR) markers have been extensively used in genetic diversity analysis of different cereal crops including rice (Ishak *et al*., 2015; Sarma *et al*., 2015), wheat (Bafghi *et al*., 2014; Hamdalla, 2014; Drikvand *et al*., 2015), sorghum (Beyene *et al*., 2014), and barley (Hua *et al*., 2015). These markers are known for their reproducibility, multi-allelic nature, co-dominant inheritance, relative abundance, and good genome coverage (Gyulai *et al*., 2012; Senan *et al*., 2014; Ali *et al*., 2015).

2.5 Genetic Markers Used in Crop Diversity Study

A variety of molecular techniques are available for detecting genetic diversity and relationship among cultivated crops. These markers have opened up new possibilities for genotyping analysis and provide

new tools for the efficient use of genetic resources. Genetic markers are measurable inherited genetic variations used to study and understand the genetics of organisms. Genetic markers are widely used by breeders and conservationists to study genetic diversity and assist crop improvement. There are different types of genetic markers which help for identification, characterization and quantification of natural population. Genetic markers may differ in genomic abundance, level of polymorphism, locus specificity, reproducibility, technical requirements, financial investment and degree of information provided (Parker *et al.,* 1998 and Rao and Hodgkin, 2002). Morphological, biochemical and molecular marker systems are the most commonly used genetic markers in diversity study and crop improvement (Schlotterer, 2004).

2.5.1 Morphological markers

Morphological markers are a marker system based on phenotypic appearance; it is the earliest and oldest methods considered as the first step in the genetic diversity assessment. They are inexpensive, simple to score and easy to apply for estimating genetic diversity (Aremu, 2011 and Fu Yong Bi, 2015). Despite their relevance, morphological markers are highly influenced by environmental factors, controlled by many genes and affected by growth stage that makes difficult to measure. They lack adequate genome coverage because of limitation of the number of markers and problems of dominance (Karp *et al.*, 1997).

2.5.2 Biochemical markers

Isozymes are the oldest biochemical markers used to study variations. They are derived from study of the chemical products of gene expressions. The assays require the extraction of proteins from plant tissues, followed by their separation using gel electrophoresis (Schulman *et al.*, 2004). Even if they are not plenty as DNA markers which under estimate the level of genetic diversity; they are reliable markers. Because their consistency in expression of diversity is irrespective of environmental factors (Kumar *et al.*, 2009). Although protein markers avoid the effects of environment, they have the drawbacks of a limitation in the number of detectable isozymes as well as tissue and developmental stage specificity. Hence, protein markers provide limited insight into the true genetic diversity of the pathogen population. These all drawbacks necessitated the development and use of DNA marker systems (Hamrick *et al.*, 1979). DNA marker systems, which were introduced to genetic analysis in the 1980s, have many advantages over the traditional morphological and protein markers that are used in genetic and ecological analyses of plant populations: firstly, an unlimited number of DNA markers can be generated; secondly, DNA marker profiles are not affected by the environment, and, thirdly

DNA markers, unlike isozyme markers, are not constrained by tissue or developmental stage specificity (Park *et al*., 2009).

2.5.3 Molecular marker

They are identifiable DNA sequences, found at specific locations of the genome, and transmitted by the standard laws of inheritance from one generation to the next. Molecular markers can be thought of as constant landmarks in the genome. Various molecular techniques existe; which are different in their principles and methodologies (Semagn *et al.,* 2006 and Kumar *et al.,* 2009). They are superior to both morphological and biochemical markers because they are relatively abundant throughout the genome, independent of environmental conditions and can be detected at any development stage of the plant (Mondini *et al.,* 2009). DNA based markers provide very effective and reliable tool for measuring genetic diversity in crop germplasm and studying evolutionary relationships. According to the approach employed in the classification, these markers are grouped as methodical principle, the aim of use, or the type of polymorphism (Gupta *et al*., 1999). Those DNA markers are divided into three classes: (a) Hybridization-based DNA markers (e.g., RFLP); (b) PCR-based DNA markers (e.g., RAPD, AFLP, SSR, and ISSR) and (c) and DNA chip sequencing-based DNA markers (SNP).In this classification, the markers are grouped on the basis of the principal method used for analysis. The classification proposed by Gupta *et al*. (1999), also reflects the "evolution" of DNA markers. The first class represents the first generation of DNA markers, which were widely used in the 1980s; then in the 1990s, PCR markers (the second group) gained key importance; and since 2000, these were succeeded by a new generation SNP markers, based on microarray technologies (the third group).DNA markers can be classified according to the aim of their use: for tagging an individual locus (RFLP, SSR, and SNP) or genome fingerprinting (RAPD, AFLP, ISSR) multilocus markers based on Southern hybridization (Kumar *et al.,* 2009). To put it differently, this classification assigns markers to mono-locus (as a rule, one locus per the diploid genome is present) and multilocus (which amplify numerous fragments of genomic DNA) ones.

2.5.3.1 Hybridization-Based (non-PCR) markers

Restriction Fragment Length Polymorphism (RFLP) markers

Restriction Fragment Length Polymorphism (RFLP) was first used as a tool for genetic studies in 1974 (Botstein *et al.,* 1980). The RFLP markers are based upon hybridization of a probe, a specific DNA sequence designed to hybridize with genomic DNA and detect a target sequence (s) in the unknown sample, on nylon membrane containing genomic DNA which has been digested with restriction enzymes. Fragments produced by restriction enzyme digestion are separated by gel electrophoresis and

transferred onto nylon membrane by methods known as Southern blotting. Differences in the sequence at or around the sequence to which the probe hybridizes may result in differences (polymorphism) in the length of the fragment detected by the probe (Grant and Shoemaker 1997). RFLP is a co-dominant marker (allelic information) which has several advantages such as no sequence information is required, unlimited in the number of loci to be detected and robustness in usage. However, this technique requires a relatively large amount of DNA in a form that is able to be digested with restriction enzymes, is labour intensive and fairly expensive, displays low levels of and is time consuming (Gupta *et al.,* 1999).

2.5.3.2 PCR-Based markers

A wide variety of PCR-based marker techniques has been developed during the last decade, each with various advantages and different shortcomings (McGregor *et al.,* 2000). These PCR-based markers differ in principle, application, complexity, informativeness, cost and time requirements (Mignouna *et al.*, 2003). The choice of marker system could often depend on the crop investigated (Russell *et al.*, 1997).

2.5.3.2.1 Random Amplified Polymorphic DNA (RAPD) Markers

Amplification of genomic DNA using single primers of arbitrary nucleotide sequence, in low stringency conditions, results in multiple amplification products from loci distributed throughout the genome (Williams *et al.,* 1990). RAPD markers became popular because of their simplicity, applicability to any genome, no sequence information requirement, relatively small DNA quantities required, results obtained quickly and high genomic abundance. Despite this, the limitations of RAPDs are numerous and include: they are dominant markers (i.e. cannot distinguish homozygotes from heterozygotes), are sensitive to laboratory changes and have low reproducibility within and between laboratories (Rafalski, 1997). The number and pattern of bands amplified can be affected with variation in template concentration and with annealing, extension and denaturing time (Bielawski and Noack, 1995) and template quality (Micheli *et al.*, 1994). The original RAPD marker analysis has been modified by using restriction enzymes to cut the genomic DNA before the amplification. Riede *et al*. (1994) and Koebner (1995) showed that in wheat, restriction digestion of genomic DNA with different endonucleases before PCR amplification improved the level of polymorphism and reduced non-specific amplification. The results were probably due to a greater efficiency in primer annealing along shorter DNA fragments where a simplified secondary DNA structure is less likely to interfere with the process.

2.5.3.2.2 Amplified Fragment Length Polymorphism (AFLP) Markers

Amplified Fragment Length Polymorphism (AFLP) is based on the selective amplification of sets of restriction fragments from genomic DNA (Vos *et al.*, 1995). DNA is first cut with restriction enzymes and double-stranded adaptors are then ligated to the fragments. PCR primers designed to match the sequence of the adaptors are used to amplify the fragments. The number of fragments amplified is intentionally limited by including nucleotides at the 3' end of the primer that extend into the unique sequence of the DNA fragment (Hoelzel and Green, 1998). The key feature of AFLP-PCR is its capacity for the simultaneous screening of many different DNA regions distributed throughout the genome (Mueller and Wolfenbarger, 1999).The great advantages of AFLPs are the large number of loci uncovered in a single assay (Glaubitz and Moran, 2000), high reproducibility (Russell *et al*., 1997), no requirement for DNA sequence information (Vienne *et al*., 2003). AFLPs and microsatellites are both likely to remain key molecular tools for the analysis of genetic variation (Mueller and Wolfenbarger, 1999).

2.5.3.2.3 Simple Sequence Repeats (SSRs) Markers

Microsatellites are also known as simple sequence repeats (SSRs) and consist of segments of DNA containing tandem repeats typically of simple motif sequences of one to six bases. These microsatellite repeats are often flanked by unique sequences, occurring only once in the genome (Glaubitz and Moran, 2000). Microsatellites were first developed for use in the human genome (Weber and May, 1989) and were later found to be abundant in plants (Morgante and Olivieri, 1993). They are useful because they are highly polymorphic markers, i.e. they often have many alleles at each locus with each allele having a different number of tandem repeats (Beckman and Weber, 1999) and length polymorphism can be revealed by electrophoresis. Microsatellites have been particularly useful for comparative genetics and genomic mapping since microsatellite sequences show a high degree of length polymorphism, abundance and distribution throughout the genome (Oh and Mao, 1999). The relative abundance, multi-allelic nature, co-dominant inheritance, high reproducibility and good genome coverage of microsatellites make them powerful molecular markers for use in genetic studies (Glaubitz and Moran, 2000). Also, SSR markers are very robust and user-friendly when compared with other marker systems and can be multiplexed during PCR or gel loading (Godwin *et al*., 2001). Furthermore, once SSR markers have been developed for use in one species it may be possible to use them in related species since the flanking regions are rather conserved and the repeats are relatively variable. The greatest challenge in using SSR markers is their discovery and initial development requiring species-specific DNA information for each marker locus in order to design primers from flanking regions (Godwin *et al*., 2001). There are observations that primers developed in one species

can amplify microsatellite loci of related species (Hormaza, 2002). SSR has been used in the construction of a genetic linkage map of soybean (Akkaya *et al*., 1995), corn (Taramino and Tingey, 1996) and sorghum (Taramino *et al*., 1997). It was used in genetic diversity analysis of rice (Cho *et al*., 2000), barley (Struss and Plieske, 1998), wheat (Prasad *et al*., 2000), sorghum (Smith *et al*., 2000), soybean (Tanya *et al*., 2001), potato (McGregor *et al*., 2000), ryegrass (Kubik *et al*., 2001), apricot (Hormaza, 2002), peach (Aranzana *et al*., 2003), carnation (Smulders *et al*., 2003) and coffee (Anthony *et al*., 2002). SSR reported that polymorphism is consistent with variation patterns revealed by the species` breeding history. Results of several investigations concluded SSR as a marker of choice for genetic diversity analysis of organisms with a narrow genetic base (Westman and Kresovich, 1999).

2.5.3.2.4 Inter simple sequence repeats (ISSR) Markers

The Inter-SSR (ISSR) also called MP-PCR (microsatellite-primed PCR) marker system was first developed by Zietkiewicz *et al*. (1994). The ISSR technique is a PCR based method, which involves amplification of DNA segments present at an amplifiable distance in between two identical microsatellite repeat regions oriented in opposite direction (Reddy *et al*., 2002). Inter-SSRs are semi-arbitrary markers amplified by PCR in the presence of one or two primers complementary to a target microsatellite. Similar to RAPD markers, ISSR markers do not require genome sequence information but provide multi-loci amplification (Bornet and Branchard, 2001). ISSR markers are as quick and easy to handle as RAPD markers, but their reproducibility is still uncertain. Some reports showed that ISSR markers are reliable in wheat (Nagaoka and Ogihara, 1997) and in cotton (Liu and Wendel, 2001) and may have the reproducibility of SSR markers because of the longer primer length (Bornet and Branchard, 2001). This means high annealing temperature applied can result in higher stringency (Gupta *et al*., 1999). However, several studies have indicated that, despite the higher annealing conditions used, these amplifications behave similarly to RAPD in terms of susceptibility to slight changes in reaction conditions such as annealing temperature, primer concentration and magnesium concentration. All these changes affect the quality of the banding patterns (Weising *et al*., 1995).

2.5.3.3 DNA chip sequencing-based DNA markers

Single Nucleotide Polymorphism (SNP) markers

SNP is defined as a nucleotide site, for which a high substitution rate has been shown among individual samples in a population (Gupta *et al*., 1999). SNPs which belong to the last-generation molecular markers occur at high frequencies in both animal and plant genomes. The development of SNP markers allows automatizing and enhancing ten folds the effectiveness of genotype analysis. These markers are compared to other DNA markers, in order to ensure adequate choice of marker type for solving various

molecular genetics problems. The restriction posed on frequency distinguishes SNPs from rare point mutations and implies the use of the former as genetic markers. Based on this definition, SNP markers formally do not include single-nucleotide insertions/deletions (indels), though some authors combine these two polymorphism types with single-nucleotide substitutions under the common term of SNPs (simple nucleotide polymorphisms) (Wang *et al*., 1998). Several types of SNPs are distinguished, according to their assignment to the structural element of genomic DNA or their functional effect (Coulondre *et al.,* 1978). In contrast to multi-allelic markers, analysis of biallelic SNP markers can be practically fully automated. Several thousands of SNPs can be analyzed simultaneously by application of DNA microarrays. SNP markers can be used in projects of different financial level. SNP is the most common type of DNA polymorphisms. High occurrence of SNP loci in the genome opens a possibility to develop on their basis molecular genetics maps of very high density, required for isolating and studying genes for resistance to various diseases as well as other agronomically valuable plant genes.

3. MATERIALS AND METHODS

3.1 Study Location and Plant Materials

A total of 64 tef accessions were used in this study (Table 1). The entire accessions were obtained from Ethiopian Biodiversity Institute (EBI). They were collected from different agro-ecologies varying in altitude, soil type, rainfall and temperature (Table 1). The study was carried out at the National Agricultural Biotechnology Research Center (NABRC) and Debre Zeit Agricultural Research Center (DZARC). The genetic diversity of tef accessions was conducted using 12 morphological parameters and 10 SSR markers. For the greenhouse experiment Completely Randomized Design (CRD) with two replications were used. Pot size of 13 cm diameter was filled with light soil and watered one day prior to seed sowing. Recommended amount of fertilizers (40 kg/ha N and 26 kg/ha P) were applied to all pots. Hand weeding and other agronomic practices were done as required.

Table 1. Passport data of the tested accessions and their geographical origins

Acc. No.	Local name	Zone	Woreda/district	Altitude (m.a.s.l.)
15309	Tef	Semen Shewa	Ataye(efeson)	1443.00
236957	Magna	Semen Shewa	Minjarna Shenkora	1700.00
236959	Magna	Semen Shewa	Minjarna Shenkora	1750.00
236960	Tekur tef	Semen Shewa	Minjarna Shenkora	1780.00
236961	Tikur tef	Semen Shewa	Minjarna Shenkora	2020.00
243493	Tikur tef	DebubWello	Tenta	2060.00
243502	Ayiro tef	DebubWello	Ambasel	1750.00
243507	Sergenga tef	Semen Wello	Habru	2020.00
243508	Manga tef	Semen Wello	GubaLafto	1835.00
243509	Fenkilew	Semen Wello	GubaLafto	1835.00
243529	Key tef	Semen Gondar	AddiArkay	1310.00
243535	Nechtef	Semen Gondar	Gondar zuria	1920.00
243536	Sergegna	Semen Gondar	Gondar zuria	1920.00
243537	Key Fesho tef	Semen Gondar	Gondar zuria	1920.00
243538	Nechtef	Semen Gondar	Gondar zuria	1920.00
243539	Key tef	Semen Gondar	Gondar zuria	2050.00
243540	Nech tef	Semen Gondar	Gondar zuria	2050.00
243544	Key tef	Semen Gondar	Gondar zuria	2050.00
243545	Deyabo tef	Semen Gondar	Gondar zuria	2050.00
243547	Nech tef	MirabGojam	Bahirdar zuria	1870.00
243548	Nech tef	MirabGojam	Bahirdar zuria	1870.00
243549	Key tef/Anshar	MirabGojam	Bahirdar zuria	1870.00
243550	Key laba tef	MirabGojam	Bahirdar zuria	1870.00
243551	Sergenga	MirabGojam	Bahirdar zuria	2070.00
243553	Key tef	MirabGojam	Bahirdar zuria	1815.00
243512	Tseada tef	Mehakelegnaw	Kola Temben	1990.00
243517	Keyih tef	Mehakelegnaw	Kola Temben	1990.00
243518	Gofgafe tef	Mehakelegnaw	Kola Temben	1990.00
243519	Tseads Dalga	Mehakelegnaw	Kola Temben	1990.00
243520	Murenayi	Mehakelegnaw	Kola Temben	1990.00

243521	Keyih tef	Mehakelegnaw	Kola Temben	2010.00
243524	Tseas tef	Mehakelegnaw	Adwa	2030.00
243525	Sergen tef	Mirabawi	Tselemti	1260.00
243526	Tseads tef	Mirabawi	Tselemti	1260.00
243527	Tseada tef	Mirabawi	Tselemti	1230.00
243528	Sergen tef	Mirabawi	Tselemti	1230.00
244789	Dalesha	Kembata	Alaba	1773.00
244790	Dorta	Kembata	Tembaro	1764.00
244791	Mitazu and Dorta	Kembata	Tembaro	1808.00
244792	Metazo	Kembata	Kachabira	1717.00
244793	Sergenagashe	Kembata	Kachabira	1714.00
244794	Yemushe	Kembata	Kachabira	1714.00
244796	Dalesha	Kembata	Alaba	1773.00
244814	Xafi	MisrakWellega	Bilaseyo	1520.00
244815	Taffee	MisrakWellega	Bilaseyo	1976.00
244816	Taffee	MisrakWellega	Bilaseyo	1946.00
244817	Tafii dima	MisrakWellega	Bilaseyo	1946.00
244818	Tafii Adi	MisrakWellega	Bilaseyo	1973.00
244821	Taffee Adi	MisrakWellega	Bilaseyo	1952.00
244822	Taffee Dima	MisrakWellega	Bilaseyo	1910.00
244850	Taffee Adi	MisrakWellega	Amurujarte	1916.00
244851	Taffee Dima	MisrakWellega	Amurujarte	1919.00
244852	Taffee Adi	MisrakWellega	Amurujarte	1876.00
244853	Taffee Adi	MisrakWellega	Amurujarte	2063.00
244854	Taffee Dima	MisrakWellega	Gidakiremu	2042.00
244856	Taffee Adi	MisrakWellega	Gidakiremu	2057.00
244876	Taffee Adi	Illubabor	Bedele	1984.00
244877	Taffee Adi	Illubabor	Bedele	1985.00
244878	Taffee Dima	Illubabor	Bedele	1925.00
244879	Taffee Adi	Illubabor	Bedele	1915.00
244880	Taffee Adi	Illubabor	Bedele	1887.00
244881	Taffee Adi	Illubabor	Bedele	1899.00
244882	Taffee Gomejen	Illubabor	Gechi	1903.00
244883	Taffee Adi	Illubabor	Gechi	1900.00

(Source: Ethiopian Biodiversity Institute (EBI)

3.2 Data Collection

3.2.1 Phenological Traits

The following phenological traits were considered during the study.

