ADDIS ABABA UNIVERSITY

Microbial, Cellular and Molecular Biology Program Unit

Applied Genetics stream

GENETIC DIVERSITY OF AN AFRICAN WILD RICE (*Oryza longistaminata* Chev. et Roehr) FROM AMHARA AND GAMBELA REGIONS AS REVEALED BY MICROSATELLITE MARKERS

By

Getachew Melaku

A Thesis submitted to the School of Graduate Studies, Addis Ababa University, in partial fulfilment of the Requirements for the Degree of Master of Science in Biology (Applied Genetics)

November 2011

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

GENETIC DIVERSITY OF THE AFRICAN WILD RICE (*Oryza longistaminata* Chev. et Roehr) FROM AMHARA AND GAMBELA REGIONS AS REVEALED BY SSR MARKERS.

By

Getachew Melaku

A Thesis presented to the School of Graduate Studies, Addis Ababa University, in partial fulfilment of the Requirements for the Degree of Master of Science in Biology (Applied Genetics)

Approved by the Examining Board:

Name	Signature
Kifle Dagne (PhD)	_____
Examiner	
Kassahune Tesfaye (PhD)	_____
Examiner	
Teklehaimanot Haileselassie (PhD)	_____
Research Advisor	
Tielye Feyissa (PhD)	_____
Research Advisor	
Fassil Assefa (PhD)	_____
Chairman	

Acknowledgements

I would like to thank my Advisors, Dr. Teklehaimanot Hayleselassie and Dr. Tileye Feyissa, not only for accepting me as a student but also for the good suggestion on the manuscript and their guidance and also for providing me everything they have for the initial motivation and encouragement that lead me to the realization of the present thesis.

I acknowledge the financial support from USAID under the Program for Biosafety Systems (PBS) lead by Dr. Teklehaimanot Haileselassie and the Microbial, Cellular and Molecular Biology Program Unit–Addis Ababa University. The invitation to do all the practical activities in the Genetics Laboratory of the University of Nairobi and BECA together with The Ethiopian Meteorological Authority for providing me the recent 17 years climatic data is also highly appreciated.

I am grateful to the technical assistance of Dr. Sam Kiboi and Franklin Irrieri from University Of Nairobi and the PhD. Candidate Mr. Asfaw Adugna from the applied Genetics stream of Addis Ababa University.

I extremely wish to thank Dr. Kifle Dagne for his more than fatherhood concern and support.

Dr. Sisay Kebede, Mr. Ayenew Sendi, Mr. Ephrem Hailu, Mr. Abel Gizaw and Mr. Yonas W/Mariam also deserve particular mention and acknowledgements for their encouragement and nice friendship under all of my difficult moments.

I certainly must acknowledge my parents and sisters for their day to day care, support and for all what they did for me.

May God bless you all!!

Contents Pages

List of Tables Pages

List of Figures

Pages

List of Abbreviations

AFLP-Amplified Fragment Length Polymorphism

ASO-Allel-specific oligonucleotides

ESTs-Expressed sequence tags

ISSR-Inter Simple Sequence Repeat

MAS- Molecular assisted selection

OLA-Oligonucleotide ligation assay

PCO-Principal coordinated Analysis

PCR-Polymerase chain reaction

QTL- Quantitative trait loci

RFLP-Restriction Fragment Length Polymorphism

RAPD-Random Amplified Polymorphic DNA

SNPs- Single Nucleotide Polymorphisms

SSR- Simple sequence repeats

Abstract

Data from microsatellite markers have been extensively used for both *in situ* and *ex situ* conservation strategies through determining the level of genetic diversity of natural populations that can widen the gene pool of cultivated plants. Such conservation practices should also lie on understanding the between and within population genetic variation and partitioning populations on the basis of geographic origin. Therefore, the objective of this study was to assess the intra-specific diversity in *Oryza longistaminata* and how the variation is partitioned within and between different geographic regions of the eight *Oryza longistaminata* populations in Ethiopia using simple sequence repeat markers. The five microsatellite markers in 320 samples generated 64 allels that revealed the presence of large amount of genetic variability (Ho =0.225; He = 0.768; Na = 7.375; Ne = 6.565 and P = 0.744). The F-statistics detected by the microsatellite loci showed Fst = 0.064 and Fis = 0.743 and there was no population in Hardy-Weinberg equilibrium. The genetic parameters estimated from this data indicated that there are high levels of genetic diversity in the populations of *O. longistaminata* studied and this diversity is higher within than between populations. Regional partitioning also revealed that most of this diversity is found between populations within regions than among regions. Among all the populations sampled, five populations (EG-4 and all Amhara populations) were identified as priorities for conservation strategies and thus, national collection and conservation strategies need to consider population and regional differences.

Key words or phrases: genetic variability, microsatellites, Molecular markers, *Oryza longistaminata*, populations and simple sequence repeats.

1. Introduction

Oryza longistaminata is a rhizomatous plant species, taxonomically placed in the genus *Oryza*, under family Poaceae (Wang *et al*., 1992). This highly diverse and allogamous species is broadly distributed throughout Africa (Sacks *et al*., 2003) including Ethiopia, particularly in Gambela, Gojam and Illubabor regions (Phillips, 1995). It is used for various purposes such as food, forage, erosion control and live fences around farmlands and homesteads (Vaughan and Sitch, 1991).

The plant *O. longistaminata* is well adapted to shallow waters in pans, woodlands and wooded grasslands, pools, in stagnant or running swamps, flood plains, edges of rivers and dams, on riverbanks, amongst cultivated rice or along canals in cultivated rice fields (Chev and Roher, 1914).

As reported by Matsuo *et al*. (1997) the perennial growth habit, drought tolerance and dual regeneration ability, both through seeds and rhizomes, enabled *O. longistaminata* to inhabit regions extending from northern Uganda to South Africa or within latitudes $3^0$38'N to $26^0$17'S and longitudes $15^0$34'E to $48^0$58'E with an altitudinal range from 400m a.s.l. to 1800 m a.s.l.

Even if this plant has the same AA genome, it provide novel genes and it hybridizes easily with cultivated rice (Lu *et al.*, 2003). Only little research has so far been done on it (Vaughan and Morishima, 2003). This gap of information is due to absence of good taxonomic key characters, habitat change throughout its lifespan (Vaughan *et al*., 2008) and availability of the species in Africa (Sacks *et al*., 2003).

O. longistaminata has different biological names assigned by different taxonomists as; *O. barthii* sensu, *O. dewildemanii*, *O. madagascariensis*, *O. perennis* Moench ssp. *Barthii* and *O. perennis* Moench ssp. *madagascarensis* (Kiambi *et al*., 2000).