1) Days to seedling emergence (DSE):-The number of days from the date of sowing to the date of 50% seeding emergence.

2) Days to heading (DH):-The number of days from the date of seedling emergence to the date when 50% of the plants in each plot have panicles tips emerged (about 5 cm) on the main shoot.

3) Days to physiological maturity(DPM):- The number of days from the date of seedling emergence up to the date when 90% of the plants in each plot attained physiological maturity as judged by visual observation of the turning of the vegetative parts to yellow (straw) color.

3.2.2 Yield and Growth Related Traits

1) Plant height (PH) (cm):- The height of five randomly selected plants from the central rows of each plot was measured from ground level to the tip of the main shoot panicle at maturity.

2) Panicle length (PL) (cm):- The average length of the main shoot panicle of five randomly selected plants from the central rows of each plot measured from the end of the peduncle to the tip of panicle at maturity.

3) Peduncle length (PDL) (cm):- The average length of the last apical internode of the main shoot culm of five randomly selected plants from the central rows of each plot measured from the last culm node up to the place where the panicle starts.

4) Culm length (CL) (cm):- The average length of the main shoot culm measured from the ground level (base) of the plant to the point where the panicle starts for five randomly selected plants from the central rows of each plot.

5) Number of total tillers per plant (TTPP):- It was recorded as the number of all tillers prod- uced per plant assessed on five randomly selected plants from the central rows of each plot.

6) Number of fertile tillers per plant (FTPP):-It was recorded as the number of panicle bearing (fertile) tillers per plant assessed on five randomly selected plants from the central rows of each plot during physiological maturity stage.

7) Hundred-seed weight (mg):- The mass/weight of 100 seeds collected from the bulk of seeds harvested from all plants in each plot and adjusted to a moisture content of 12.5%.

3.3 DNA Extraction

Ten seeds from each accession were grown in a greenhouse and fresh leaves were collected from 20 days old plant for gDNA extraction. About 20 days after sowing, when three to four leaves developed (Figure 1a), seedling leaves were detached and placed under silica gel for one week. After the leaves were dried, about 50 mg of the dried leaves were placed in 2 ml autoclaved Eppendorf tubes and grinded using Geno Grinder (MM-200, Retsch) with 25 rpm/s for 5 minutes. Then genomic DNA was extracted by modified cetyltrimetyl ammonium bromide (CTAB) method as described by Wang *et al*. (1996). Even if the original protocol lacks the use of RNase; in order to increase the quality of gDNA 2µl of RNase was added to each sample and incubated at 37°C for 30 minutes.

The quality and quantity of gDNA was assessed using 0.1 % agarose gel and 1xTAE buffer using a lambda standard ladder, (100 µg) as reference band. The DNA bands were visualized using gel documentation, 3UV-trans illuminator (Bio-Doc). The quantity and purity of gDNA was measured by Nano drop spectrophotometer (ND-8000, Thermo scientific) (Figure 1 a and b). Samples with high band intensity, purity with 1.8 to 2 were selected for further PCR analysis.

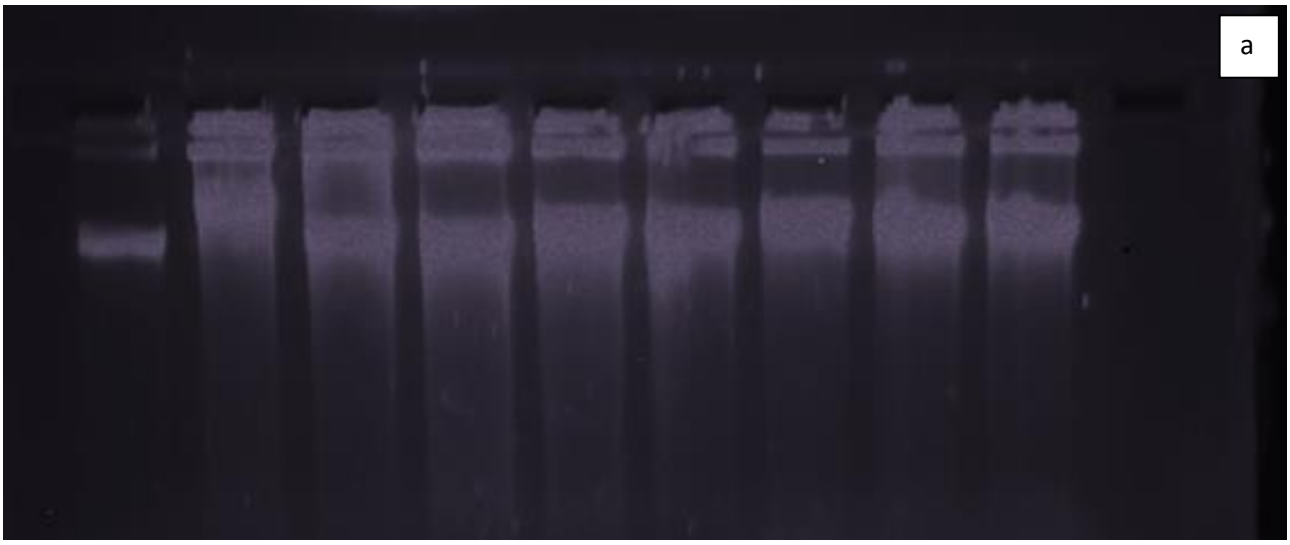
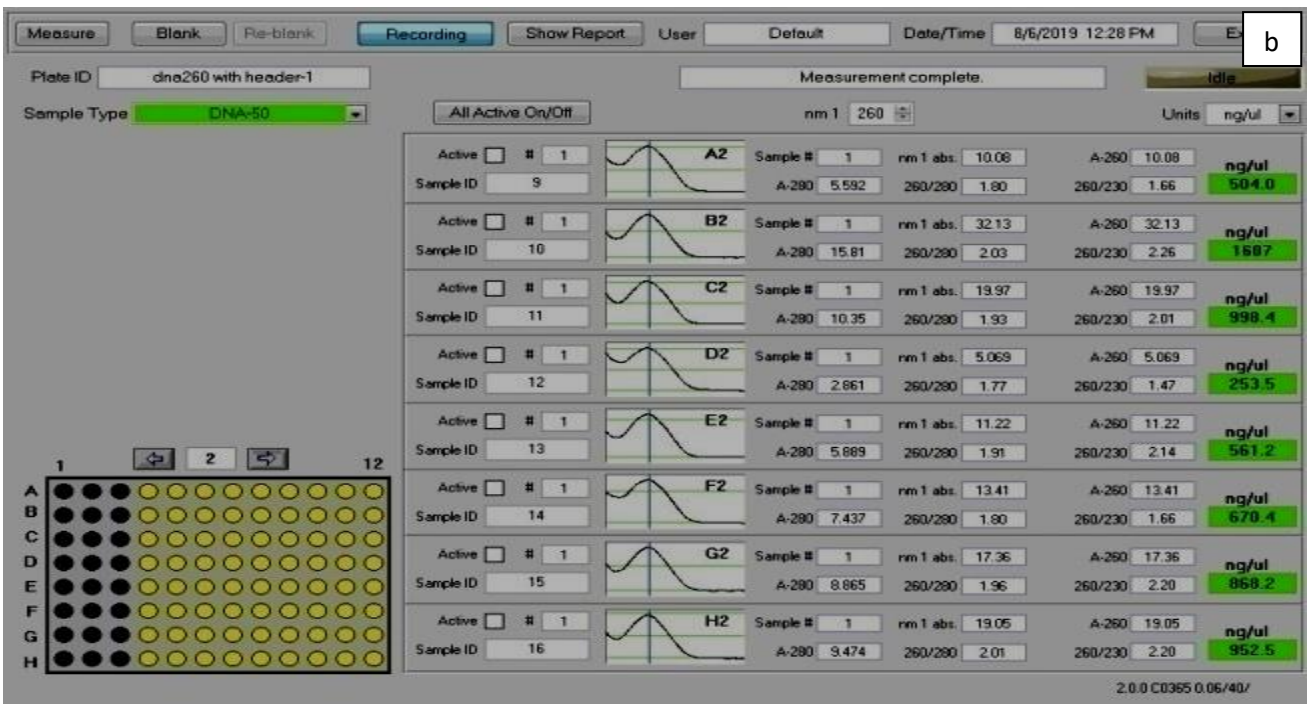


Figure 1. a) Genomic DNA image



b) Nano drop spectrophotometer

3.4 Primers Screening and PCR Optimization

A total of 15 SSR primer pairs were used for initial screening and PCR amplification. The markers were developed by Zeid, (Zeid *et al*., 2011) and the sequences were extracted from the supplementary materials. Screening of the SSR primers were done on 20 representative tef accessions. Out of the tested fifteen primers, 10 SSR primers were selected for final analysis based on their level of

polymorphism and specificity to the target loci (Table 2). For further PCR amplification different protocols were tested and the best amplification profiles were selected.

Table 2. Locus name, expected size, annealing temperature and sequence of the 10 selected microsatellite primers used in this study

SSR Markers	Expected size (bp)	Annealing Temperature(°C)	Forward primer sequence	Reverse primer sequence
CNLTs2	160-260	58.2	CAGCAGGGAGAGAGAGGAGA	GGGGTCAAGTTATTGCTTAGAGAA
CNLTs5	250-310	57.2	CCCAAAGTGATGCAAAAACA	GATAGATAGAGACACAGACACACACA
CNLTs6	90-260	58.2	AATTCGCAGCTGATCTACGC	CTCGTCGATATACGTGCAAAA
CNLTs19	180-300	58.2	CATTTCTTGCTGCTGGATCA	AGTATGGTGGCCTTGGTGAG
CNLTs22	100-260	55.7	CATGCTCGTTCAGAGTCCAA	GGGGGATCTAGGAGAGAGAGA
CNLTs24	135-180	55.7	GAGAGCGGTTTTGTCCTACG	GGAATAGGGAGGCGAGGTAG
CNLTs25	150-320	57.2	TTGGAATGAGATGGCATTTG	GAAGCGGGGTAAGATTTGAA
CNLTs458	130-250	59.1	AACAAGAACCACACAACA	AAAGGAACCCACAGGGGTAA
CNLTs461	220-320	63.9	GTCTTGATGGTGGCGGAATAG	TCATCATCCTGCTCGAATCA
CNLTs463	190-260	59.1	TGCTAGGATGGTCCTGTTGAG	AGCACCAAATCCCTATGCAC
CNLTS27	200-220	58.2	TTGGAATGAGATGGCATTTG	GAAGCGGGGTAAGATTTGAA
CNLTS460	190-220	59.1	TCAGACGAGCACAAGTTTGG	GACGGTGTGTTGACGAGAACT
CNLTS464	250-280	58.2	GTCTTGATGGTGGCGGAATAG	TCATCATCCTGCTCGAATCA

(Source: New England Bio Labs)

3.5 PCR Amplification and Gel Electrophoresis

Polymerase chain reaction (PCR) based simple sequence repeat (SSR) marker system was used to investigate the relationships among sixty four accessions of tef. PCR amplifications were carried out in 96-well plates (Eppendorf Master Cycler Pro S) as per the protocol suggested by Williams *et al*. (1990). The PCR reaction was done in 25 µl reaction mixture composed of 6.25µl 1X One Taq Master mix (New England Bio Labs), of each forward and reverse primers, 0.25µl Dimethyl- sulfoxide (DMSO, Fisher Scientific), 2 µl nuclease free water and 2 µl 50ng/µl template genomic DNA, respectively. Amplification condition was programmed to initial denaturation at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 1minute, optimized annealing temperature based on the individual primer requirement (55.7- 63.9°C) for 1minute, primer extension at 72°C for 2 minute which was followed by final extension at 72°C for 10 minutes and holding temperature at 4°C. Amplified DNA fragments were separated out on a 3% agarose gel electrophoresis using 1× TAE buffer. The samples were mixed with 3 µl of 6X loading dye containing gel red, then loaded on gel with 6.5 µl of SSR amplification product and run for 3 to 3:30 hrs at 100 constant voltages. The

molecular weight of each amplified product was estimated by comparing the DNA bands with SMOBio (100-1500bp) DNA ladder loaded in the peripheral wells of the gel on the side of the sample wells. The patterns were photographed using the Gel documentation system (Bio Doc Imaging System) and stored as digital pictures.

3.6 Data Analysis

3.6.1 Analysis of Morphological Diversity Parameters

3.6.1.1 Analysis of Variance

After checking for outliers and normality of residuals, the morphological data were subjected to analysis of variance (ANOVA) using the SAS statistical package version 9.3 (SAS Institute, 2010) following SAS statement for completely randomized design for single location.

3.6.1.2 Cluster Analyses

Before proceeding to the multivariate analysis; the average mean value of genotypes was standard - ized by subtracting the mean and dividing by standard deviation. Clustering of tef accessions were performed by average linkage clustering method. The number of clusters were determined by the points where local peaks of the pseudo F statistic (Calinski and Harabasz, 1974) join with small values of the pseudo t2 statistic (Duda and Hart, 1973) followed by a larger pseudo t2 for the next cluster fusion was examined to decide the number of clusters (SAS Institute, 1996). This assumption gives very accurate results in determining the number of clusters. The method is conservative and appropriate for hierarchical clustering methods (Milligan and Cooper, 1985). The dendrogram was generated using MINITAB version 16.

3.6.1.3 Principal Component Analysis

Principal component analysis (PCA) was performed in order to highlight the resolving power of the ordination. PCA, using correlation matrix of SAS version 9.3, was computed to project the data into lower dimensions and to show genetically similar genotypic clusters. It was also used to show those traits, which were responsible for the significant variations in tef accessions. The characters with larger absolute values closer to unity within each principal component influence the clustering more than those with lower absolute values closer to zero (Chahal and Gosal, 2002). In this analysis, only PCs with eigen values greater than one are considered as important. As suggested by Johnson and Wichern (1988) trait coefficient or eigenvector greater than half divided by the standard deviation (square root) of the Eigen value of the respective PC was employed as general guideline for weighing the relative significance of traits constituting the PCs.