Besides, local consideration of *O. longistaminata* as noxious weed growing under diverse environmental conditions in tropical Asia, Africa, Australia and Central and South America (Khush, 1997), climatic change, overgrazing and human population pressure are also implicated in reduction of the genetic diversity of the species (Kiambi *et al*., 2000).

Therefore, a study such as the present one would contribute to assessing the genetic status and creating potential for effective genetic resources conservation of *O. longistaminata* in Ethiopia, particularly in Gambela and Gojam areas based on microsatellite molecular markers.

2. Literature Review

2.1 Phylogeny and Biogeography of *Oryza*

The genus *Oryza* includes two cultivated species and more than 20 wild species grouped into four clearly defined species complexes called *O. granulata*, *O. ridleyi*, *O. sativa* and *O. officinalis* (Ren *et al*., 2003). However, it is believed that less attention is given to assessing its origin, divergence and diversity which made this genus to remain largely unclear (Vaughan and Morishima, 2003).

2.1.1 Phylogeny of *Oryza*

Despite few isozyme based studies that were used to estimate the divergence time of *Oryza* and its related genera (Second, 1991), the proposed divergence event were not consistent (Vaughan and Morishima, 2003). Thus, at recent times, sequences of low-copy nuclear genes and chloroplast DNA fragments are widely applied and clearly considered middle Miocene (14–15 MYA) as the time of *Oryza* to be separated from *Leersia* and started distribution from Asia to Africa (Liang *et al*., 2010).

The evolutionary relationships among the eight diploid A-genome group species have long been controversial and still not well studied at the sequence level (Vaughan *et al*., 2008). But the evidence and many works revised by Sharma (2003) showed that, divergence of these lineages is the most recent within the genus and effected at two phylogenetic dates.

Roughly 2–3 MYA is suggested as the earlier time of divergent between ancestor of the perennial African AA genome species, *O. longistaminata* and Asian AA genome *Oryza* (Zhu and Ge, 2005) and 0.64 MYA as the latter date for the divergence of AA *Oryza* gene pool between Asian and African AA genome species (Ma and Bennetzen, 2004).

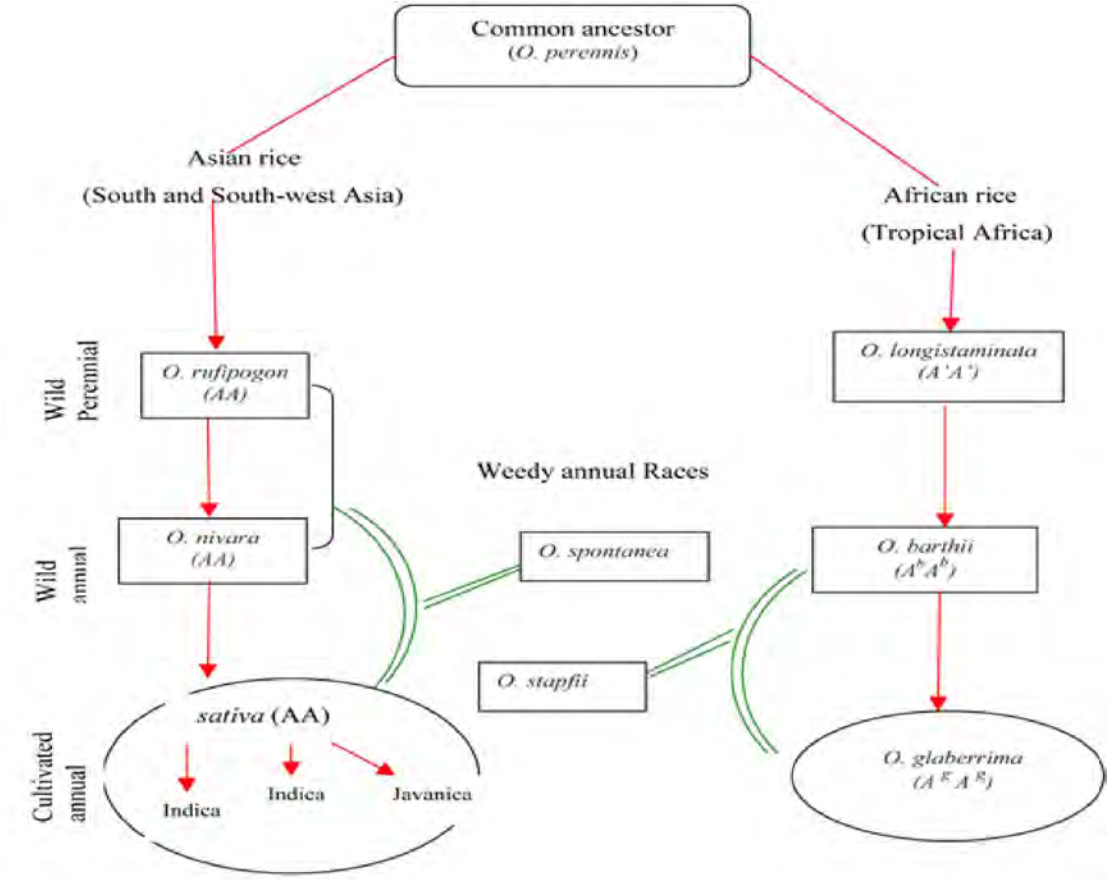

Figure 1. Schematic representation of the evolutionary pathways of Asian and African cultivated rice .Arrows with solid line indicates direct descent, arms with double lines indicate introgressive hybridization (Source: Randhawa *et al*., 2006).

2.1.2 Biogeography of *Oryza*

The tropical Asia, Africa, Australia and Central and South America are geographical regions for the abundance of *Oryza* members (Khush, 1997). Majority of the current biogeographcal studies of *Oryza* species proposed Asia as center of origin (Liang *et al*., 2010). Despite their center of origin is still debatable (Vaughan *et al*., 2008), the assumption on multiple introduction events proposed the distribution of *Oryza* from Asia to Africa and Australia (Vaughan *et al*., 2005). This prediction is also in agreement with the speculation stating Southeast Asia and New Guinea as distribution

sites for the two 'primitive' *Oryza* complexes and the pan-tropical distribution for the two recent species complexes (Vaughan *et al*., 2008).

Table 1 Species of *Oryza*, chromosome numbers, genome group/symbol and geographic distribution (Source: Randhawa *et al*., 2006).