3.6.2 Analysis of Molecular Diversity Parameters

The clear and reproducible alleles amplified by each SSR markers among 64 accessions were scored according to their fragment size (bp) using PyElph1.4 software package (Ana and Cristian, 2012). Bands with the same fragment size were treated as monomorphic (identical fragments). On the other hand, bands with different fragment size were scored as polymorphic. Genetic diversity analyses were carried out on the basis of the scored bands. GenAlEx ver. 6.502 statistical software packages were employed to compute the standard indices of genetic diversity. Locus based diversity indices including major allele frequency (MAF), the number of different allele (Na), gene diversity, polymorphic information contents (PIC), heterozygosity and fixation index (F) were computed using Power marker v3.25 soft ware (Liu and Muse, 2005). Effective number of alleles, Shannon's Information index (I), gene flow (Nm), allelic frequency, observed heterozygosity (Ho), expected heterozygosity (He), number of differ -rent allele(Na), number of effective allele (Ne) and estimate of the deviation from Hardy-Weinberg equilibrium (HWE) over the entire populations were determined using GenAlEx ver 6.502 software (White and peakall, 2015).

$$PIC = 1 - \sum_{□=1}^{□} p^2 ij$$

Where k is total number of alleles detected for SSR marker, Pij is the frequency of j^{th} SSR allele for the i^{th} marker

$$\text{Percent polymorphism (\%P)} = \frac{\text{Number of polymorphic bands}}{\text{Total number of bands}} x100$$

The genetic variability of a population can be estimated by measuring gene and genotypic frequencies. Nei (1973) introduced the concept of gene diversity to describe genetic variability in both sexually and asexually reproducing populations. Gene diversity (H) is defined as the probability of obtaining two different alleles at a locus when two haploid individuals are sampled randomly from a population. This can be calculated by:

$$H = 1- \Sigma Xi^2$$

Where H is the gene diversity of the population and Xi is the frequency of different alleles at a particular locus. The minimum value of gene diversity is 0 and it is when the locus is monomorphic

(only one allele per locus). In dominant markers, the value of gene diversity range from 0 to 0.5 and H= 0.5 when the frequencies of the two allele in particular locus are equal. However, in multi-allelic markers like microsatellites, the maximum value for gene diversity will increase with increasing the number of alleles per locus. For instance, if the number of alleles per locus becomes 3, 4 and 5, the gene diversity will become 0.67, 0.75 and 0.8, respectively. The other important genetic parameter to be considered in assessing the gene diversity is determining how the variability is distributed among and within the different subpopulations. Nei (1973) proposed a method for partitioning the genetic diversity of a population into different components. He suggested that genetic differentiation between subpopulations can be estimated as:

$$GST = (HT - Hs) / HT$$

Where, HT is the total gene diversity of the population, Hs is the average gene diversity within subpopulations and GST is the proportion of the total genetic variation accounted for by variation among sub populations. If the amount of gene diversity within subpopulations is high, but among subpopulations is low, the GST value will be small and vice versa (Nei, 1973). GenAlEx ver. 6.502 software package was used to compute population differentiation test: Wrights fixation index (FST) and pair wise FST at 1000 bootstraps. Rarified allelic richness (Ar) and private rarified allelic richness (Arp) were computed using HP-Rare 1.1 software (Kalinowski, 2005). Analysis of molecular variance (AMOVA) and estimate of the variance components were conducted using GenAlEx ver. 6.502 software (White and peakall, 2015).

To examine the genetic relationship between the different accessions, a genetic dissimilarity matrix was calculated using Jaccard's formula and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) based Neighbor-Joining tree and hieratical clustering (dendrogram) were generated using DARwin ver. 6.0 (Perrier and Jacquemoud-Collet, 2006). Population structure and admixture patterns were determined using STRUCTURE software ver. 2.3.4 based on Bayesian algorithm (Pritchard *et al.*, 2000). To estimate the true number of population cluster (K), a burn-in period of 100,000 was used in each run, and data were collected over 200,000 Markov Chain Monte Carlo (MCMC) replications for K = 1 to K = 10 using 20 iterations for each K. The optimum K value was predicted following the simulation method of Evanno *et al.* (2005) using the web-based STRUCTURE HARVESTER ver. 0.6.92 (Dent and Bridgett, 2012). Bar plot for the optimum K was determined using Clumpak beta version (Kopelman *et al.*, 2015).

4. RESULTS AND DISCUSSIONS

4.1 Genetic Diversity Analysis Using Morphological Markers

Tef is highly diverse and variable in terms of morphological and agronomic characters. The distribution of the crop in different agro-ecological zones coupled with the selection by farmers on the basis of their preferred traits has resulted in a number of varieties with unique characters. Genetic diversity analysis of tef accessions facilitates the development of improved varieties with high productivity and yield stability. From the current study panicle forms like semi-compact, semi-loose, very loose and compact are observed (Figure 2). Similar result was reported by Mengesha *et al*. (1965) which was the first study on phenotypic diversity in tef germplasm.

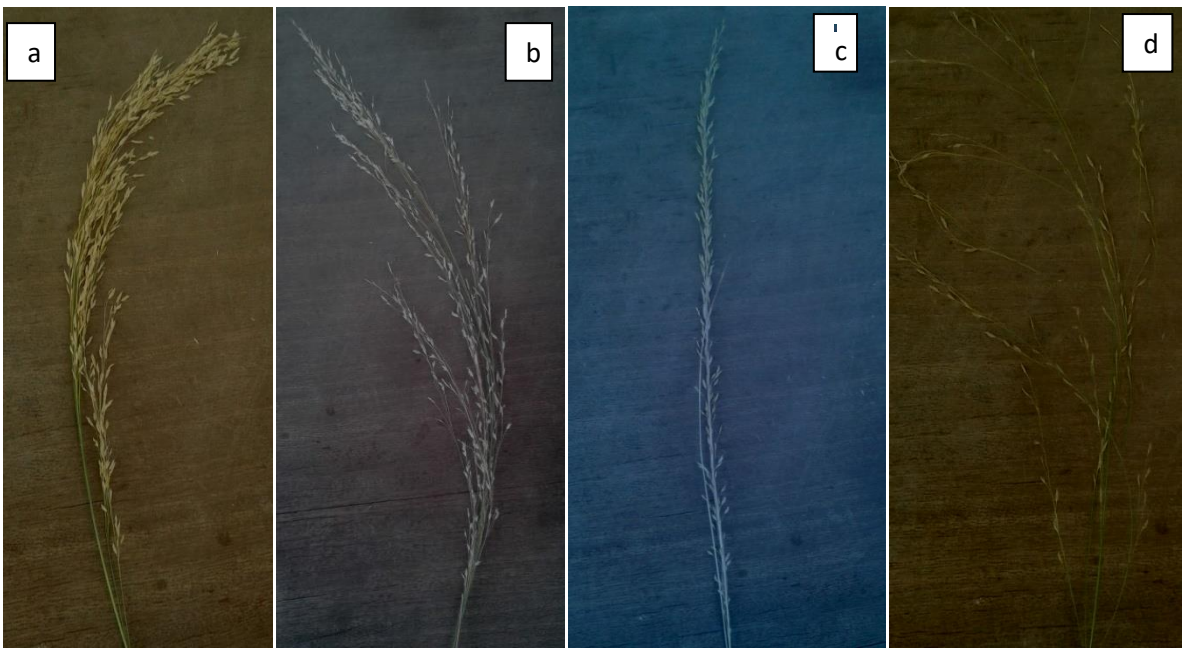


Figure 2. Types of panicle observed in the current study a) Semi-compact, b) Semi-loose, c) Compact d) Very loose

4.1.1 Analysis of Variance for Quantitative Traits

The Analysis of variance (ANOVA) for 12 quantitative characters of the 64 tef accessions showed significant ($P< 0.01$ or $P < 0.05$) differences among the accessions for all the studied traits (Table 3). The mean squares due to genotypes were highly significant ($P< 0.01$) for days to heading, total number of nodes, first culm internode, and second culm internode and significant ($P < 0.05$) for plant height, culm length and peduncle length (Table 3). Where-as, mean square due to genotypes was not significant for the studied traits days to maturity, total tiller, fertile tiller and hundred seed weight. This indicates that existence of considerable genetic variability for selection and hybridization. Similar finding was also reported by Adnew *et al*. (2005) and Assefa *et al.* (2000) from the experiment

conducted to evaluate the performance of tef accessions. Most of the traits showed significant difference ($P < 0.05$) among the tested tef accessions. This might be because of controlled growth condition of the tested accessions.

Table 3. Analysis of variance for 12 traits of 64 tef accessions tested in2019 at Debre Zeit Agricultural Research Center

Traits	MSG	MSE	CV%	R-Square
Days to heading	26.66**	24.25	11	0.52
Days to maturity	22.68^{ns}	39.04	7	0.37
Plant height	93.90*	104.84	12	0.48
Panicle length	27.80^{ns}	34.12	18	0.45
Culm length	37.27*	43.77	14	0.46
Peduncle length-	8.90*	10.15	24	0.46
Total number of node	0.57**	0.36	16	0.61
First culm internode	4.45**	3.11	23	0.59
Second culm internode	8.94**	6.88	21	0.56
Total tiller	0.24^{ns}	0.24	10	0.00
Fertile tiller	0.66^{ns}	0.86	23	0.44
Hundred seed weight	29.25^{ns}	25.71	21	0.46

*, ** and *** Significant at 0.05, 0.01 and 0.001 probability level respectively and ns on-significant MSG= Mean Square due to Genotype, MSE= Mean Square due to error, CV%= Coefficient of variation in percentage.

Genetic variability is the most essential requirement for any crop improvement program. A wide range of variation was observed for almost all the traits in the material under study. The tef accessions evaluated in this study showed significant phenotypic variability. These implied that the underutilized genetic variability of the accessions studied should be future priority in tef improvement and genetic conservation programs. The existence in the tef germplasm lines of broad ranges of mean values between the lowest and highest extremes for each of the 12 traits (Table 3) offers ample chances for the genetic improvement of the crop through selection and recombination of lines with the desired levels of expression. For instance, the range of variability in phenology implies that it is possible to develop tef varieties suitable for the versatile agro-ecological zones of Ethiopia. This result was similar with the previous findings (Assefa *et al*., 2000).They reported significant genetic variation among tef accessions with regards to the evaluated quantitative characters. According to the report by Ebba (1975) significant phenotypic variability was observed among the tested tef accessions collected from different Ethiopian tef growing regions for descriptions of 35 tef accessions were given based on phenology, plant vigor, shoot and root related traits, panicle form, spikelet size, growth habit, and lemma and caryopsis color.

4.1.2 Mean Performance and Range of Parameters

The 64 accessions studied showed wide range of variability for most of the traits evaluated in the study (Table 4). Wide range of variability was observed for plant height ranged from 60cm to 112cm with mean value 86 cm. With respect to days to maturity, it ranged from 67 days to 98 days with mean value of 82.5 days. Higher result range of maturity (84-100.5) report by Kebebew *et al*. (2000). This showed accessions incorporated in the present study have a wide range of variability because of genetic factor and the geographic difference for the tested locations may be the probable reason for the differences. Culm length ranges from 40.1cm to 62cm with mean value of 51.5cm. Accessions number 244814 (114cm), 243535 and 244793 (111cm) showed longest height than the rest and Accessions named 243508 (8cm) and 243507(9 cm) were exhibited the shortest height.

Table 4. Mean performances of 64 tef accessions for 12 traits tested at Debre Zeit Agricultural Research Center in 2019 cropping season

	Traits											
Acc. No.	**DTH**	**DTM**	**PLH**	**PAL**	**CL**	**PDL**	**TN**	**FCIN**	**SCIN**	**TT**	**FT**	**HSW**
15309	46.5a	84a	72.5a	30.9b	43.5a	9.9a	4.1c-f	7.5e-p	13.4f-m	5a	5a	25.35b
236957	40.5a	83a	78.4a	32.9b	48.2a	11.8a	3.9e-h	6.9g-q	10.9o-u	5a	5a	32.75b
236959	42.5a	73.5a	93.7a	36.6b	54.1a	15.3a	3.8f-i	7.7e-n	11.6l-t	5a	4.5a	25.83b
236960	43a	83.5a	75.3a	30b	44.4a	11.4a	4d-g	5.2n-q	12.6f-q	5a	5a	27.30b
236961	51a	90a	82.5a	33.4b	48.3a	12.9a	4.1c-f	5.1o-q	8.7v-w	5a	5a	25.75b
243493	41a	86a	79.3a	29.7b	50a	10.5a	4.3b-d	5.2n-q	10.9o-u	5a	4.5a	25.20b
243502	38.5a	84a	71.9a	26.7b	42.1a	13.2a	4.3b-d	9.8b-e	11.5l-u	5a	4.5a	24.20b
243507	37a	84.5a	74a	25.7b	47a	13a	3.3k-m	8.7b-j	11n-u	4.5a	4.5a	23.20b
243508	40.5a	84a	71.3a	29.2b	41a	10.6a	3.4j-l	9.5b-f	12.6f-q	5a	5a	26.60b
243509	47a	81.5a	87.6a	30.9bb	51.2a	12a	3.9e-h	7.1f-p	10.4r-v	5a	5a	165.75a
243529	47a	81.5a	84.8a	25.8b	55a	16.1a	4d-g	6.3i-q	11.8j-t	5a	5a	25.50b
243535	38a	87a	82.6a	26b	52.2a	14.5a	3.2l-m	7.1i-p	12.7f-q	5a	5a	24.20b
243536	50a	90.5a	86.9a	38.4b	51.7a	12.7a	3.4j-l	8.6b-j	14.1c-h	5a	5a	31.50b
243537	38a	88a	79.5a	31.1b	47.9a	12.1a	3.7g-j	6.7i-q	12.2i-t	5a	4.5a	21.35b
243538	43.5a	89a	78.4a	27.9b	47.5a	10.7a	4d-g	7.8d-m	13.9d-j	5a	4.5a	24.00b
243539	43a	86.5a	82.1a	32.9b	49.8a	15.9a	3.4j-l	8.3b-m	11.4m-u	5a	4.5a	20.15b
243540	41.5a	85a	70.6a	28b	42.1a	14.4a	3.7g-j	6.2j-q	12.7f-q	5a	4a	22.95b
243544	50a	80.5a	82.5a	28.7b	51.3a	13.8a	3.5i-l	7.4e-p	11.8j-t	5a	4.5a	25.75b
243545	42.5a	82a	82.6a	29.9b	50a	10.3a	3.6h-k	7.9c-m	12.5f-r	5a	4.5a	20.70b
243547	50a	87a	85.1a	27.7b	51.8a	11a	3.6h-k	7.5e-p	13.7d-k	5a	4.5a	18.25b
243548	42.5a	84.5a	79.8a	29.6b	49a	12.1a	3.4j-l	8.5b-k	13.1f-n	5a	4a	30.15b
243549	50a	85.5a	86.1a	26.5b	52.7a	13a	3.6h-k	6.6h-q	13.9d-j	5a	3.5a	22.05b
243550	48.5a	84a	70.2a	26.7b	41.1a	13.7a	3.8f-i	6.2j-q	11.9i-t	5a	4.5a	22.85b
243551	40.5a	83a	82.4a	34.1b	50.1a	16.1a	3.4j-l	9b-f	14.2c-h	5a	4.5a	24.20b
243553	45a	84a	84.3a	36.6b	50.4a	12.6a	3.7g-j	8.4b-l	10.6q-v	5a	4.5a	27.55b
8243512	45a	88a	70.5a	31.6b	41.6a	16.5a	4.4b-c	5.8m-q	10.9o-u	5a	5a	30.20b
243517	40.5a	84a	73a	30.7b	42.1a	10a	4.4b-c	8.2b-m	11.8j-t	5a	4.5a	20.30b
243518	46a	86a	79.9a	31.7b	47.2a	13.7a	3.4j-l	8.2b-m	14.4n-u	5a	4.5a	23.95b
243519	47a	87a	82.6a	28.4b	52.9a	15.5a	3.8f-i	6.8h-q	12.7f-q	5a	4.5a	23.95b
243520	40.5a	86a	87.6a	33.2b	52.7a	9.7a	3.6h-k	8b-m	12.9f-p	4.5a	4a	26.35b
243521	46.5a	85.5a	85.1a	32.6b	49.7a	10.9a	3.6h-k	8b-m	13.4f-m	4.5a	4a	23.80b
243524	49a	90a	72.8a	63a	40.3a	11.1a	3.6h-k	6k-q	13.6e-l	4.5a	4a	25.65b
243525	41a	85.5a	81.7a	30.8b	49.2a	11.8a	3.6h-k	6.6h-q	13.5f-m	5a	4.5a	25.45b
243526	40.5a	90a	71.2a	30.5b	50.6a	14.1a	3.4j-l	5.2n-q	15.7b-e	5a	4a	26.75b
243527	50.5a	88a	84.4a	33.8ba	51.2a	10.6a	4d-g	6.6i-q	12.2i-t	4.5a	3a	27.40b
243528	42.5a	85.5a	87.4a	30.3b	52.8a	11.7a	4d-g	6.6i-q	10.6q-v	5a	3a	28.60b
244789	42.5a	87.5a	76.8a	33.2b	43a	15.4a	4.8a	10.2b-d	13.4f-m	5a	4a	25.40b
244790	42.5a	83a	85.8a	32.7b	50.5a	12.5a	4.8a	10.5b	13.1f-m	5a	4.5a	30.50b
244791	43a	78a	75.2a	23.2b	46.9a	16.3a	2.8n	6.4i-q	9.4u-v	5a	4.5a	19.35b
244792	46.5a	84a	69.4a	29.7b	40.1a	14.8a	3.8f-i	6.7h-q	10.7q-v	4.5a	4a	24.50b
244793	50.5a	80.5a	92.1a	35.2b	55.6a	14.4a	3.6h-k	6.7h-q	11.8j-t	5a	3.5a	15.80b
244794	42	84a	76.6a	34.6b	44.6a	12.7a	3.9e-h	7.7d-n	14c-h	5a	3.5a	14.95b
244796	42.5a	80.5a	78.6a	34.4	46.8a	13.3a	3.8f-i	9b-h	10.8p-v	4.5a	3.5a	28.55b
244814	46a	86a	75a	32b	42.3a	12.6a	3.2l-m	6.6i-q	14.5c-g	4.5a	3.5a	3.50b
244815	41a	85a	87.1a	38.1b	50.4a	14.9a	4d-g	10.4b-c	12.7f-q	5a	4.5a	27.90b
244816	46.5a	86.5a	73.1a	28b	42a	9.7a	4.2b-d	10.4bc	13f-o	5a	5a	24.20b
244817	41a	88a	81.1a	36.2b	47.4a	12.9a	4d-g	6k-q	8.7vw	5a	4.5a	21.10b
244818	49a	85.5a	77.1a	34.5b	43a	10.3a	3m-n	6.4i-q	11.4m-u	4.5a	4.5a	23.70b
244821	51a	82a	90a	34.3b	51a	15.5a	3.8f-i	8.2b-m	10.2t-v	5a	5a	24.75b
244822	39a	79.5a	77.6a	29.5b	46.8a	13.8a	3.6h-k	4.4q	10.3s-v	5a	5a	23.80b