Complex/Species	Chr. no.	Genome group	Distribution
O.schlechteri Pilger	48	Unknown	Papuua New Guinea
O.brachyantha Chev.et Roehr	24	FF	Africa
O.ridleyi complex			
O.longiglumis Jansen	48	HHJJ	Papuua New Guinea
O.ridleyil Hook	48	HHJJ	South East Asia
O.meyeriana complex			
O.granulata Nees et Arn ex watt	24	GG	South and South East Asia
O.meyeriana (Zoll.et Mor .ex Steud.) Baill	24	GG	South East Asia

O.sativa complex			
O.glaberrima (*cultigen*) Steud.	24	AA	West Africa
O.barthii A.Chev.	24	AA	Africa
O.longistaminata Chev.et Roehr.	24	AA	Africa
O.sativa(cultigen) L.	24	AA	World wide
O.nivara	24	AA	Tropical Asia
O.rufipogon sensu Lacto.	24	AA	Tropical Asia
O.meridionalis Ng	24	AA	Tropical Australia
O.glumaepatula Steud.	24	AA	South America

2.2 Taxonomic position of *Oryza longistaminata*

Oryza longistaminata is a monocotyledonous plant belonging to the genus *Oryza*, under the grass family Poaceae (Gramineae) (Wang *et al*., 1992) of the subfamily Oryzoideae within the rice tribe Oryzeae (Vaughan, 1994).

Many taxonomic literature used spikelet and awn length, rhizomatous nature and habitat requirement to categorize and list the wild *Oryza* members (Takeota, 1962). But now a days, environmental influence on such characters made these taxonomic keys to be out of use (Liang *et al*., 2010).

For this reason, almost all current ecological, genetic and morphological studies demonstrate that the whole *Oryza* members with the exception of the three *Oryza* species called *O. brachyantha*, *O. coarctata* and *O. schlechteri* fall into four clearly defined species complexes as *O. sativa*, *O. officinalis* Wall ex Watt., *O. ridleyi* Hook. and *O. granulate* complexes (Vaughan *et al*., 2008).

Among these complexes, the presence of the two cultivated rice species along with their easily intermating AA genome wild relatives including *Oryza longistaminata* Chev. et Roehr lead many workers to study diversity among and within species of the *O. sativa* complex through different mechanisms (Jiang-Hong *et al*., 2005).

For instance, the biotic long-distance dispersion reflected differentiation between the Austral-Asian species and the African species (Vaughan *et al*., 2008). Studies from a wide geographic areas along with molecular data reviewed by Doi *et al*., (2000) further separate Austral-Asian species as the Australian strains and the Asian rice (Jiang-Hong *et al*., 2005).

Furthermore, complete sequencing of the rice genome and establishment of BAC libraries for most genomes of the genus (Ammiraju *et al*., 2006) resulted from high concern of researchers on the *O. sativa* complex to simplify the task of taxonomy and phylogeny of AA genome groups and even the genus *Oryza* in general (Vaughan, 1989).

2.3 Botanical description of *Oryza longistaminata*

2.3.1 Morphology

Oryza longistaminata is a robust perennial herb with an erect, soft and spongy stem and has about 2.5m height (Cope, 2008). This grass has an alternate, simple and entire leaf with a spongy, pale green to brownish, smooth and glabrous leaf sheath and an acuminate, bright to dark green and glabrous blade (Chev and Roher, 1914).

Most of grass families have a short and membranous structure in between the leaf sheath and the blade called ligule. Among other species of the

family, *O. longistaminata* is easily distinguished from other wild *Oryza* species (Kiambi *et al*., 2000). As the ligule is very long (ranges between 15 and 45 mm) and pointed projection along with the long, branched, creeping and extensive rhizomes at the lower nodes (Cope, 2008).

According to Chev and Roher (1914), *O*. *longistaminata* is partly self-incompatible and allogamous grass and it fills its elliptic pollen inside the narrow asymmetrically elliptical-oblong, 7–15 mm long, pale green to brownish, deciduous and obliquely articulated to the pedicel called spikelet.

The inflorescence is 16–40 cm × 2.5–8 cm in size, dense, slight drooping with an obliquely ascending manner and with an almost erect branches and the fruit as an oblong caryopsis (grain), 7.5–8.5 mm long, glabrous, pale brown and glossy (Cope, 2008).

2.3.2 Reproductive Biology

O. longistaminata populations are larger than other *Oryza* species. This is probably because it is perennial and has two simultaneous modes of reproduction (Cope, 2008). The continuous reproduction of *O. longistaminata* both in time and space is facilitated through its rhizomes and seeds (Kiambi *et al*., 2005).

Its unique strictly outcrossing feature or high rate of self-incompatibility impede the population size from being large enough (Jones *et al*., 1996). Even if the reason for this high level of sterility in *O. longistaminata* is not obvious, the perennial growth habit of the species which causes relatively longer phonological cycle and variability in maturity of seeds along with scarcity of water level in its habitats are suggested as the two driving forces (Kiambi *et al*., 2005).

As a result, *O*. *longistaminata* is restricted to have natural reproduction mainly by its rhizomes and sets only few seeds (Chev and Roher, 1914). A small amount of the seeds could be the result of hybridization with the cultivated species, *O. glaberrima* and *O. sativa* (Vaughan, 1989).

2.4 Uses of *O. longistaminata*

Like most of the wild *Oryza* species, *Oryza longistaminata* is used as food during time of scarcity (famine), for constructing homes and good grazing for cattle in countries like Ethiopia and Sudan (Vaughan and Sitch, 1991). This species can also accomplish useful gene transfer to *O. sativa* through conventional hybridization, selection and backcrossing procedures (Brar and Khush, 1997).

This African wild rice species is genetically diverse, strongly perennial and vigorously rhizomatous (Tao *et al*., 2000). Thus, the plant is the best donor of perennial traits to develop a permanent ground cover and perennial upland rice (Sacks *et al*., 2003). Disease resistance (Tao *et al*., 2000), high pollen production, long stigmas and drought tolerance traits were transferred to the cultivated rice (Khush *et al*., 1990).

Currently, the tedious task for an effective transfer of such important and novel traits from *O. longistaminata* to the cultivated rice is made easier (Brar and Khush, 1997). Mostly through *in situ* conservation, BAC library, fluorescence *in-situ* hybridization and *in vitro* embryo rescue are widely used (Sacks *et al*., 2003).

2.5 Marker systems and their applications in genetic diversity analysis

2.5.1. Morphological markers

Morphological marker systems are the earliest, simple and inexpensive genetic markers which lie on phenotypic appearance (Vithanage *et al.*, 1995). However, morphological traits that exhibit continuous variation between individuals in a population often obscure the evaluation of genetic diversity. Moreover, pleiotropism and a multifactorial basis to morphological traits further conceal the characterization of the plant populations (Rahman *et al*., 2010). Thus, the very plastic nature of morphological traits in turn needs extra time and resources for evaluation in the field and greenhouse (Bretting and Widrlechner, 1995).