244850	47a	82a	69.3a	26.5b	42.2a	12.1a	4.4b-c	6k-q	12.5f-r	5a	5a	25.45b
244851	40.5a	86a	77.2a	31.4b	42.6a	10.8a	3.6h-k	8.8b-h	11.5l-t	4.5a	4a	23.15b
244852	46.5a	91.5a	81.2a	27.7b	49.3a	15.3a	4.5a-b	9.8b-e	14.6c-f	4.5a	4a	29.00b
244853	42.5a	86a	75.6a	34.8b	44.7a	11.8a	3.4j-l	7.8d-m	16.1a-c	4.5a	4a	22.20b
244854	46a	79.5a	72.5a	27.5b	43.5a	14.2a	3.6h-k	8.2b-m	12.4g-s	4.5a	4.5a	24.25b
244856	48a	89a	67.9a	22.7b	42.9a	14.1a	3.6h-k	6.5h-q	6.8w	4.5a	4a	23.25b
244876	50.5a	82a	77.5a	35.7b	4+4.5a	11.9a	2.2o	7.6e-m	11.8j-t	4.5a	5a	21.25b
244877	42.5a	83.5a	77a	31.9b	47.6a	13.1a	4.3b-d	9b-h	13.2f-m	4.5a	4.5a	22.30b
244878	40a	79a	78.5a	29.7b	48.3a	12a	3.4j-l	5.9l-q	13.4f-m	5a	5a	28.20b
244879	50a	87a	81.5a	35.2b	47.6a	12.8a	4d-g	9b-h	13f-o	4.5a	4.5a	29.30b
244880	42.5a	82a	82.1a	34.7b	49.4a	15.1a	4d-g	13.2a	17.1a-b	4.5a	5a	25.85b
244881	46.5a	84.5a	82.6a	33.6b	48.1a	16.5a	3.6h-k	5p-q	15.8b-d	5a	4.5a	24.65b
244882	43a	81.5a	87.2a	34b	51.7a	14.7a	3.2l-m	9.4b-g	14.6c-f	4.5a	3.5a	27.10b

DTH =Days to heading, DTM=Days to maturity, PLH=Plant height, PAL=Panicle length, CL =Culm length, PDL=Peduncle length, TN=Total node, FCIN=First culm internode, SCIN =Second culm internode, TT=Total tiller, FT=Fertile tiller, HSW=Hundred seed weight

Number of total tiller ranged from 4.5 for accession 243524 to 5 for accession 15309 with the mean value of 4.5. The length of first culm internode ranged from 4.4 to 13.2. The longest first culm internode was recorded by accession 15781 (14 cm) and accession 244879 (13 cm). However, the present finding is different when compared to the report of Zeid *et al*. (2012). Peduncle length ranged from 6 cm to 30cm; this trait exhibited the highest variation between the smallest and the biggest value. Panicle length varied from a minimum of 22.5 cm to a maximum of 46cm which showed wide variation with a mean of 34 cm. There was wide variation observed between the minimum and maximum value for length of first basal internode and second basal internode. Hundred seed weight also showed variability among the accessions tested. It ranged from 14 mg to 43 mg with a mean value of 28.5 mg per hundred seed. The reason for these variations was from the results, the variability observed among accessions of tef for all characters may indicate the possibilities for genetic improvement of the crop through continuous selection and cross breeding. The mean comparisons between accessions showed high variability. It may indicate the availability of desirable traits in every direction; therefore, any improvement and conservation programs should give attention for all geographic locations and consider every location as a source of important traits. According to the study, accessions from South Wollo, North Gonder, Tigray Central zone, East wollega and Illubabor showed lowest length which is important trait related to lodging resistant. The top 8 highest panicle length accessions were from North Shewa, East Wollega, Kembata Tembaro and Tigray East zone accessions. This implies the potential of the regions for further yield improvement program. The present study revealed the existence of considerable genetic variability among the accessions in the above regions. Therefore, for further conservation, breeding and any other improvement program of tef in Ethiopia due attention and exploiting of the existing potential of the above regions is essential.

4.1.3 Cluster Analysis

Cluster analysis was performed on the Euclidean distance matrix utilizing Ward's linkage method, and the resulting dendrogram is given in Figure 4, using MINITAB software version 16. The cluster analysis has grouped the 64 tef accessions into eight distinct clusters based on their quantitative traits

similarity (Figure 3). Members within a single cluster being considered as having more close relationships with each other than they are with those in significantly distant clusters. The result showed that, in most cases the accessions resulting from different area were clustered together, this may be attributed to an exchange of genetic materials (gene flow) between the parents. The different accessions grouped within a given cluster were assumed to be more closely related in terms of the studied traits than those accessions grouped into different clusters. A wide range of genetic diversity among all accessions was observed. The results have shown that it is possible to both classify the genetic diversity of elite accessions and select accessions or cultivars for the highest genetic diversity, as indicated by cluster analysis.

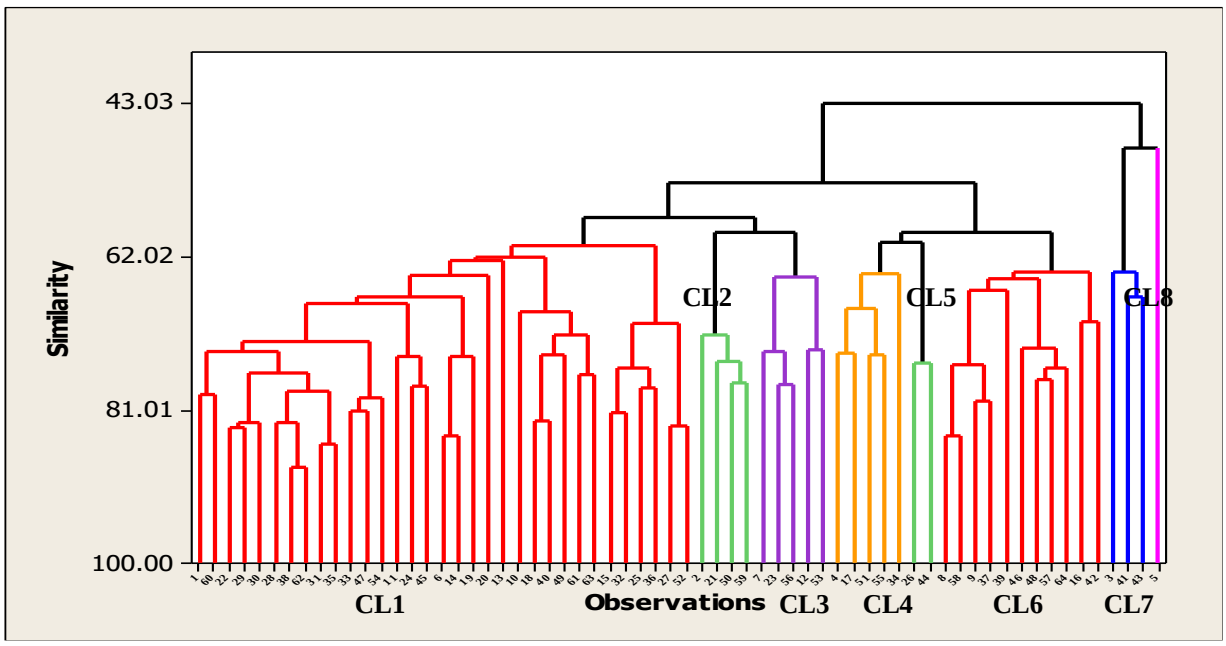


Figure 3. Dendrogram showing relationship among 64 tef accessions based on average linkage and Euclidean distance using the mean of 12 quantitative traits.

The distribution of accessions into different divers1ity classes was presented in Appendix 6. Each cluster comprises different number of accessions and each cluster has a unique characteristic. The current analysis grouped all the accessions in to eight clusters which indicate that the variation present in tef accessions. Similar result with the current finding was reported by Adnew *et al*. (2005). According to their report, 144 accessions of tef were grouped in to eight major clusters. Even though, there are accessions from similar geographic origins grouped together; clustering of germplasm was not associated with the geographical distribution and mainly grouped due to their morphological differences. Because of the material transfer and existence of gene flow between the regions; accessions in the present study do not group exclusively on the basis of their origin. The independent events of evolutionary forces such as genetic drift, mutation, migration, selection and germplasm exchange might play roles in assigning the accessions to different clusters. These factors might be

separated them into related but different gene pools. Geographically distant accessions are grouped together; which indicates that accessions from different geographic origin might have similar genetic background. The random and independent clustering pattern and discordance among diversity patterns and geographical distribution of accessions in this investigation showed that geographical isolation is not the only factor causing genetic diversity in tef. Therefore, for tef hybridization parental lines should be selected based on genetic diversity rather than the geographical distribution.

4.1.4 Clusters Mean and Comparison of Clusters

The first cluster constitutes 33 accessions which were collected from the whole regions of the study area. Accessions from north Wollo and central zone constituted more than 42% of the group. The whole accessions from Tigray central zone except one were grouped in cluster I Appendix 6. The second largest cluster (CL 6) has 11 accessions; contains accessions from four different populations North Wollo, Kembata Tembaro, East wollega and Illubabor. The commercial nature of the crop may favor the possible exchange of genetic material by human. The fifth cluster is containing only single accessions. Similar trend was observed in the previous tef clustering studies. Adnew *et al.* (2005) reported that two solitary clusters were observed from the total eight clusters. The solitary clusters were different from the rest of the clusters by superior values of hundred seed weight. Evaluated 144 tef accessions grouped into eight clusters of which two of them are solitary clusters (Adnew *et al.,* 2005). Based on cluster mean values; the first cluster (CL1) was explained by longer days to heading and days to maturity. This cluster was also characterized by lowest mean value for the first culm internode. The second cluster (CL2) was differentiated by longest peduncle and culm length, highest panicle length and shortest culm length. Good performance was observed for the peduncle length and highest tiller number. The third cluster (CL3) which form separate cluster was characterized by shorter plant height and lower hundred seed weight. The fourth cluster (CL4) was described by longest peduncle and highest mean value for hundred seed weight was recorded. The cluster was among the top high yielder genotype and top in hundred seed weight. The average performance for number of seed per plant was also higher. The shortest peduncle was described by (CL5). Compared to the other clusters, cluster (CL6) generally showed highest mean performance in most of the desirable traits. The 7th cluster exhibited early maturity and highest tiller number per plant, number of total node and length of second culm internode. The cluster has shorter days to maturity, longer plant height and higher hundred seed weight. The longest culm length was recorded by cluster eight (CL8). It is important to note that the superiority of a particular tef accessions in respect to a given character gets weak by other accessions that are related and grouped in the same clusters which are inferior or intermediary for that character in question.

4.1.5 Principal Component Analysis

In order to reduce a large set of variables to a small set and to assess the pattern of variations; a well-known method of dimension reduction, principal component analysis was done by considering all the 12 variables simultaneously. Five of the 12 principal components accounted for more than 62 % of the total variation (Table 5). The first principal component alone accounted for about 16.6 % of the total variation. While four of the 12 traits considered exerted positive effects on this component, the rest eight while four of the 12 traits considered exerted positive effects on this component, the rest eight (DM, DH, PH, PAL, PDL, TT, FT and FCIN) traits employed negative effects of different magnitudes. Among those traits having positive contribution effect on the first PC include number of second culm internode, total tiller per plant, fertile tiller per plant and hundred seed weight. On the other hand, days to heading, days to maturity, panicle length, peduncle length, culm length, total node, first culm internode and plant height had all negative weights on this component. Interpretation of the principal components is based on the studied variables which are most strongly correlated with each component (i.e., which of these Eigenvectors are large in magnitude, the farthest from zero in either positive or negative direction). As Holland (2008) suggested, standard criteria permit to ignore components whose Eigen values are less than 1 when a correlation matrix is used. In the present study, the principal components analysis revealed that five principal components with Eigen-values greater than unity accounted for 62.62% of the gross variability in 12 phenomorphic characters (Table 5). Similarly, Kebebew *et al*. (1999, 2000) reported that about 71-79% of the variation in 320 tef germplasm lines was explained by five PCs. The first principal component alone explained 16.6% of the total variation, while PC2, PC3, PC4 and PC5 in that order accounted for 15.54%, 11.4%, 10.46%, and 8.61% of the gross observed variation among the tef accessions. The first three PCs together accounted for a cumulative of 43.55% of the total variation indicating that much of the variability among the test accessions originated from the traits included in these PCs.

Table 5. Eigen value, percentage, cumulative variance and Eigen vector of the first five principal components for 12 traits in 64 tef accessions

Parameters	**Principal components (PCs)**				
	PC1	PC2	PC3	PC4	PC5
Eigen value	2.32	2.18	1.59	1.46	1.20
Variance (%)	16.6	15.54	11.4	10.46	8.61
Cumulative variance	16.6	32.15	43.55	54.01	62.62
Traits	Eigen Vector				
Days to heading	-0.04	0.64	-0.01	-0.17	-0.01
Days to maturity	-0.31	0.18	-0.01	0.40	-0.33
Plant height	-0.27	-0.07	-0.17	0.58	0.08

Panicle length	-0.03	-0.04	0.60	0.07	0.14
Culm length	-0.16	0.03	0.36	0.34	-0.21
Peduncle length	-0.16	0.09	0.52	0.04	-0.03
Total number of node	-0.03	0.11	0.26	-0.26	0.23
First culm internode	-0.15	0.09	-0.05	0.25	0.67
Second culm internode	0.21	0.25	-0.11	0.20	0.42
Total tiller	0.07	0.27	-0.29	0.17	-0.22
Fertile tiller	0.01	0.01	0.01	0.01	0.00
Hundred seed weight	0.18	-0.03	0.02	0.15	0.26

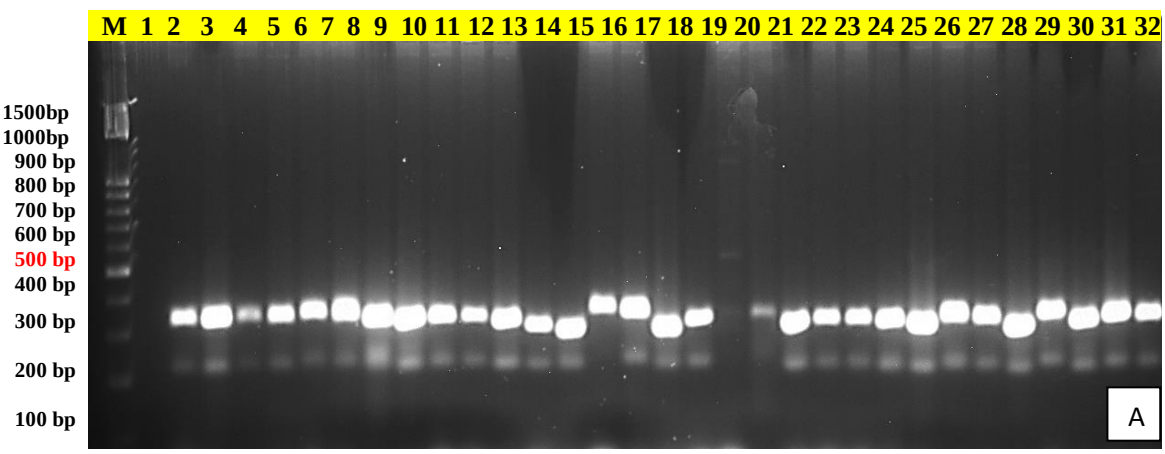
PC1, PC2, PC3, PC4 and PC5 = Principal component 1, 2, 3, 4 and 5 respectively

The second component accounting for an addict dsional 15.54% of the total variation primarily illustrates the patterns of variations in days to heading, second culm internode, total tiller and total number of node; which were found to have positive and high impacts on the second component while panicle length, plant height and hundred seed weight had negative coefficients. The third principal component accounted 11.4 % of the total variation and it was predominantly accounted by variation in panicle length and peduncle length, culm length and Total number of node. Whereas, days to heading, days to maturity, plant height, total tiller, second culm internode and first culm internode contributed negative value. The fourth principal component accounted 8.61 % of the total variation and was indicated with high variation in plant height and days to maturity.

4.2 Genetic Diversity Analysis Using SSR Markers

4.2.1 Microsatellite Markers level of polymorphism

Polymorphism information content (PIC) value is commonly used in genetics as a measure of polymorphism for a marker locus used in linkage analysis. In this thesis work, only polymorphic SSR markers were selected for diversity analysis (Figure 4 A and B). PCR amplification of the SSR markers was undertaken using the ten primer pairs that were polymorphic among the tef accessions and a total of 914 bands were scored from the examined 10 SSR markers.



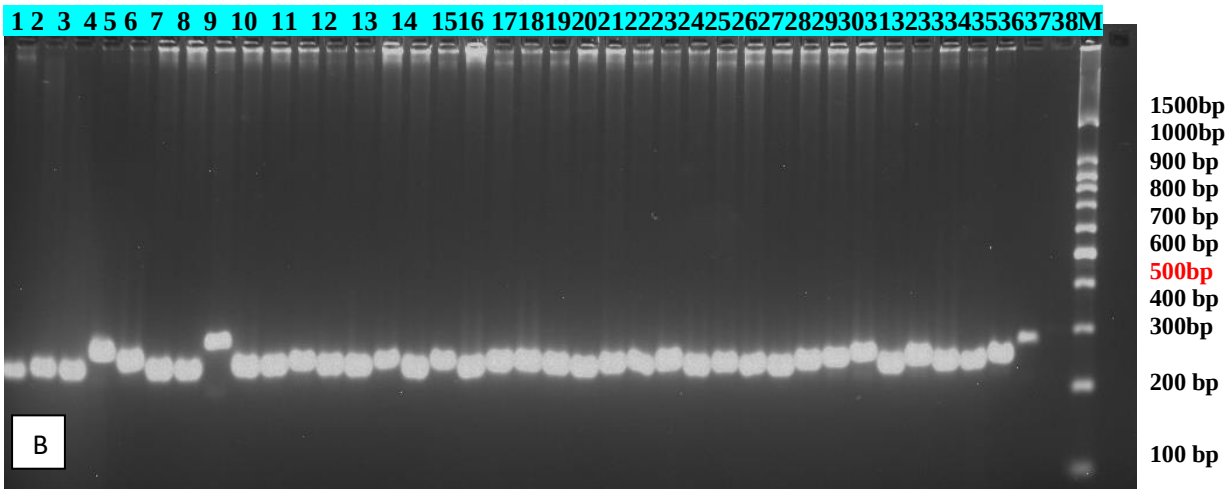


Figure 4. Sample of gel electrophoretic pattern of SSR markers (A, CNLTS 461 and B, CNLTS5) generated from tef accessions collected from different geographical areas in Ethiopia. Where; M is a 100bp DNA ladder and the numbers in lanes indicates the accessions.

4.2.2 Analysis of Molecular Variance

The spread/exchange of single genes, genotypes and even the whole population in different regions resulted by the processes of exchange of gametes, individuals and population on geographic scale (which is considered as gene flow) in conjugation with other evolutionary forces is sufficient to prevent differentiation between populations (McDermott and McDonald, 1993). To explain the properties of subdivided populations; the magnitude of differentiation between and within populations can be quantified using F-statistics (Fit, Fis and Fst) also known as fixation indexes (Wright, 1951). Fixation index (Fst) is a measure of population differentiation due to genetic structure. According to Wright (1951), the threshold to determine the level of Fst value ranges from 0 to 0.05 considered as low, 0.05 to 0.15 moderate, 0.15 to 0.25 large and those greater than 0.25 mean very large genetic differentiations among populations. Analysis of molecular variance (AMOVA) partitioned the total molecular variance within and among the accessions based on their source of origins. There were highly significant ($P<0.001$) molecular variances among populations (AP), among individuals (AI) and within individuals (WI). The highest proportion 56 % of the variation was attributed to genetic variability within individuals (WI), while 39 % was due to variation among individuals within the same population. In contrast, a smaller portion 5 % of the total variation was among populations (AP) (Table 6). This low genetic differentiation among population may be due to the communality of seed source and use. This

leads to an increase in the distribution of alleles among different populations regardless of their geographical distance.

Table 6. Analysis of molecular variance (AMOVA) based on standard permutation across the full data set of tef accessions collected from different geographical origin

Source of Variation	Df	SS	MS	Estimated Genetic Variance	%	*P Value*	F-statistics
Among Populations	8	68.36	8.539	0.210	5	0.001	Fst=0.05
Among individuals	55	308.380	5.607	1.632	39	0.001	Fis=0.41
Within population	64	150.000	2.344	2.344	56	0.001	Fit =0.44
Total	127	526.695		4.185	100		Nm =4.742

Df = degree of freedom, SS=sum of squares and MS=mean squares

Supportive results with the present study were reported by different authors. Abraha *et al.* (2016) based on the molecular characterization of sixty tef accession using 10 SSR markers. According to the result the variance component within population accounted 63 % while among population accounted only 2%. The observed Fst value among the studied tef populations was 0.05 (Table 6). According to the rule the extent of genetic differentiation among the nine populations in terms of allele frequencies measured was small. This implied inter mating among individuals within similar populations was not significant because tef is self-pollinating. The overall observed gene flow (Nm) or gene migration value was (4.742), which showed the approximate number of individual's migration from one population to the other. Based on Slatkin (1985) and Waples (1987), Nm values grouped into three categories: Nm > 1.00 high, 0.25-0.99 intermediate and 0.000 – 0.249 low. Therefore, the high Nm value observed in this study indicates high gene flow between populations which will agree with the AMOVA result showing low variation between populations. The pair-wise genetic differentiation among accessions within the source of origin ranged from 0.19 between East Wollega and Kembata Tembaro to 0.99 and 0.90 between Illubabor and Tigray West zone and between Kembata Tembaro and Tigray West zone populations respectively (Table 7). The low value of Fst implies that there is high frequency of identical alleles among accessions. This might be due to the commercial nature of the crop from both directions of East wollega to Kembata Tembaro or vice versa resulting in admixture among the two populations because of seed exchange. The lowest Fst value observed between the two populations can be explained by high level of seed exchange, which leads to genetic lose. Illubabor populations have larger Fst value between Illubabor and Tigray West zone. This might be because of low seed exchange with Illubabor and low distribution of the accessions in Tigray West zone.

Table 7. Pair wise population matrix Fst Values for assembled nine tef populations

	NS	SW	NG	WG	CZ	WZ	KT	EW	IL
NS	0								
SW	0.51	0							
NG	0.46	0.61	0						
WG	0.73	0.48	0.59	0					
CZ	0.48	0.33	0.46	0.43	0				
WZ	0.60	0.97	0.57	0.72	0.64	0			
KT	0.88	0.46	0.62	0.58	0.55	0.90	0		
EW	0.75	0.36	0.40	0.51	0.48	0.74	0.19	0	
IL	0.60	0.52	0.51	0.89	0.82	0.99	0.71	0.46	0

NS=North Shewa; SW= South Wollo; NG= North Gondar; WG=West Gojjam; TCZ=Tigray Central Zone; TWZ= Tigray West Zone; KT=Kembata Tembaro; EW=East wollega; IL= Illubabor

4.2.3 Analysis of Diversity Parameters

The study identified a total of 914 alleles, giving an average of 14.5 alleles per locus for the evaluated 10 microsatellites. The number of alleles detected in the present study was lower than the number of microsatellite alleles reported by Abraha *et al*. (2016); according to the report number of alleles ranged from eight to twenty three per locus. This might be the number and type of genotypes used for the study was different from the previous studies and the genotypic difference of the tested accessions may favor the difference. The highest number (109) of bands per locus was recorded for CNLTs463 marker, out of which 60 (55.05%) were polymorphic. CNLTs19 resulted in the highest percentage of polymorphic bands (76.19%), whereas CNLTs5 gave the smallest percentage (39.77%) of polymorphic bands (Table 8). The percentage of polymorphic bands produced by the other microsatellite markers were 42.17% (CNLTs24), 72.83% (CNLTs2), 61.95% (CNLTs6), 60.95% (CNLTs22), 64.63% (CNLTs25), and 67.01% (CNLTs458) and 67.74% (CNLTs461) (Table 8). The study resulted in a total of 914 alleles (average 14.50 alleles per locus), out of which 84 (62.77%) were scarce (frequency between 0.01 and 0.05) (Table 6). The frequency of 16 (11.94%) alleles was between 0.05 and 0.1, whereas 34 alleles (25.37%) had a frequency of 0.1 or higher (Table 8).

Table 8. Summary of the number of alleles with their respective frequencies

Markers	Number of alleles with frequency			
	Scarce (0.01 - 0.05)	0.05 - 0.1	0.1 or higher	Total
CNLTs2	8	1	4	13
CNLTs5	9	2	4	15
CNLTs6	9	4	3	16
CNLTs19	9	0	4	13
CNLTs22	10	1	3	14
CNLTs24	7	2	3	12
CNLTs25	12	2	3	17
CNLTs458	10	1	4	15
CNLTs461	10	2	3	15
CNLTs463	8	1	3	12
Total	84	16	34	134
Percentage	62.77	11.94	25.37	

4.2. 4 Genetic distance between populations

The analysis showed that the major allele frequency within population ranged from 0.23 (CNLTs458) to 0.07 (CNLTs5) with mean frequency of 0.33 per locus (Table 9). The highest gene diversity (0.91), allelic richness (6.33), private allelic richness (0.22), polymorphic information content (0.90), number of alleles (13), effective number of alleles (5.09) and Shannon's Information Index (1.63) were recorded for the microsatellite marker CNLTs6. On the other hand, the highest major allele frequency (0.66) and the lowest gene diversity (0.63), allelic richness (4.11), polymorphic information content (0.61), number of alleles (14.5), effective number of alleles (3.58), and Shannon's information index (1.38) were revealed by CNLTs5 (Table 10). The study also revealed that all the markers were highly informative with the PIC ranging from 0.61 (CNLTs5) to 0.90 (CNLTs6). The highest gene flow (2.80), the lowest fixation index (0.02) were observed for CNLTs463 and the mean observed heterozygosity (Ho) of 0.76 was greater than the average expected heterozygosity (He) of 0.74 CNLTs463 (Table 9). Highest and lowest numbers of effective alleles were observed for marker CNLTs6 (5.09) and CNLTs2 (2.89). The mean value of effective alleles was lower than earlier report Zeid *et al.* (2012).

Marker CNLTs463 exhibited the peak gene flow (2.80) but the lowest value was 0.75 for marker CNLTs2. The lowest value of Shannon information index (I) was recorded for marker CNLTs2; whereas, CNLTs6 exhibited the highest value. The polymorphic information content (PIC) value provide an estimate of discrimination power of markers by taking number of alleles and relative frequencies of alleles. Markers polymorphism with >0.5 are considered as highly informative, between

0.5 and 0.25 moderate and those with values below 0.25 are low informative (Anderson, *et al.*, 1993; Smith, *et al.*, 1997). In the present study the PIC value ranges from 0.61 for marker CNLTs5 to 0.90 for marker CNLTs6 with a mean value of 0.87. This showed that the markers used in the study were highly informative in expressing the genetic heterogeneity of the tasted tef accessions (Table 9).

Table 9. Informativeness and levels of different diversity indices of the SSR loci across populations of tef collected from different agro ecological origin

Marker	MAF	Na	Ne	I	He	Ho	H	GD	Nm	PIC
CNLTs2	0.48	4.44	2.89	1.15	0.59	0.36	0.36	0.78	0.78	0.76
CNLTs5	0.66	5.11	3.58	1.38	0.70	0.35	0.36	0.63	1.11	0.61
CNLTs6	0.14	6.22	5.09	1.63	0.77	0.51	0.53	0.91	1.50	0.90
CNLTs19	0.46	5.11	3.38	1.32	0.66	0.32	0.33	0.89	1.43	0.86
CNLTs22	0.34	5.33	3.83	1.42	0.71	0.66	0.64	0.89	1.92	0.86
CNLTs24	0.35	4.88	3.25	1.29	0.67	0.33	0.33	0.80	1.40	0.78
CNLTs25	0.38	5.67	4.16	1.47	0.71	0.42	0.42	0.85	1.57	0.84
CNLTs458	0.23	5.44	3.73	1.44	0.71	0.51	0.50	0.86	1.32	0.85
CNLTs461	0.34	6.11	4.49	1.58	0.75	0.47	0.45	0.85	2.12	0.84
CNLTs463	0.37	5.56	4.27	1.51	0.74	0.76	0.77	0.82	2.80	0.88
Mean	0.33	5.00	3.32	1.30	0.66	0.44	0.47	0.81	1.65	0.87

Where MAF = Major allele frequency, NA= Number of different alleles, Ne = Effective number of alleles, GD= Gene diversity, PIC = Polymorphic information content, I = Shannon's Information Index Nm= Gene flow

The result indicated that the efficacy of the markers for further genetic diversity study of tef accessions in the future. In the present study the observed higher PIC value is the same as to Zeid *et al.* (2012).

4.2.5 Genetic Variability Within and Among the Populations

Summary of the different genetic diversity indices over the ten markers for the nine populations are presented in Table 10. The comparative analysis showed that there is no much difference among the nine populations of tef with regard to genetic diversity indices including similar number of alleles, effective number of alleles, private allelic richness, Shannon's information index, observed heterozygosity, expected heterozygosity, percentage of polymorphic loci and fixation index. Comparatively, the populations of East Wollega scored greater values in number of different alleles (7.7) and effective number of alleles (4.91), private allelic richness from North Gondar (0.8), and fixation index from West Gojjam (0.42).Whereas, expected heterozygosity in North Gondar population is higher. However, both North Gondar and Illubabor populations showed same value in observed heterozygosity (0.56) and percentage of polymorphic loci (100) (Table 10).

Table 10. Summary of different population diversity indices averaged over the 9 loci for each population

No.	Populations	N	Na	Ne	Arp	I	Ho	He	PPL	F
1	**N. Shewa**	**5**	**5.10**	**4.14**	**0.5**	**1.47**	**0.52**	**0.73**	**100**	**0.30**
2	**S. Wollo**	**5**	**3.80**	**2.87**	**0.1**	**1.15**	**0.37**	**0.63**	**100**	**0.41**
3	**N. Gondar**	**9**	**6.70**	**4.54**	**0.8**	**1.64**	**0.56**	**0.76**	**100**	**0.27**
4	**W. Gojjam**	**6**	**5.50**	**4.01**	**1.0**	**1.49**	**0.43**	**0.72**	**100**	**0.42**
5	**C. Zone**	**7**	**3.80**	**2.55**	**0.5**	**1.06**	**0.34**	**0.58**	**100**	**0.37**
6	**W. Zone**	**4**	**3.70**	**2.99**	**0.1**	**1.15**	**0.48**	**0.64**	**100**	**0.27**
7	**K. Tembaro**	**7**	**6.00**	**4.58**	**0.5**	**1.56**	**0.46**	**0.73**	**100**	**0.33**
8	**E. Wollega**	**13**	**7.70**	**4.91**	**0.6**	**1.70**	**0.48**	**0.75**	**100**	**0.36**
9	**Illubabor**	**8**	**6.10**	**4.21**	**0.4**	**1.56**	**0.56**	**0.74**	**100**	**0.25**
	`Mean	**7.11**	**5.38**	**3.87**	**0.5**	**1.42**	**0.47**	**0.70**	**100**	**0.33**

N= Number of samples, Na= Number of different alleles, Ne = Effective number of alleles, Arp = Private allelic richness, I = Shannon's Information Index, Ho = Observed heterozygosity, He = Expected heterozygosity, PPL= Percentage of polymorphic loci, F = Fixation Index

4.2.6 Cluster Analysis and Population Structure

These analyses resulted in three clusters (CL1, CL2 and CL3). All the further grouped into two sub-clusters (Figure 5). The first two major clusters (CL1 and CL2) are composed of individuals from the 9 and 7 populations respectively. The third cluster is composed of only 6 populations. Although there could be considerable intermixing, sub-clustering resulted in the individuals to be grouped according to their geographical regions of collection Sub-clustering of the major cluster CL1 gave two groups with 5 and 7 populations. In the same way, sub-clustering of CL2 resulted in two group composed of six and five populations (Figure 5). Likewise, accessions from different populations clustered together, which may imply the existence of gene flow between and within populations. Illubabor accessions were grouped with geographically distant accessions, Kembata Tembaro in Cluster III. This indicates accessions in one cluster might evolved from different lines of ancestry. In addition the independent events of evolutionary forces such as genetic drift, mutation, migration, selection and germplasm exchange might separated them into related but different gene pools (Slatkin and Takahata, 1985 and Keneni *et al.*, 2012). These also indicated that varieties from different seed types might have similar genetic background for microsatellite markers.