2.5.2. Biochemical markers

Biochemical markers are protein polymorphism based markers (Weising *et al*., 2005). The two general forms of protein data are gathered from the simultaneous electrophoretic methods of isozymes and allozymes (Murphy *et.al*., 1996). The primary observed evidence in studies of electrophoretic variation of isozymes in natural populations is coloured bands in a slab of starch gel (Gottlieb, 1977). Distinguishing a particular enzyme among hundreds that may be present in a crude tissue extract is possible due to the combination of electrophoresis and staining specificity (Chamberlain, 1998). The biochemical consequence of the change on amino acids in the polypeptides or enzymes can be distinguished if these changes affect their electrophoretic migration (Gottlieb, 1977). Since the amino acid sequence of a polypeptide is collinear to the nucleotide sequence of its coding structural gene locus, isozyme analysis using electrophoresis is an analysis of a gene (Crawford, 1990).

However, these advantages of isozyme analysis are offset by certain limitations. The first limitation is that enzymes subjected to electrophoresis are a tiny and probably non-representative sample of the total array of proteins present in an organism and electrophoretic differences are only one kind of difference (Weeden and Wendel, 1990). Underestimation of the actual amount of genetic difference between taxa is another problem of isozyme analysis (Gottlieb, 1977).

A single mobility class on a gel may sometimes contain more than one enzyme (Murphy *et.al*., 1996). Electrophorectic evidence also lacks information on the number of amino acid differences or mutational steps that cause differences in enzyme mobilities (Gottlieb, 1977). Secondary isozymes (conformational isozymes), that differ in secondary or tertiary structure can also result in alterations on mobility and may mislead the number of loci encoding an enzyme system (Murphy *et.al*., 1996).

2.5.3. Molecular markers

According to Bretting and Widrlechner (1995), molecular markers are "land marks" which can be identified on the genome and offer the best possible means of identifying individuals from biological samples. Such markers are also reliable in genetic mapping, marker assisted selection (MAS), population genetics and molecular systematics (Weising *et al*., 2005). Genetic markers were described before the discovery of proteins and DNA (Rahman *et al*., 2010). DNA marker systems, which were introduced to genetic analysis in the 1980s, have many advantages over the traditional morphological and protein markers (Gupta, 1999).

Although the vast majority of DNA-based studies used microsatellites or Simple Sequence Repeats (SSRs) (Abdlkrim *et al*., 2009), other genetic markers like; RFLPs (Restriction Fragment Length Polymorphisms), Amplified Fragment Length Polymorphisms (AFLPs), Single Nucleotide Polymorphisms (SNPs), mitochondrial (mtDNA) and chloroplast (cpDNA) have been developed and applied to a range of plants (Rahman *et al*., 2010).

RFLPs that resulted from point or chromosomal mutations in restriction enzyme recognition sites are the first Southern Blot based DNA marker systems (Tanksley *et al*., 1989). The RFLP technique employs molecular hybridization of cDNA or genomic DNA probes with genomic DNA fragmented by restriction enzymes (Botstein *et al*., 1980). While the RFLP technique utilizes mostly low copy number probes, the fingerprinting technique uses highly repetitive minisatellite DNAs as probes (Kim *et al*., 2000). These southern blot based marker systems also have other limitations. Therefore, most of the genetic analysis became based on the PCR revolutionized technique (Mullis *et al*., 1986).

AFLP is another PCR-based molecular marker system which obviates the need for template DNA sequence information (Vos *et al*., 1995). Genomic DNA digested with two different restriction enzymes is ligated with specific adaptor sequences which then amplified, denatured, separated and stained with silver (Rahman *et al*., 2010). Partial digestion of the template DNAs by

contamination or organ specific methylation produces spurious AFLP bands (Donini *et al*., 1997). Such marker types are usually coded in a binary way by the presence or absence of fragments (Bonin *et al*., 2007).

Mitochondrial and chloroplast DNA have low mutation rate and evolve at comparatively slower rates (Takazaki and Nei, 2008). But these markers are often mediated entirely by one sex (Scribner *et al*., 2001). Thus, they can reveal genetic differences in the long-term accumulated effects of sex-bias in animals and pollen or seed dispersion in plants. Actually, the involvement of multiple markers having different modes of inheritance is very crucial for better estimation of genetic diversity (Smouse, 2010).

The other form of molecular markers is the system that utilizes SNPs (single nucleotide polymorphisms) (Choi *et al*., 2007). Compared to the gel-based molecular marker systems, SNP detection and analysis can be carried out with through-put analysis, by hybridization with allele-specific oligonucleotides (ASO) (Hashimoto *et al*., 2004), primer extension (Sybanen, 1999), oligonucleotide ligation assay (OLA) (Iannone *et al*., 2000) and invasive cleavage (Lyamichev, *et al*., 1999). In *indica* type rice, the comparison of genomic sequences from Yu *et al*., (2002) demonstrated one SNP in every 170 bp. The frequency of the SNPs in all plant species can range from approximately one per 30 bp to one per 500 bp (Rahman *et al*., 2010). These new generation marker systems are powerful tools in genotyping large number of loci and reliable genetic variability investigations (Smouse, 2010). Despite, they are only amenable to use in those species for which extensive nucleotide sequence information is available (Rafalski *et al*., 2002).

The most technically efficient and cost-effective PCR-based markers are the Microsatellites or Simple Sequence Repeats (SSRs) (Abdlkrim *et al*., 2009).
In the 1960s, simple repeats that are scattered throughout eukaryotic genomes were identified in density gradient centrifugations of randomly sheared genomic DNAs by the technique of a ‘satellite peak’ (Moustacchi *et al*., 1966). Isolation and sequencing of these satellite DNAs revealed repeat

motifs of variable length from just a single base to thousands of bases (Pardue *et al*., 1970). Subsequently, satellites of 10–30 bp repeat motifs, termed minisatellites, were isolated in mammals (Jeffreys *et al*., 1985). Finally, satellites with 1-6 bp repeat motifs, called microsatellites, were isolated (Blakenhol *et al*., 2009). Furthermore, Weber and May (1989) demonstrated that SSR polymorphisms (SSRPs) could be easily detected in animals by PCR through two flanking primers (Weissenbach *et al*., 1992). In plants, the presence of SSRs was first demonstrated by the hybridization of oligonucleotide probes of poly (G-T) and poly (A-G) on the phage libraries of tropical tree genomes (Condit and Hubbell, 1991). A search of published DNA sequences reveals that SSRs are also highly abundant in diverse plant genomes (Wang *et al*., 1994). SSRs have, therefore, become the preferred molecular marker system for analysis in plant genetics and ecology (Park *et al*., 2009).