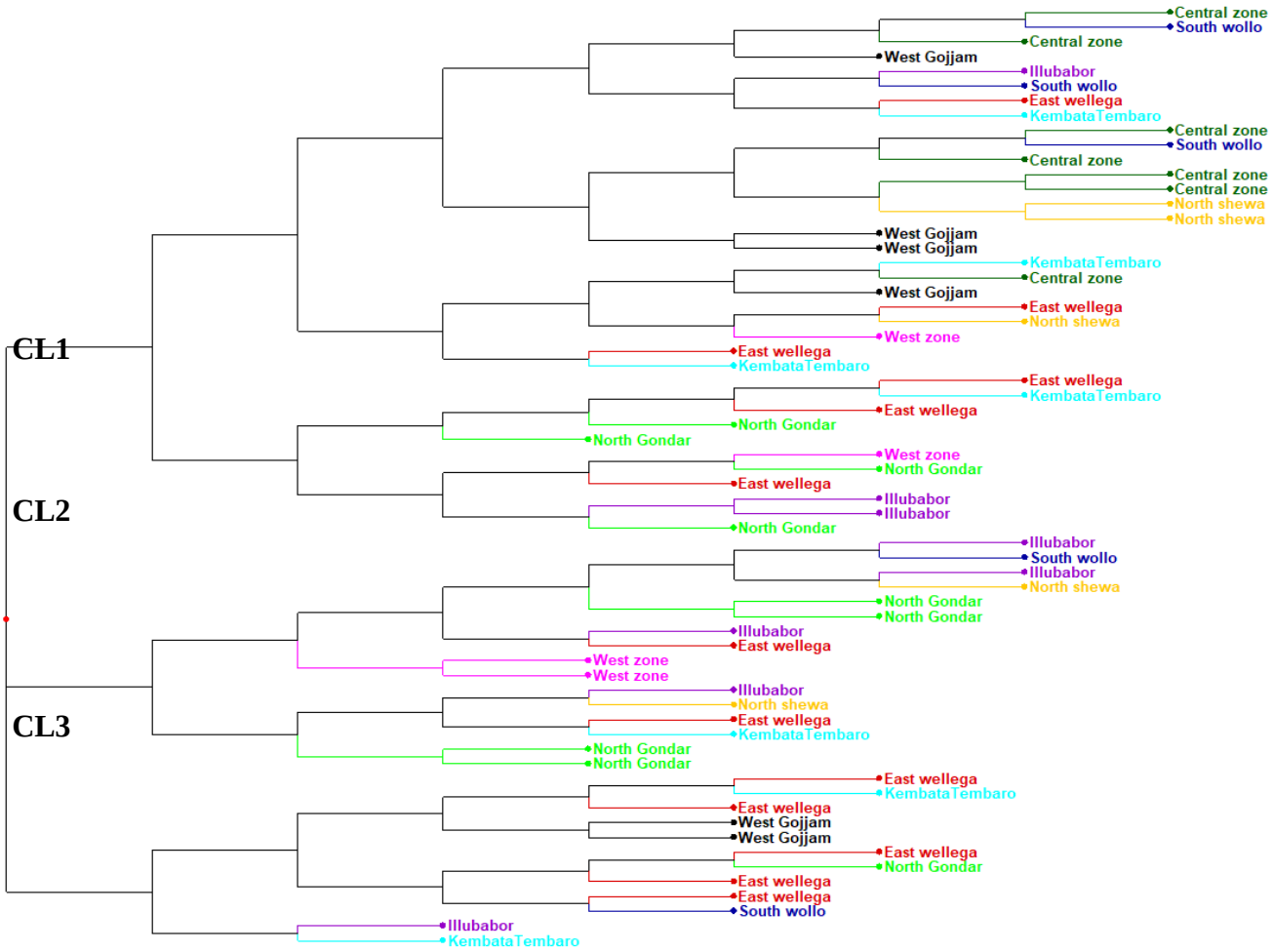


Figure 5. Unweighted Neighbor Joining (NJ) dendrogram showing genetic relationship of 64 tef accessions using 10 SSR markers

4.2.7 Principal Coordinate Analysis

Principal Co-ordinate Analysis (PCoA) showed that the first three coordinates accounted for about 30.72% of the genetic variation present in SSR molecular data derived from the study. The first, second, and third principal coordinates respectively explained about 13.13 %, 9.20 % and 8.39 % of the gross variation respectively. The PCoA analysis in the two-dimensional plot displayed in Figure 6 showed that accessions from different collection sites often clustered together. There was no separate group formed by a single population. This, in turn, agrees with the results of the dendrogram in that there was no unique clustering among accessions from the same population. The overall grouping pattern of PCoA corresponds with the clustering dendrogram (Figure 6 and 5) which explains about conformity of results obtained from the cluster analysis. The result in the PCoA further supports previous results of Zeid *et al.* (2012) and Solomon *et al.* (2009) where mixed clustering was observed among accessions from different origin of populations. In some cases, accessions of the same population such as East Wollega and North Wollo formed sub cluster in the major groups. The presence of seed exchange and gene flow between and within populations or collection sites may be the probable explanation behind the mixed clustering of accessions from different populations. The first three principal ordinate axes explained 30.72% of the total variation (Figure 6). The long standing issue

in population genetics is the identification of genetically homogeneous groups of individuals. In this regard, a recent Bayesian algorithm implemented in the software STRUCTURE allows the identification of such groups. The model based approach described by Pritchard *et al.* (2000) was applied in the population structure analysis. The *ad hoc* statistics (ΔK) was calculated according to Evanno *et al.* (2005) to determine true K value. The algorithm allows estimating the true number of clusters (K) in a sample of individuals under study. The analysis revealed that maximum ΔK (which is good indicator of the true number of clusters) becomes peak at K = 3 (Figure 8A). Based on this value, Clumpak result (bar plot) showed wide admixtures and hence there was no clear geographic origin-based structuring of populations (Figure.7B).

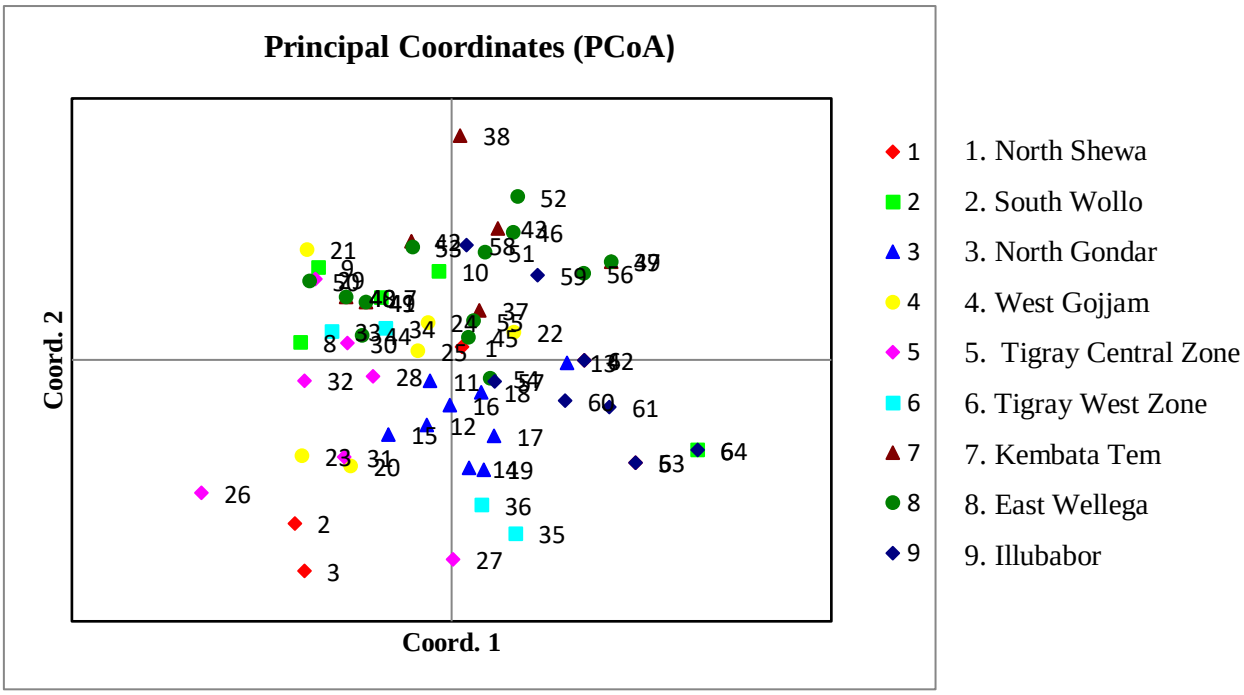
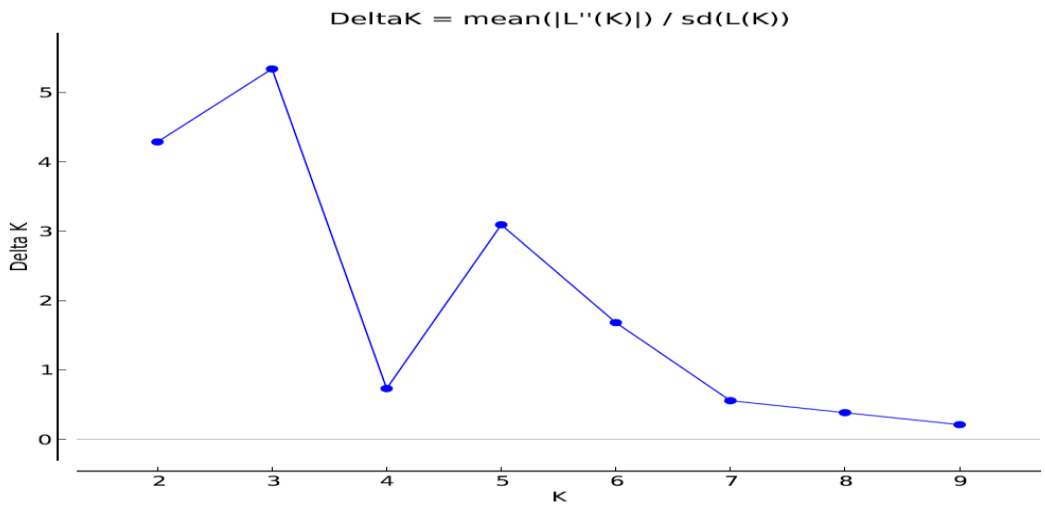


Figure 6. Principal coordinates analysis (PCoA) bi-plot showing the clustering pattern of 64 tef from the nine populations. Samples coded with the same symbol and colors belong to the same population

A



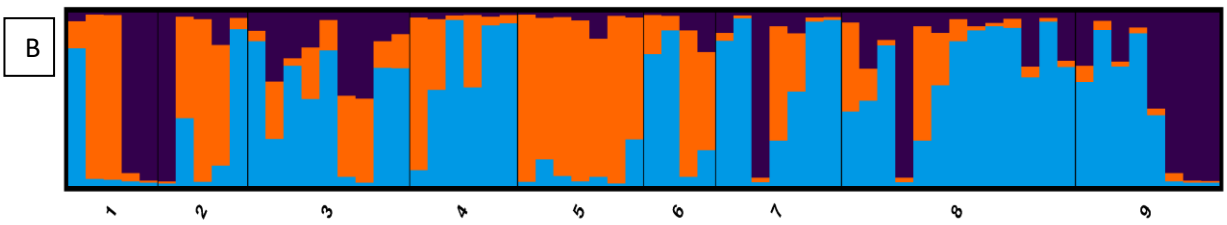


Figure 7. Inferred population structure of 64 tef accessions A) K value estimated using Evanno *et al.* (2005) method (highest peaks at K=3), and B) Bayesian model- based estimation of population structure (K=3) where each color represents a different cluster and the length of the colored segment shows the accessions's estimated proportion of membership in that cluster as calculated by STRUCTURE. The numbers 1-9 on figure 7B indicates the population.

5. CONCLUSIONS AND RECOMMENDATIONS

Genetic diversity of crop plants is the foundation for the sustainable development of new varieties. So there is a need to characterize the diverse genetic resources using different statistical tools and utilize them in the breeding programme. Morphological data in combination with molecular data are used for precise characterization of germplasm resources. With the advent of high throughput molecular marker technologies it is possible to characterize larger number of accessions with limited time and resources. It can be studied using morphological as well as molecular markers such as SSR markers. This study was conducted to evaluate the extent and pattern of genetic diversity of sixty four tef accessions and provide information on the status of genetic variation for further improvement and conservation using 12 morphological parameters and 10 SSR markers. The broad spectrum of genetic diversity in tef implies great opportunities for genetic improvement through either direct selection or intra-specific hybridization between parental lines with desirable traits. Cluster analysis based on the morphological data assigned the accessions into eight clusters based on the characters studied. The analysis of variances (ANOVA) table revealed that there is diversity among the accessions of tef for all characters studied; in other words, there were highly significant (PH, CL and PDL) and significant (for DH, TN, FCIN, and SCIN.) differences ($p<0.01$ and $p<0.05$) respectively.

The five principal components (PCs), with Eigen values greater than 1, explained about 62.62% of the total variation among accessions for all traits. The study found the existence of high levels of diversity among 64 tef accessions which are good for the introduction of new genes. The current study found the existence of high levels of diversity among 64 tef accessions which are good for the introduction of new genes in the existing genotypes. The dendrogram constructed to identify the genetic similarities among these accessions showed that accessions from the same regions were found to cluster mostly together implying a correlation between molecular groupings and their source of collection. The present finding also indicated the effectiveness of SSRs in detecting polymorphisms among different accessions of tef. Clustering pattern on the basis of SSR markers provides ample information in identification and confirmation of tef accessions. This information is significantly crucial for the development of pure tef breeding lines. The study conducted on tef were limited to either on morphological marker or molecular markers and only few accessions are evaluated at a time which is difficult for better estimation of the genetic diversity. The result of this study suggested that future germplasm collection and utilization strategies should take into consideration the magnitude and pattern of genetic diversity established both at genotypic and phenotypic levels.

6. REFERENCES

Abraha, M., Hussein, S., Laing, M. and Assefa, K. (2016). Performance of tef [*Eragrostis tef* (Zucc.)Trotter] genotypes for yield and yield components under drought stressed and non-stressed conditions. *Crop Sci*. **56**:1-8.

Abraham, R. (2015). Achieving food security in Ethiopia by promoting productivity of future world food tef: A review. *Adv Plants Agric Res*. 2.

Adnew, T., Seyfu, K., Hailu, T. and H. Sridhara, (2005). Genetic diversity in tef [*Eragrostis tef* (Zucc.)Trotter] germplasm. *Genet Reso. and Crop Evol*.**52**: 891–902.

Akkaya, M. S., Bhagwat, A.A. and Cregan, P.B. (1995). Length polymorphisms of simple-sequence repeat DNA in soybean. *Gene.* **132:** 1131-1139.

Alaunyte, I., Stojceska, V., Plunkett, A., Ainsworth, P. and Derbyshire, E. (2012). Improving the quality of nutrient-rich Tef (*Eragrostis tef*) breads by combination of enzymes in straight dough and sour dough bread making. *J Cereal Sci.* **55**: 22–30.

Ali, M.A., Gyulai, G., Hidvégi, N., Kerti, B., Al-Hemaid, F.M.A., Pandey, A.K. and Lee, J., (2015). The analysis and genetic advance in phenolmorphic and agronomic traits of tef [*Eragrostis tef)*. *J Pl Physi. and Biochem*.**2**: 1–2.

Ana, B. P. and Cristian, I. V. (2012). PyElph a software tool for gel images analysis and phylogenetics. *BMC Bioinfo.***13**: 9.

Anderson, J. A., Churchil, G. A., Autrique, J. E., Tanksley, S. O. and Sorrels, M. E. (1993). Optimizing parent selection for genetic linkage maps.*Genome*.**36:** 181-186.

Anthony, F., Combes, M.C., Astorga, C., Bertrand, B., Graziosi, G. and Lashermes, P. (2002).The origin of cultivated Coffee Arabica L. varieties revealed by AFLP and SSR markers. *Theor Appl Genet*. **104**: 894–900.

Aremu, C.O. (2011). Genetic diversity: A review for need and measurements for inter species crop improvement. *J of Microbial Biotech Res.* **1**: 80-85.

Assefa, K., Cannarozzi, G., Girma, D., Kamies, R., Chanyalew, S., Plaza-Wüthric, S., Blosch, R., Rindisbacher, A., Rafudeen, S. and Tadele, Z. (2015). Genetic diversity in tef [*Eragrostis tef* (Zucc.)Trotter]. *Front P Sci*. **6**: 1-13.

Assefa, K., Ketema, S., Tefera, H., Kefyalew, T. and Hundera, F. (2000). Trait diversity, heritability and genetic advance in selected germplasm lines of tef *[Erugrostis ref* (Zucc.)Trotter]. *Hered itus,* **133***:* 29-37.

Assefa, K., Merker, A., and Tefera, H. (2003a). Inter simple sequence repeat (ISSR) analysis of genetic diversity in tef [*Eragrostis tef* (Zucc.)Trotter]. *Hereditus,* **139**: 174–183.

Assefa, K., Y, J.K., Zeid, M., Belay, G., Tefera, H., Sorrells, M.E. (2011). Breeding tef [*Eragrostis tef* (Zucc.) Trotter]: conventional and molecular approaches. *P. Breed*.**130**: 1-9.

Ayalneh, T., Amsalu, A., Habtamu, Z., Dessalegn, T. and Wondale, L. (2012). Multivariate diversity, heritability and genetic advance in tef landraces in Ethiopia. *Afr. Crop Sci J*.**19**: 201-212.

Ayele, M., Dolezel, J., Vanduren, M., Brunner, H. and Zapata arias, F. J. (1996). Flow cytometric analysis of nuclear genome of the Ethiopian cereal Tef [*Eragrostis tef* (Zucc) Trotter]. *Genetica*, **8**: 211–215.

Beckman, J. S., and Weber, J.L. (1992). Survey of human and rat microsatellites.*Geno.***12**: 627-631.

Bekele, E. and Lester, R. N. (1981).Biochemical assessment of the relationships of *Eragrostis tef* (Zucc) trotter with some wild *Eragrostis* species (Gramineae). *Ann Bot.* **48**: 717–725.

Bielawski, J. P., and Noack, K. (1995). Reproducible amplification of RAPD markers from Vertebrate DNA. *Bio. Techniques,* **18**: 856-860.

Bornet, B., and Branchard, M. (2001). Non-anchored Inter Simple Sequence Repeat (ISSR) Markers: Reproducible and Specific Tools for Genome Fingerprinting. *Pl Molec Biol. Reporter,* **19**: 209–215.

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

Brookes, A. (1999). The Essence of SNPs, *Gene.* **234**: 177–186

Calinski, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*. **3:** 1–27.

Cannarozzi, G., Plaza-Wuthrich, S., Esfeld, K., Larti, S., Wilson, Y.S., Girma, D., De Castro, E., Chanyalew, S., Blosch, R., Farinelli, L., Lyons, E., Schneider, M., Falquet, L., Kuhlemeier, C., Assefa, K., and Tadele, Z. (2014). Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (*Eragrostis tef*). *BMC genomics* 15.

Chahal, G.S. and Gosal, S.S., (2002). Principles and Procedures of Plant Breeding*:* Biotechnological and conventional approaches. Narosa Publishing House, New Delhi.

Chanyalew, S.,Tefera, H., Harjit, S. and Sorrells, M.E. (2007). Comparison of AFLP,EST-SSR, ISSR and SSR markers for polymorphism among recombinant inbred lines of tef (*Eragrostis tef*). *J Genet Breed*. **61**: 27–34.

Cheng A, Mayes S, Dalle G, Demissew S, Massawe F. (2017). Diversifying crops for food and nutrition security - A case of tef. *Biol Rev Camb Philos Soc*. **9**: 188-98.