Even if SSRPs are varied in length by different number of repeat units, it is not certainly known (Richard *et al*., 2008). However, several studies have demonstrated that such tandem repeats in genomes arose by unequal crossing-over between repeat units during meiosis (Jeffreys *et al*., 1998) and DNA replication slippage (Strand *et al*., 1998). DNA replication slippage can also occur during *in vitro* amplification of the SSRs which often obscure non-parental SSR bands in genetically segregating populations (Ellegren, 2004).

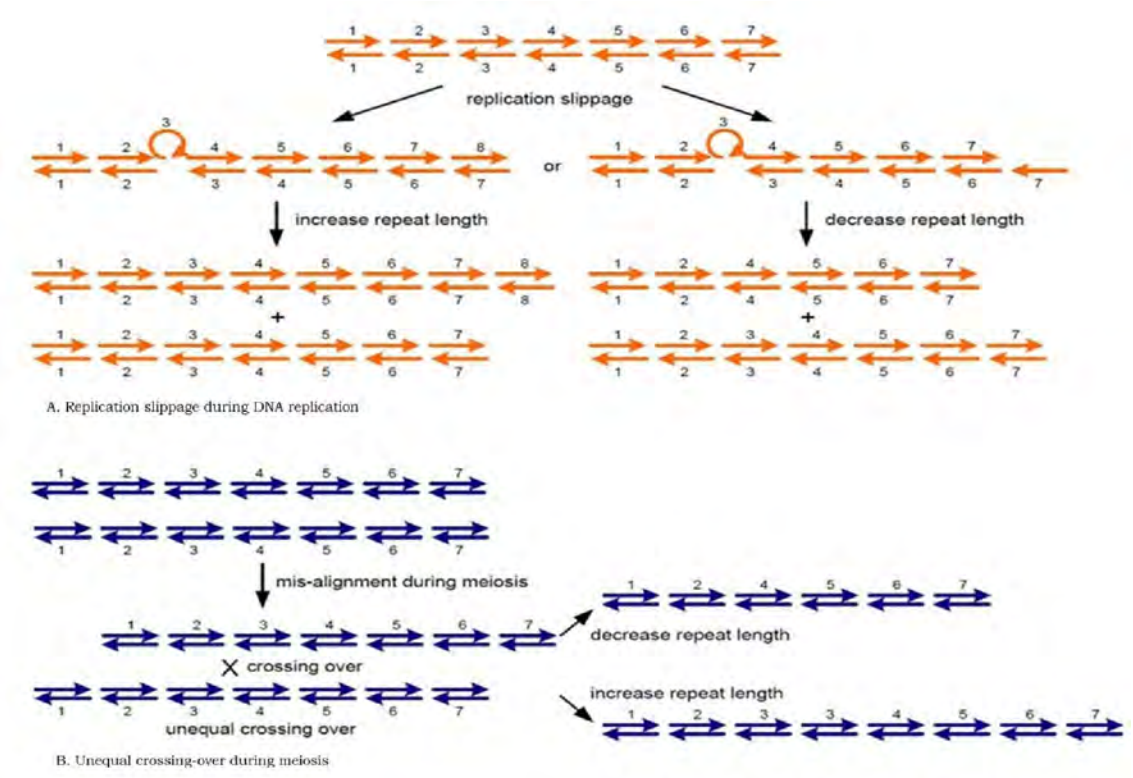

Figure 2. The mechanisms of unequal crossing-over and replication slippage for producing SSRPs (Source: Park *et al*., 2009).

Commonly, different species harbour different SSR motifs preferentially and SSRs with longer repeats are more polymorphic than those of shorter ones (Ellegren, 2004).

In rice, microsatellites are abundant and well distributed throughout the genome (Akagi *et al.,* 1996). Previous studies on rice have contributed to the development of several hundred microsatellite markers and a genetic map consisting of 320 SSRs (Temnykh *et al.,* 2001). For characterization and documentation, this technique has been recently used in crop species including rice, wheat, maize, barley, rapeseed, soybean, potato and other crops (Rahman *et al.,* 2006). In Bangladesh, nine soybean cultivars were identified by microsatellite markers, which have provided identity and might work as protection (Islam *et al.,* 2007). During the last five decades 133 maize cultivars grown in France were also characterized using microsatellite fingerprinting (Clerc *et al.,* 2005).

Generally, SSR markers have many advantages over the other marker systems. Some of the advantages are their high reproducibility, simple experimental procedures and requirement of only a small amount of template DNA (Park *et al*., 2009).

The second advantage of the SSR marker system is that it does not require restriction with enzymes, the template DNA is not found to be ultra pure (Manen *et al*., 2003) and has high polymorphic genetic information contents (Boder *et al*., 2003). In a comparative study of the utility of RFLP, RAPD, AFLP and SSR marker systems for germplasm analysis, SSRs showed the highest expected heterozygosity (Park *et al*., 2009). The third advantage has to do with the co-dominant nature of SSR polymorphisms (Scott *et al*., 2000). The co-dominant nature of SSRPs is also suitable for genetic analysis in segregating F2 populations or parentage analysis in hybrids (Slavov *et al*., 2005).

The fourth advantage is their abundance and distribution in genomes (Toth *et al*., 2000). In search of SSRs longer than 12 bp in a 57.8 Mb of the publicly available rice (*Oryza sativa* L.) sequence, Temnykh *et al*., (2001) showed that many kinds of SSRs are present every 16 kb.

A fifth advantage of the SSR marker system is that SSRs are preferentially associated with nonrepetitive DNA (Andersen and Liberstedt, 2004). Genomic sites of SSR markers, derived from genomic libraries, fall into either the transcribed region (genic SSRs) or the non-transcribed region (genomic SSRs). The SSRs, derived from ESTs or cDNAs, are mostly genic SSRs, which have the potential for application in such areas as gene function characterization (Ronning *et al*., 2003), association analysis for gene tagging and QTL analysis (Zeng *et al*., 2009).