Cho,YG., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, SR., Park, WD., Ayres, N. and Cartinhour, S. (2000). Diversity of microsatellites derived from genomic libraries and Gene Bank sequences in rice (Oryza sativa L.). *Theor Appl Genet.* **100**: 713–722.

Coulondre, C., Miller, J.H., Farabaugh, P.J., and Gilbert, W. (1978). Molecular Basis of Base Substitution Hot Spots in *Escherichia coli*. *Nature,* **274**: 775–780.

Crymes, AR. (2018). The international footprint of tef: Resurgence of an ancient Ethiopian grain.

CSA. (2014). Agricultural Sample Survey for2013/14.Statistical Bulletin 532.Addis Ababa, Ethiopia: Central Statistical Agency.

D'Andrea, A. (2008). Tef (*Eragrostis tef*) in ancient agricultural systems of Highland Ethiopia. *Econ Bot.* **62**: 547–566.

Datta, D., Gupta, S., Chaturvedi, S.K. and Nadarajan, N. (2011). Molecular Markers in Crop Improvement. Indian Institute of Pulses Research, *Kanpur* **2**: 8-24.

Demissie, A. (1991). A decade of germplasm exploration and collecting activities by the Plant Genetic Resources Centre /Ethiopia. Cambridge: Cambridge University Press.

Drikvand, R.M., Elham, G.N., Salahvarzi, A. (2015). Investigation of genetic diversity of some durum and bread wheat genotypes using SSR markers. *J of Biodiv. and genet.*

Duda, R. O. and Hart, P. E. (1973). Pattern classification and scene analysis, New York: John Wiley & Sons. **15**: 77-79.

EarlDent.A. and Vonholdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Cons Genet Resor.* **4**: 359-361.

Ebba, T. (1975). Tef Cultivars: Morphology and Classification. Dire Dawa: University, College of Agriculture.

Espelund, M., Bekele, E., Holst-Jensen, A., Jakobsen, K.S. and Nordal, I. (2000). A molecular genetic analysis of tef (*Eragrostis tef* (Zucc.)Trotter: non-coding regions of chloroplast DNA, 18Sr DNA and the transcription factor VP1. *Hereditas* **132**: 193–202

Evanno, G., Regnaut, S. and Goudet, J. (2005). Detecting the number of clusters of individua -ls using the software STRUCTURE: a simulation study. *Mol Ecol.* **14**: 11-20.

FAO. (2015). Analysis of price incentives for Tef in Ethiopia Technical notes series, MAFAP, by Assefa B. Demeke M., Lanos B, Rome.

Gepts, P. and Papa, R., (2003). Possible effects of (trans) gene flow from crops on the genetic diversity from landraces and wild relatives. *Enviro Biosafety Resear***. 2**:89-103.

Glaubitz, J. C., and Moran, G. F. (2000). Genetic tools: the use of biochemical and molecular markers. In Young, A., Boshier, D., and Boyle, T., eds. Forest Conservation Genetics-Principles and Practice. CSIRO Publishing and CABI Publishing. Pp. 39-59.

Godwin, I. D., Mace, E.S. and Nurzuhairawaty (2001). Genotyping Pacific Island Taro (*Colocasiaesculenta* (L.) Schott) Germplasm. In Henry, R. J., ed. Plant Genotyping: the DNA Fingerprinting of Plants. CABI publishing. Pp. 109-128.

Gomez, A.K. and Gomez, A.A. (1984). Statistical Procedure for Agricultural Research. John Wiley and Sons. London.

Govindaraj, M., Vetriventhan, M. and Srinivasan, cM. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Grown in Brunei Darussalam and assessment of their tolerance to saline environment.

Gupta, P.K., Varshney, R.K., Sharma, P.C., and Ramesh, B. (1999). Molecular Markers and Their Applications in Wheat Breeding. *Plant Breed.***118**: 369–390.

Hamrick M.S. (2014). Assessment of the degree of genetic divergence among genetically *Hereditas*133: 29–37. Hamrick JL, Linhart YB, Mitton JB. Relationship between life history characteristics and electrophoretically detectable genetic variation in plants. *Annu. Rev Ecol Syst.* **10**: 173-200.

Hoelzel, A. R., and Green, A. (1998). PCR protocols and population analysis by direct DNA sequencing and PCR-based DNA finger printing. In Hoelzel, A. R., ed. Molecular genetic Analysis of Populations-A practical approach. Oxford University Press. Pp. 201-233.

Hua, W., Zhang, X., Zhu, J., Shang, Y., Wang, J., Jia, Q., Li, C. and Yang, J. (2015). A study of genetic in Jatropha: a review. *The Open Renewable Energy*, **8**: 1–6.

Jingura, T., Kamusok, R.M. and Jingura, R. (2015). Utility of markers for determination of genetic diversity in Jatropha: a review The Open Renewable Energy. *Sci. Front. Res.* **8** : 1-6.

Johnson, R.A. and D.W. Wichern, (1988). Applied Multivariate Statistical Analysis. Prentice Hall, Englewood Cliffs, NJ.

Jonah, PM., Bello, L. L., Lucky, O., Midau, A., Moruppa, SM. and Moruppa, SM. (2011). Review: The importance of molecular markers in plant breeding programmes. *Glo J of Sci Front Res.* **11**: 5–12.

Jones, B. M. G., Ponti, J., Tavassoli, A. and Dixon, P.A. (1978). Relationships of Ethiopian cereal tef (*Eragrostis tef* (Zucc)Trotter) evidence from morphology and chromosome-number. *Ann. Bot.* **42**: 1369–1373.

Kalinowski S. T. (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes*. **5**: 187–189.

Karp, A., Kresovich, S., Bhat, K.V., Ayad, W.G. and T. Hodgkin, (1997). Molecular tools in plant genetic resources conservation: a guide to the technologies. International Plant Genetic Resources Institute (IPGRI) Technical Bulletin, No.2: Rome, Italy.

Kebede, H., Johnson, R.C. and Ferris, D.M. (1989). Photosynthetic response of tef (*Eragrostis tef*) to temperature. *PhysiolPlant.* **77**: 262–266.

Keneni, G., Bekele, E., Imtiaz, M. and Dagne, K. (2012). Genetic vulnerability of modern crop Cultivars: Causes, mechanism and remedies. *IJ of Plant Res*. **2**: 69-79.

Ketema, S.(1997). Tef *(Eragrostis tef (Zucc.)Trotter.* Rome: Institute of Plant Genetics and Crop Plant Research, Gatersleben / International Plant Genetic Resources Institute.

Ketema, S. (1993).Tef *(Eragrostis tef):* Breeding, Agronomy, Genetic Resources, Utilization, and Role in Ethiopian Agriculture. Addis Ababa: Institute of Agricultural Research.

Khlestkina, E. K. and Salina, E. A. (2006).SNP markers: Methods of analysis, ways of development, and comparison on an example of common wheat *Genetika*, **42**: 725-36.

Koebner, R. M. D. (1995).Pre-digestion of DNA template improved the level of polymorph -ism of Random Amplified Polymorphic DNAs in wheat. Genetic Analysis. *Biomol. Engin*. **12**: 63-67.

Kumar, P., Gupta, V.K., Misra, A.K., Rodi, D.R. and Panday, B.K., (2009). Potential of molecular markers in plant Biotechnology. *Southern Cross J.* **2**: 141-162.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg N. A., Mayrose, I. (2015). CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Res*. **15**: 1179–1191.

Kubik, C., M. Sawkins, W.A. Meyer, and B.S. Gaut, (2001). Genetic diversity in seven perennial ryegrass (Loliumperenne L.) cultivars based on SSR markers. *Crop Sci*. **41**:1565-1572.

Laca, M. (2009.). Grain production of 15 tef varieties grown in churchill county, Nevada during 2009.

University of Nevada Cooperative Extension.

Liu, K. J. and Muse, S. V. (2005). Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinfor.* **21** : 21-28.

Liu,B., and Wendel, J. F. (2001). Inter simple sequence repeat (ISSR)polymorphisms as a genetic marker system in cotton. *Mol Ecol Notes,***1**: 205-208.

McDermott, J. M. and McDonald, B.A. (1993). Gene flow in plant pathosystems. *Annu. Rev. phytopathol***. 3**: 353-373.

McGregor, C. E., Lambert, C. A., Greyling, M. M., Louw, J. H. and Waenich, L. (2000). A comparative assessment of DNA finger printing techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*solanumtuberosum* L.). *Euphytica*, **113**: 135-144.

Micheli, M. R., Bova, R.,Pascale, E., and D'Ambrosio, E. (1994). Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucl Acids Res.* **22**: 1921-1922.

Mignouna, H.D., Abang, M. M. and Fagbemi, S. A. (2003). A comparative assessment of molecular marker assays (RAPD, SSR and AFLP) for white yam germplasm characterization. *Ass App lBiol*. **142**: 269-276.

Milligan, G. W. and Cooper, M. C., (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometri*. **50**: 159–179.

Minten, B., Tamru, S., Engida, E. and Kuma T. (2016). Transforming staple food value chains in Africa: The case of tef in Ethiopia. *J Dev Stud*, **52**: 27-45.

Minten, B.Tamru, S. Engida, E. and Kuma T. (2016). (a) Feeding Africa's cities: The case of the supply chain of tef to Addis Ababa. *Econ Dev Cult Change*. **64**: 65-97.

Morgante, M., and Olivieri, A. M. (1993). PCR amplified microsatellites as markers in plant genetics. *J P Sci.* **3**:175-182.

Mueller, U. G., and Wolfenbarger, L.L. (1999). AFLP genotyping and finger printing. *Trends in Ecol Evol.***14**: 389-394.

Nagaoka, T. and Ogihara, Y. (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theor Appl Genet*. **94**: 597-602.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**: 583-590.

Nei, M. and Li, W. (1979). Mathematical method for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Acad of Sci*. **76**: 5256–5273.

Oh, Y., and Mao, L. (1999). Microsatellite protocols. In Innis, M. A., Gelfad, D. H., and Sninsky, J. J., eds. PCR applications Protocol for functional genomics. Academic Press.

Park, Y., Lee, J.K. and Kim, N., (2009). Simple Sequence Repeat Polymorphisms (SSRs) for evaluation of molecular diversity and germplasm classification of minor crops. pp. 4546–4569.

Peakall, R., Smouse, P.E. (2007). GENALEX 6.5 Genetic Analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes.* **6**: 288–295.

Perrier, X. and Jacquemoud-Collet, J. P. (2006). DARwin software.http://darwin.cirad.fr/.

Petros, Y., Merker, A. and Zeleke, H. (2007). Analysis of genetic diversity of *Guizotiaabyssinica* from Ethiopia using inter simple sequence repeat markers. *Hereditas*.**144**: 18-24.

Phillips, S. (1995). Flora of Ethiopia and Eritrea Poaceae (Gramineae). National Herbarium, Addis Ababa University, Addis Ababa, Ethiopia and Department of Systematic Botany, sala University, Uppsala, Sweden. Pp. 2.

Prasad, M., Varshney, R.K., Roy, J.K., Balyan, H.S. and Gupta, P. K. (2000). The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. *Theor Appl Gen.* **100**: 584-592.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multi-locus genotype data. *Genet*. **155**: 945–59.

Rafalski, J. A. (1997). Random amplified polymorphic DNA (RAPD) analysis. In Caetano- Anolles, G., ed. DNA Markers Protocol, Applications, and Over views. Wiley-VCH.

Ranade, S.A. and Yadav, H. (2014). Universal molecular markers for plant breeding and genetics.

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

Riede, C. R., Fairbanks, D. J., Andersen, W. R., Kehrer, R.L. and Robinson, L. R. (1994). Enhance ment of RAPD Analysis by restriction endonuclease digestion of template DNA in wheat. *P Breed*. **113**: 254-257.

Russell, J. R., Fuller, J.D., and Macaulay, M. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and repeats. *Plant Genet Res.* 1-11.

SAS Institute, (2010). SAS/STAT Guide for Personal Computers, Version 13.1 editions. Cary, N.C., SAS Institute Inc.

Scandalios. J .G. (1969). Genetic control of multiple forms of enzymes in plants; a review. *Biochem Genet*. **3** : 37-79.

Senan, S., Kizhakayil, D., Sasikumar, B. and Sheeja, T.E. (2014). Methods for development of microsatellite markers: an overview. *Nat Sci Bio*. **6**: 1–13.

Seyfu, K. (1987). Research recommendations for production and brief outline of strategy for the improvement of tef [*Eragrostis tef* (Zucc.)Trotter]. In: Proc. 19th Natl. Crop. Imp. Conf. IAR. Addis Ababa, Ethiopia.

Schulman, A., Flavell, A. and Ellis, T., (2004). The application of LTR retro transposons as molecular markers in plants. *Methods Molec Biol*. **260**: 145-173.

Shashidhar, H.E., Kanbar, A., Toorchi, M., Raveendra ,G.M., Kundur, P., Vimarsha, H.S., Soman, R, Kumar, NG, Bekele, B.D. and Bhavani, P. (2013). Breeding for drought resistance using whole plant architecture: conventional and molecular approach.

Slatkin, M. and Takahata, N. (1985). The average frequency of private alleles in a partially isolated

population. *Theor Popul Biol*. **28**: 314-331.

Smith, J. S. C., Kresovich, S., Hopkins, M.S., Mitchell, S.E., Dean, R.E., Woodman, W.L., Lee, M. and Porter, K. (2000). Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. *Crop Sci.* **40**: 226–232.

Smith, J.S.C., Chin, E.C.L., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovich, S. and Ziegle, J. (1997). An evaluation of the utility of SSR loci as molecular markers in maize (Zea mays L.): comparisons with data from RFLPs and pedigree. *Theor and Appl Genet.* **95**: 163-173.

Smulders, M.J.M., G. Bredemeijer, W. Rus-Kortekaas, P. Arens, and B.Vosman. (2003). Use of short microsatellites from database sequences to generate polymorphisms among Lycopersiconesculentum cultivars and accessions of other Lycopersicon species. *Theor Appl Genet.* **97**: 264-272.

Somers, D.J., Kirkpatrick, R., Moniwa, M., and Walsh, A. (2003). Mining Single-Nucleotide Polymorphisms from Hexaploid Wheat EST. *Genome,* **49**: 431–437.

Spaenij-Dekking, L., Kooy-Winkelaar, Y., and Koning, F. (2005). The Ethiopian cereal tef in celiac disease. *The New Engl. J of medici,* **353:** 1748-1749.

Tadele, Z. and Assefa, K. (2012). Increasing food production in Africa by boosting the productivity of understudied crops. *Agron.* **2**: 240–283.

Tadessa, D. (2017). Nutritional and Soio-Cultural Values of Tef (*Eragrostis tef*) in Ethiopia. *Intern J of Food Sci. and Nutrit*. **2**: 50-57.

Tadesse, E. (1975). Tef (*Eragrostis tef*) cultivars Morphological and classification. Part Iagri cultural Experimental Station Bulletien, 66, Addis Ababa University, College of Agriculture, Dire Dawa, Ethiopia.

Taramino, G. and Tingey, S.V, (1996). Simple sequence repeats for germplasm analysis and mapping in maize. *Geno.,* **39**: 277-287.

Tavassoli, A. (1986). The cytology of Eragrostis with special reference to E. tef and its relatives. Ph.D. dissertation, London University, London, UK

Tefera, H., Assefa, K. and Belay, G. (2003a). Evaluation of interspecific recombinant inbred lines of *Eragrostis tef* x *E pilosa*. *J Genet Breed.* **57**: 21–30.

Tesema, A. (2013). "Genetic diversity of tef in Ethiopia,"inachievements and prospects of tef Improvement, edsA. Assefa, S. Chanyalew, and A. Tadele (Bern: EIAR-University of Bern), 15–20.

VanBuren, R., Wai,C.M., Keilwagen,J. and Pardo, J. (2018). A chromosome-scale assembly of the model desiccation tolerant grass *Oropetium thomaeum. Ameri society of pl boil.*

Vavilov, N. (1951). The origin, variation, immunity and breeding of cultivated plants. Translated from the Russin by K. Starrchester, The Ronald Press Co. New York, USA.

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

Wang, D.G., Fan, J.B. and Siao, C.J. (1998). Large-Scale Identification, Mapping, and Genotypin-g of Single-Nucleotide Polymorphisms in the Human. *Genome Sci.* **280:** 1077–1082.

Wang, M. Oppedijk, B. J. and Lui, X. (1996). Apoptosis in barley aleurone during germination and its inhibition by Abscisic acid. *Plant Mol Biol.* **32**: 1125-1134.

Waples, R.S. (1987). A multi species approach to the analysis of gene flow in marine shore fishes. *Evol*. **41**: 385-400.

Weber, J. L. and May, P. E. (1989). Abundant class of Human DNA polymorphism which can be typed using the Polymerase Chain Reaction. *Amer J Hum Genet*. **44**: 388-396.

Weising, K., Atkinson, R. G., and Gardner, R. C. (1995). Genomic fingerprinting by Micro- satellite-primed PCR: A critical evaluation. PCR Methods Application **4**: 249 -255.

Westman, A. L. and S. Kresovich, (1999). Simple sequence repeat (SSR)-based marker variation in Brassicanigra gene bank accessions and weed populations. *Euphytica,* **109**: 85-92.

Zeid, M., Assefa, K., Haddis, A., Chanyalew, S. and Sorrells, M.E. (2012). Genetic diversity in tef

(*Eragrostis tef*) germplasm using SSR markers. *Field Crops Res.* **127**: 64–70.