Although analysis with SSR markers has many merits compared to other marker systems (Abdlkrim *et al*., 2009), there are a few inherent problems associated with it. Other than being costy in-start (Park *et al*., 2009), SSR

polymorphisms are sometimes derived from slippage during polymerase chain reactions (Hauge and Litt, 1993). Slippage during PCR produces 'stutter bands' that differ in size from the main product by multiples of the length of repeat unit (Smulders et al., 1997). The stuttering produces many ladder bands in polyacrylamide gel separation and leads to quasi-scoring, if there are no prominent bands among the ladders (Park *et al*., 2009). The other limitation of SSR analysis lies on homoplasy which can occur if two bands are similar in size but not identical in sequence (Shinde *et al*., 2003). The relative advantages and disadvantages of these molecular and biochemical markers are summarized below.

Table 2. Comparison of the most commonly used molecular and biochemical marker systems in plant genome analysis (Park *et al*., 2009).

Feature	Isozyme	RFLP	AFLP	SSR	SNP
Abundance	Low	Medium	Very high	High	Very high
Types of Polymorphism	Amino acid change	Single base change	Single base change	Repeat length variation	Single base change
DNA quality	-	High	High	Medium	Medium
DNA sequence Information	-	Not required	Not	Required	Required
Level of Polymorphism	Low	Medium	High	High	High
Inheritance	Codominance	Codominance	Dominance	Codominance	Codominance
Reproducibility	Medium	High	Medium	High	High
Technical complexity	Medium	High	Medium	Low	Medium
Developmental cost	Medium	High	Low	High instart	High
Species Transferability	High	Medium	High	Medium	Low

3. Objectives

3.1 General objective

- To investigate genetic diversity of *O. longistaminata* populations in Amhara and Gambella regions using SSR markers.

3.2 Specific objectives

- To assess the genetic diversity within and between populations of *O. longistaminata*.
- To compare the genetic variability of Amhara and Gambela *O. longistaminata* populations.
- To determine the eco-geographic distribution of the wild rice (*Oryza longistaminata*) and generate information for sustainable use and conservation.

4. Materials and Methods

4.1 Plant materials

A total of 320 samples representing eight populations were collected from Gambella and Amhara regions. Four of the populations were collected from Amhara and four were from Gambella regions. Here, sampling was done randomly with an approximate distance of 10 m from each other.

Localities of the populations used in the present study, the site maps and climatic diagram based on data of the recent 17 years, obtained from the Ethiopian Meteorological Authority, are given in Table 3 and Fig. 3, 4a and Fig. 4b respectively.

Table 3. Localities of *O. longistaminata* used in the present study.

Population ID	No of individual	Region	Locality	Latitude	Longitude	Altitude (m.a.s.l.)
EG1	40	Gambella	Kera area	8°11'53"	34°16'36"	435
EG2	40	Gambella	Komboni	8°09'35"	34°23'25"	438
EG3	40	Gambella	Abobo	7°48'34"	34°20'31"	460
EG4	40	Gambella	Agricultural institute	7°43'43"	34°26'10"	455
BD1	40	Amhara	Fogera plains	12°01'24"	37°43'12"	1799
BD2	40	Amhara	Tikur	11°41'58"	37°18'37"	1727
BD3	40	Amhara	Zege	11°31'04"	37°28'23"	1801
BD4	40	Amhara	Tana Hotel	11°36'24"	37°23'45"	1800

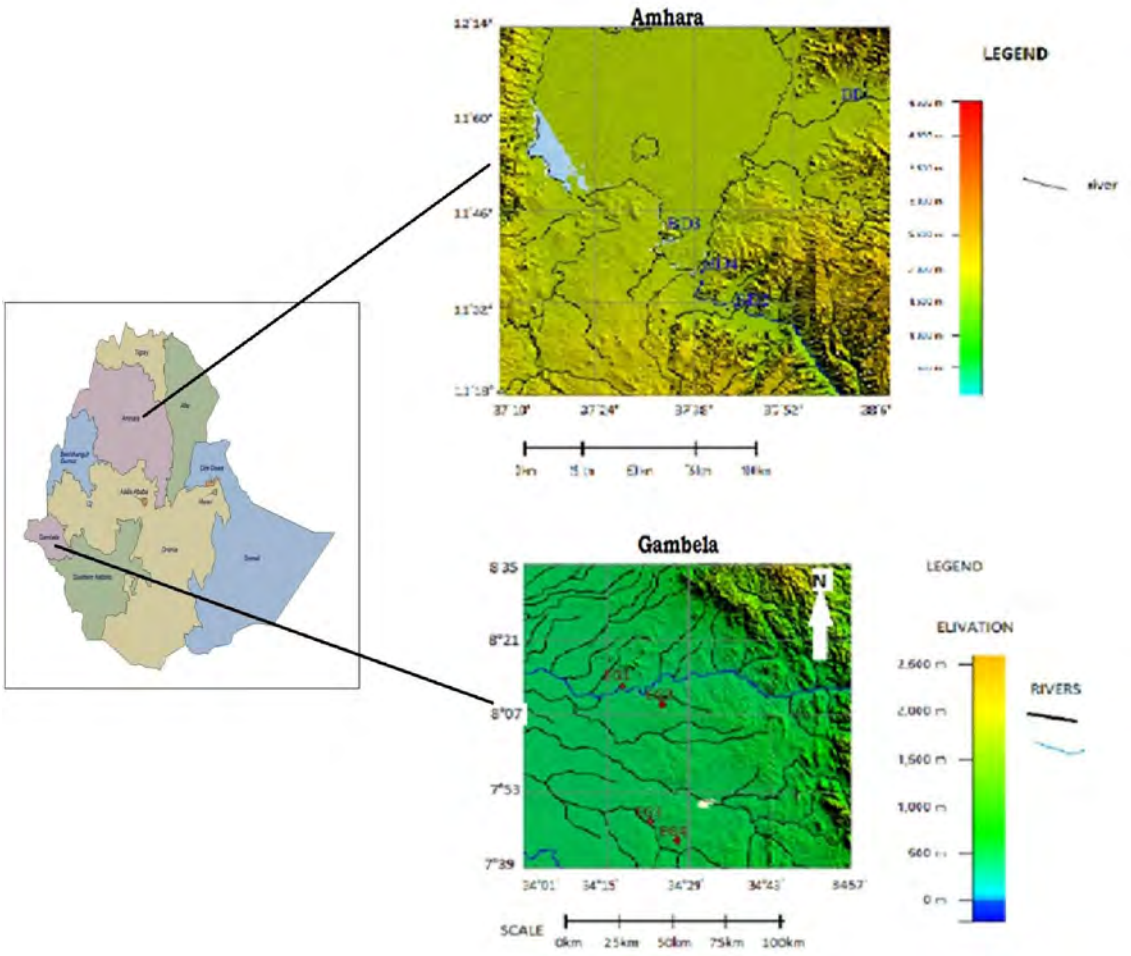

Fig. 3. Site maps for where the samples were collected in Gambela and Amhara.

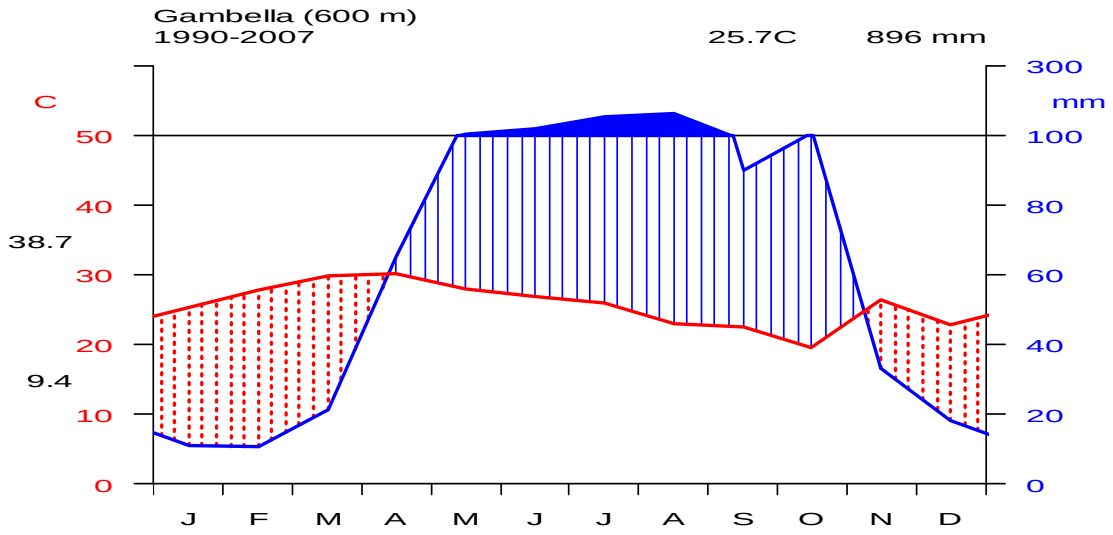

Fig. 4a. Climograph of Gambela at 600m a.s.l. with 25.7 ºc mean temperature and 896 mm annual rainfall.

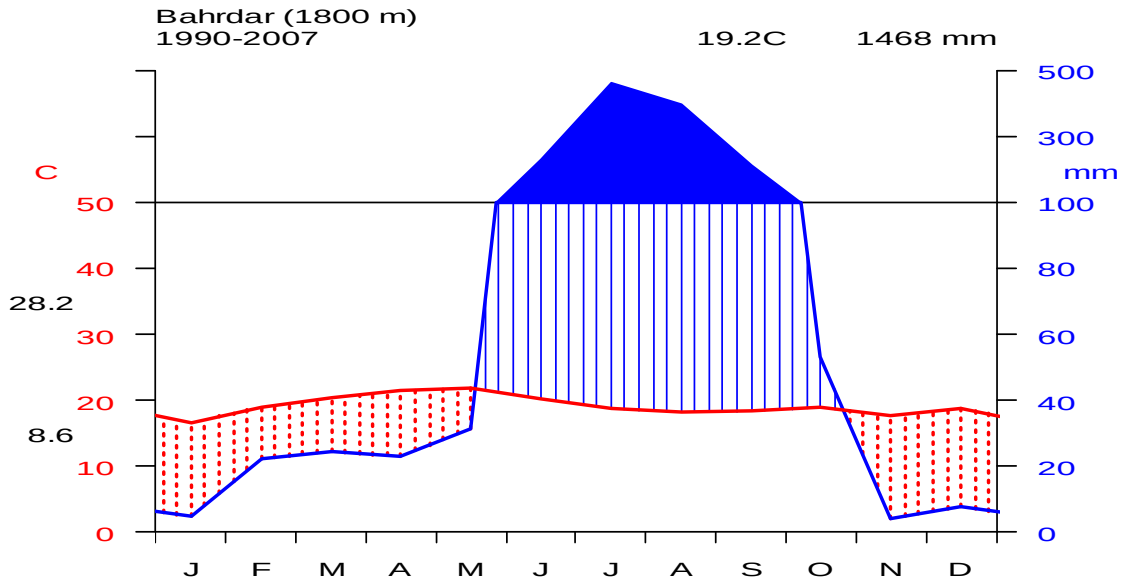

Fig. 4b. Climograph of Bahir Dar at 1800m a.s.l. having 19.2 ºc mean temperature and 1468 mm annual rainfall.

4.2 DNA Extraction

DNA was extracted from the leaf part of the sampled plants. Pair of pliers along with a clean white garment were used for pressing and soaking the leaf sample extract on to both sides of the FTA paper (Lin *et al*., 2000; Drescher and Graner, 2002).

4.3 FTA processing for PCR

Processing of FTA for PCR was performed as described by the manufacturer, with skipping the second sterilized water rinsing step. A single FTA disc measuring 1.2 mm was punched from the pigmented FTA disc and incubated for 15 min in a 50 µl FTA wash solution filled PCR plate well. These were followed by 5 min rinsing with a 100µl of double distilled water and dried in an oven for 15 min at 56 ^{o}C.

4.4 PCR Amplification

The PCR master mix consisting of 310 µl Sterile distilled water, 500 µl 1x Reddymix, 40 µl of 0.5 µM each forward and reverse primers with 10 µl of 25 mM $MgCl_2$ was prepared through 30 s vortexing and one minute centrifugation at 1,200rpm in a 1.5 eppendorf tube. Nine µl of the master mix was dispensed to each of the PCR tubes having a single disc. Next, these PCR tubes were set up in PTC-100 or Programmable Thermal Controller from MJ RESEARCH, INC. for DNA amplification.

The five selected pair of primers described by McCouch *et al*. (2002) and Song *et al*. (2003) with their No of cycles, periodical and thermal profile used in microsatellite DNA profiling analysis are stated in Table 4 and Table 5 respectively.

Table 4. Details of Primers used in microsatellite DNA profiling analysis as described by McCouch *et al*. (2002) and Song *et al*. (2003).