Zhu, F. (2018).Chemical composition and food uses of tef (*Eragrostis tef*). *Food Chem*; **239**: 2-15.

Zietkiewicz, E., Rafalski, A., and Labuda, D. (1994). Genome fingerprinting by Simple Sequence Repeat (SSR)-Anchored Polymerase Chain Reaction Amplification. *Genomi*. **20**: 176-183.

7. APPENDICES

Appendix 1. Plant DNA Extraction Protocol for DArT

BUFFER STOCK SOLUTIONS

EXTRACTION BUFFER STOCK
To make 500 ml:

0.35 M sorbitol	31.9 g sorbitol
0.1 M TrisHCl pH 8.0	50 ml 1M TrisHCl pH 8.0
5 mM EDTA pH 8.0	5 ml 0.5 M EDTA pH 8.0
fill up to 500 ml MiliQ H2O	

LYSIS BUFFER STOCK
To make 500 ml:

0.2 M TrisHCl pH 8.0	100 ml 1M Tri HCl pH 8.0
0.05 M EDTA pH 8.0	50 ml 0.5 M EDTA pH 8.0
2M NaCl	200 ml 5 M NaCl
2% CTAB	10 g CTAB fill up to 500 ml with MilliQ H20

SARCOSYL STOCK 5% (w/v)

FRESH BUFFER WORKING SOLUTION*:

0.5 % (w/v) sodiumdisulfite(= sodium metabisulfite)

2 % (w/v) PVP-40 (K29-32) Sigma dissolve in required volume of extraction buffer stock; add same volume of lysis buffer stock and 0.4 volume of extraction (=lysis) buffer stock of sarcosyl stock.

For example to make 120 ml:

Add 0.6 g sodium disulfite (= sodium metabisulfite) and 2.4 g PVP-40 (K29-32) to 50 ml extraction buffer stock and dissolve; add 50 ml lysis buffer stock and 20 ml sarcosyl stock

F or example to make 30 ml:

Add 0.15 g sodiumdi sulfite (= sodium metabisulfite) and 0.6 g PVP-40 (K29-32) to 12.5 ml extraction buffer stock and dissolve; add 12.5 ml lysis buffer stock and 5 ml sarcosyl stock

*This buffer may settle into two layers on standing. Heat to 65ºC and shake immediately before adding to extraction tubes.

PROTOCOL
For 15 ml Sarsted tubes:

- ✓ Aliquot 6 ml of freshly prepared preheated to 65ºC well mixed "fresh buffer solution" and place tubes to the 65ºC incubator or water bath, (3, 4 days old "fresh buffer solution" works fine),
- ✓ Grind required amount (same across all samples) of plant material in mortar and pestle under liquid nitrogen to fine powder,
- ✓ Suspend powder in 6 ml "fresh buffer solution" kept at 65ºC (make sure there are no clumps, vortex if necessary),
- ✓ Incubate at 65ºC for 1 h (can extend for another 30 min), invert tubes in every 20 minutes or incubate with gentle shaking,
- ✓ Cool down for 5 min and add 6 ml of chloroform : isoamyl alcohol (24 : 1) mixture,
- ✓ Mix well for 30 min,
- ✓ Spin 20 min, 3000 x g, RT,
- ✓ Transfer water phase to fresh tube, add same volume of ice cold isopropanol and invert tube ~ 10times, nucleic acids should become visible,
- ✓ Spin 30 min, 3000 x g, RT,
- ✓ Discard supernatant, wash pellet with 2 ml 70 % EtOH,
- ✓ Discard EtOH, dry pellet and dissolve in 200 µl – 500 µl 1 x TE (10 mMTrisHCl pH 8.0, 1 mMEDTA pH 8.0),
- ✓ Check DNA quality and quantity on 0.8 % agarose gel. (If RNA quantity is several fold less than DNA, RNase treatment is not necessary for DArT applications).
- ✓ For 2 ml Eppendorf tubes:
- ✓ Aliquot 1 ml of freshly prepared preheated to 65ºC, well mixed "fresh buffer solution" and place tubes to the 65ºC incubator or water bath, (3, 4 days old "fresh buffer solution" works fine),
- ✓ Grind required amount (same across all samples) of plant material in mortar and pestle under liquid nitrogen to fine powder,
- ✓ Suspend powder in 1 ml "fresh buffer solution" kept at 65ºC (make sure there are no clumps, vortex if necessary),
- ✓ Incubate at 65ºC for 1 h (can extend for another 30 min), invert tubes in every 20 minutes or incubate with gentle shaking,
- ✓ Cool down for 5 min and add 1 ml of chloroform: isoamyl alcohol (24 : 1) mixture,
- ✓ Mix well for 30 min,
- ✓ spin 20 min, 10000 x g, RT,

- ✓ Transfer water phase to fresh tube, add same volume of ice cold iso-propanol and invert tube ~ 10times, nucleic acids should become visible,
- ✓ Spin 30 min, 10000 x g, RT,
- ✓ Discard supernatant, wash pellet with 2 ml 70 % EtOH,
- ✓ Discard EtOH, dry pellet and dissolve in 250 µl of 1 x TE (10 mMTrisHCl pH 8.0, 1 mM EDTA pH8.0),

Check DNA quality and quantity on 0.8 % agarose gel. (If RNA quantity is several fold less than DNA, RNase treatment is not necessary for DArT applications).

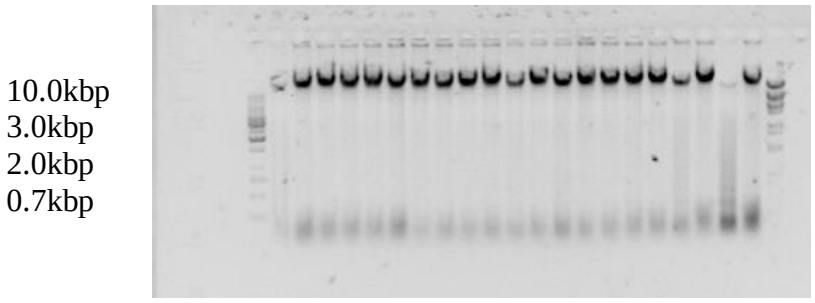


Figure 1. An example of sugarcane DNA extracted with the described method

Appendix 2. DILUTION OF PRIMERS FOR PCR

1. Currently primers are ordered from SIGMA GENOSYS or others depending on price (remember to consider delivery charges and difference costs with redundancies).

2. Use the website order system that contains correct information about scales available - if you are ordering more than c. 6 primers, the bulk submission is simpler.

3. Order small scale: 0.05 um or (less) 3 OD (but note from Sigma-Genosys that the minimum scale is the same price as the next size up so order the larger; sometimes the minimum scale is not available for oligos containing redundancies or modifications).

4. When the oligos come, put the original datasheet and gel picture in the Oligo Master Book. Make a photocopy for yourself.

5. In the laminar flow hood, reconstitute the dried oligos in (e.g. SIGMA) molecular biology grade water to make a 100 µM (micro-molar) stock solution; in addition to the original stock tube make two 20µl (micro-litre) stock aliquots. (Oligos will be more stable in slightly buffered salts, e.g. TE, but this can inhibit enzyme reactions.) The data sheet will give the number of ul of water that are needed to make this dilution - typically 300 to 800 µl. (They also give the nanomoles in the synthesis - e.g. 46.6 nmole. To make the 100 uM stock, multiply this by 10 and add than many ul of water (e.g. 233 µl for 46.6 nmol = 200 µM stock solution).

6. Write the name of the oligo on the cap of the stock tube. Then put the original stock in a common oligo box (normally large fluorescent boxes in tall freezers in Trude's or Pat's drawer).

7. Small aliquots of specific oligo stocks can be kept in individual boxes in your freezer compartment..

8. Make a small amount of working solution by diluting aliquoted 100 µM stock (in the laminar flow hood) with molecular biology grade water. 1:10 giving a 10 µM solution for genomic PCR If you use 1 µl of this stock in a 20 µl PCR mix, you are using 0.4 µM of primer in the final mix. 1:20 giving a 5µM for plasmid PCR

9. PCR set-up using York Bio or Promega Taq polymerase (formerly we used Bioline, now more expensive):

- water (add first)
- 1.5 µ1 of 50mM MgCl (final conc. 1.5 mM)
- 5 µl PCR buffer (comes with Taq)
- 1 µl 10 mMdNTPmix or 4µ1 of 2.5 mMdNTPmix (final conc. 200 µM; see below)
- 2.5 µl of each primer'
- 5-10 ng of DNA
- 0.1-0.5 µl Taq polymerase (=1 unit to 5 units)
- Total 50µl

Notes

· Scaling down of PCR set-up to 20-30µl for marker typing and testing (even as low as 7µl is possible). Only use 50 µl wher you need to cut out and clone the dNTP mix: we used to make a 2.5mM working solution in Tris-HCI but ROCHE recommends to make a 10mM solution by

adding 5µ1 of dATP, dCTP, TTP and dGTC (each 100mM; as supplied by ROCHE) to 30µl sterile double distilled water.

BE AWARE THAT DILUTED OLIGOS CAN BREAK DOWN WHEN DEFROSTED SEVERAL TIMES.

ALWAYS WEAR GLOVES WHEN HANDLING OLIGOS.

M13 primer STOCK AND DILUTED ALIQUOTS WILL BE KEPT IN THE GENERAL PCR BOX.

Appendix 3. 6X DNA Loading Dye

6X DNA Loading Dye is ideal for preparing DNA markers and samples for loading on agarose or polyacrylamide gels. The solution allows for easy visual tracking of DNA migration during electropho resis.

Advantages

- Loading DNA samples into gel electrophoresis wells
- Easy visual tracking of DNA migration during electrophoresis
- Dry Ice is not required for shipping

Composition

- 10 mM Tris-HCl (pH 7.6)
- 60 mM EDTA
- 0.1% bromophenol blue
- 0.1% xylene cyanol FF
- 50% glycerol
- Quality Control

The quality of the 6X DNA Loading Dye is tested on a lot-to-lot basis according to Geneaid's ISO-certified quality management system.

Storage

4ºC up to 12 months or -20ºC >12 months

Add a 1/5 volume of 6X DNA Loading Dye to the DNA sample.

Appendix 4. General PCR Protocol

Sally A. Green, 2-17-96

Here in the Maddock Lab, we do 25µL PCR reactions in 0.5mL microfuge tubes. You can do PCR in different size reaction volumes and in smaller tubes as long as they fit in the thermo cycler.

Materials:

Template DNA (genomic, plasmid, cosmid, bacterial/yeast colony, etc.) primers (resuspended to a known concentration with sterile TE) buffer (usually 10X, usually sold with Taq polymerase or you can make your own)

note: some different buffer recipes follow at the end of this protocol $MgCl_2$ (25mM is convenient)

Taq polymerase

dNTPs (2mM stock)

note: a 2mM stock of dNTPs means that the final concentration of each dNTP (dATP, dCTP, dGTP, and dTTP) is 2mM -- NOT that all dNTPs together make 2mM. dNTPs come as 100mM stocks -- thaw and add 10µL of each dNTP to 460µL of ddH_20 to make 2mM. Store at -20°C.

sterile ddH_20 glovesPCR machine aerosol tips, if desired

PCR is very sensitive to contamination from outside DNAs. Steps should be taken to reduce the chance for contamination, such as wearing gloves, using aerosol tips (tips with a wad of cotton at the top), and not spitting in the tubes. I don't use aerosol tips and have dispensed with the gloves -- just be careful. Something that IS important is to assemble your reactions on ice.

The final concentrations of reagents in PCR reactions are as follows:

buffer: 1X, usually comes as 10X stock. For 25µL reactions, this means 2.5µL.dNTPs: for most general PCR, you want the final concentration to be 200µM, so a 2mM stock is essentially 10X -- use 2.5µL per reaction .primers: a good place to start with primer concentration is 50pmol of each primer per reaction. If you don't get your desired product, you can increase to 75pmol or 100pmol. This usually does the trick. I wouldn't go past 200pmol of primers unless it's a special protocol that recommends using more. note: hints on resuspending primers follow this protocol

Template: it's not usually necessary to be incredibly fastidious about how much template you add to a reaction. You can get product with incredibly small amounts of starting DNA. I usually do a 3mL plasmid prep and use 1/6 of a micro liter per PCR reaction. You can use more or less -- it doesn't seem to matter that much.

$MgCl_2$**:** this is the greatest variable in PCR. The success of a PCR is very dependent on how much magnesium is present in the reaction. For this reason, it is usually advisable to do a magnesium optimization when performing new PCRs. I do my reactions in sets of six, keeping all variables constant except for magnesium. I usually go from 1mM to 6mM $MgCl_2$. Since the stock is 25mM, usually, this means that 1µL of stock equals 1mM MgCl2 in a 25µL reaction -- it's convenient. Given that you're probably going to do at minimum six PCR reactions with six different magnesium concentrations, and you will have to add several different ingredients, it's useful to take steps to reduce the number of pipetting steps, which in turn reduces time and change for cross-contamination. The way I set up my reactions is as follows:

Appendix 5. Tef populations codes used for the morphological and molecular data analysis

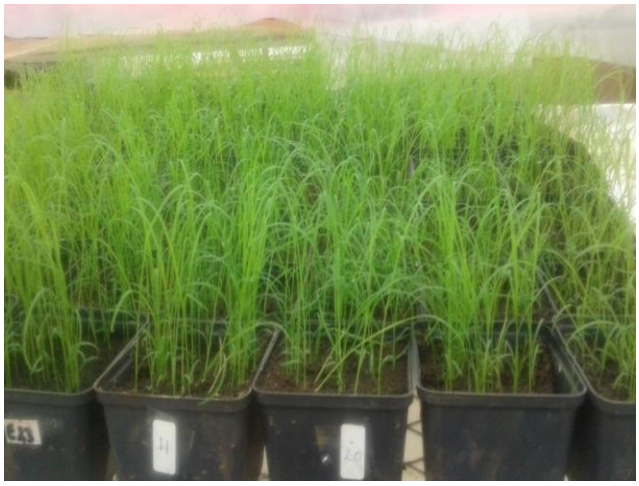
Acc. Number	Accession code	Population Group	Population code
15309	1	SEMEN SHEWA	1
236957	2	SEMEN SHEWA	1
236959	3	SEMEN SHEWA	1
236960	4	SEMEN SHEWA	1
236961	5	SEMEN SHEWA	1
243493	6	DEBUB WELLO	2
243502	7	DEBUB WELLO	2
243507	8	SEMEN WELLO	2
243508	9	SEMEN WELLO	2
243509	10	SEMEN WELLO	2
243529	11	SEMEN GONDAR	3
243535	12	SEMEN GONDAR	3
243536	13	SEMEN GONDAR	3
243537	14	SEMEN GONDAR	3
243538	15	SEMEN GONDAR	3
243539	16	SEMEN GONDAR	3
243540	17	SEMEN GONDAR	3
243544	18	SEMEN GONDAR	3
243545	19	SEMEN GONDAR	3
243547	20	MIRAB GOJAM	4
243548	21	MIRAB GOJAM	4

243549	22	MIRAB GOJAM	4
243550	23	MIRAB GOJAM	4
243551	24	MIRAB GOJAM	4
243553	25	MIRAB GOJAM	4
243512	26	MEHAKELEGNAW	5
243517	27	MEHAKELEGNAW	5
243518	28	MEHAKELEGNAW	5
243519	29	MEHAKELEGNAW	5
243520	30	MEHAKELEGNAW	5
243521	31	MEHAKELEGNAW	5
243524	32	MEHAKELEGNAW	5
243525	33	MIRABAWI	6
243526	34	MIRABAWI	6
243527	35	MIRABAWI	6
243528	36	MIRABAWI	6
244789	37	KEMBATA ALABANA TEMB	7
244790	38	KEMBATA ALABANA TEMB	7
244791	39	KEMBATA ALABANA TEMB	7
244792	40	KEMBATA ALABANA TEMB	7
244793	41	KEMBATA ALABANA TEMB	7
244794	42	KEMBATA ALABANA TEMB	7
244796	43	KEMBATA ALABANA TEMB	7
244814	44	MISRAK WELLEGA	8
244815	45	MISRAK WELLEGA	8
244816	46	MISRAK WELLEGA	8
244817	47	MISRAK WELLEGA	8
244818	48	MISRAK WELLEGA	8
244821	49	MISRAK WELLEGA	8
244822	50	MISRAK WELLEGA	8
244850	51	MISRAK WELLEGA	8
244851	52	MISRAK WELLEGA	8
244852	53	MISRAK WELLEGA	8
244853	54	MISRAK WELLEGA	8
244854	55	MISRAK WELLEGA	8
244856	56	MISRAK WELLEGA	8
244876	57	ILLUBABOR	9
244877	58	ILLUBABOR	9
244878	59	ILLUBABOR	9
244879	60	ILLUBABOR	9
244880	61	ILLUBABOR	9
244881	62	ILLUBABOR	9
244882	63	ILLUBABOR	9
244883	64	ILLUBABOR	9

Appendix 6. Distribution of 64 tef accessions in to eight clusters using 12 quantitative traits

Cluster	Number of accessions	Accessions included in the cluster
1	33	1 6 10 11 13 14 15 18 19 20 22 24 25 28 29 30 31 32 33 35 36 38 40 45 47 49 52 54 60 61 62 63
2	4	2 21 50 59
3	5	7 12 23 53 56
4	7	4 17 34 51 55
5	2	26 44
6	11	8 9 16 37 39 42 46 48 57 58 64
7	3	3 41 43
8	1	5

Appendix 7. A) Tef seedlings B) Tef leaf under silica gel used during DAN extraction.



A

B