FWD/REV Primers	Sequence	Band Size (bp)
RM44-FWD	5'- ACG GGC AAT CCG AAC AAC C-3'	90 – 112
RM44-REV	5'-TCG GGA AAA CCT ACC CTA CC-3'	90 – 112
RM180-FWD	5'-CTA CAT CGG CTT AGG TGT AGC AAC ACG-3'	90 – 130
RM180-REV	5'-ACT TGC TCT ACT TGT GGT GAG GGA CTG-3'	90 –130
RM234-FWD	5'-ACA GTA TCC AAG GCC CTG G-3'	99 – 166
RM234-REV	5'-CAC GTG AGA CAA AGA CGG AG-3'	99 – 166
RM263-FWD	5'-CCC AGG CTA GCT CAT GAA CC-3'	124 170
RM263-REV	5'-GCT ACG TTT GAG CTA CCA CG-3'	124–170
RM280-FWD	5'-ACA CGA TCC ACT TTG CGC-3'	148–172
RM280-REV	5'-TGT GTT CTT GAG CAG CCA GG-3'	148–172

Table 5. Periodical and Thermal profile for the PCR reaction of the SSRs.

Marker type	No of cycles	PCR profile	Phases of the PCR reaction					
			Initial Denaturation	Denaturation	Annealing	Extention	Final extention	Cooling
RM 44	34	Temp (°c)	95	95	56	72	72	15
		Time (min)	7	0.5	20	0.66	10	α
RM-180	34	Temp (°c)	95	95	55	72	72	15
		Time (min)	7	0.5	0.5	1	0.5	α
RM-234	34	Temp(°c)	95	95	55	72	15	15
		Time (min)	7	0.5	0.5	1	12	α
RM-263	29	Temp (°c)	95	95	57	72	72	15
		Time (min)	5	0.5	0.66	0.66	10	α
RM-280	39	Temp (°c)	95	95	57	72	72	15
		Time (min)	7	0.5	0.5	0.5	40	α

4.5 Electrophoresis

PCR products were detected by gel electrophoresis through loading 1.8 μl of each genomic DNA samples in a 2.5% agarose gel (50 ml, 1.5x TBE, 1.25g agarose and 1.5 μl of 1mM EtBr). After electrophoresis at constant voltage of 100V for 45 minutes, the gel was exposed to the UV transilluminator.

4.6 Fragment Analysis

Preparation for fragment analysis was started from synthesizing a master mix consisting of Hi-Di formamide and GeneScan LIZ-500 size standard in a 75:1 ratio. 9 µl of the master mix were dispensed in to 1.8 µl of the PCR product. To minimize the cost for analysis, two or three differently dye labelled PCR products were multiplexed.

These dye-labelled PCR products of the reaction plate were electrophoresed in ABI Prism 3730 capillary sequencer and their entire process; base lining, peak detection, peak sizing and size calling were completely automated by GeneScan Analysis Software version 3.1.

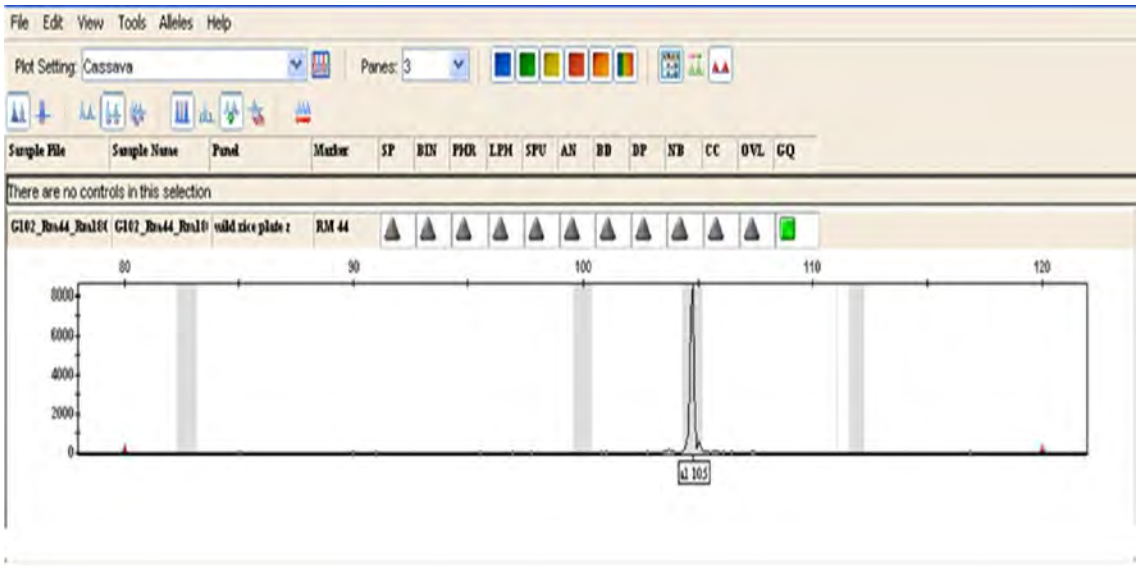

Fig. 5 Electropherogram showing the analysis of multiplexed FAM, TET and HEX labelled microsatellite PCR products.

Here, peak represents fluorescence Intensity (vertical axis) with size estimate in nucleotides (horizontal axis) and Microsatellite alleles which are identified above the peaks.

4.7 Data analysis

Different parameters were used to assess the genetic variation in all populations. Measures of variation include the polymorphic information containt (PIC), allelic frequencies in all polymorphic simple sequence repeat (SSR) loci, the actual and effective number of alleles, observed heterozygosity (Ho) and expected heterozygosity (He). These analyses were performed using the software program Fstat version 2.9.3 (Goudet, 2001). The F- statistics parameters (Fst, Fis and Fit) estimated from this genetic data were calculated using (Weir and Cockerham, 1984).

Genetic distances (D) were computed using Nei's (1972) and the dendrogram was drawn based on the genetic distance. To further examine the patterns of variation among individual samples, a principal coordinated analysis (PCO) was performed based on Jaccard's coefficient (Jaccard, 1908). The calculation of Jaccard's coefficient was made with PAST soft ware version 1.18 (Hammer *et al.,* 2001).

Definitions for the parameters used to measures genetic variation:

(i) Na = The mean observed number of allels.

(ii) Ne = Effective number of allels.

(iii) PIC = Polymorphic Information Containt.

(iv) Ho = proportion of observed heterozygotes.

(v) He = proportion of heterozygotes under Hardy-Weinberg equilibrium.

(vi) Fst = the degree of genetic differentiation between populations.

(vii)Fis=deviation of genotypic frequencies from the Hardy-Weinberg equilibrium.

(viii) Fit = Total inbreeding coefficient.

(ix) D = gene differences between populations.

5. Results

The five pairs of primers amplified a total of 64 alleles in 320 plants from the 8 populations of *O. longistaminata* (Table 5). In this study, all of the loci showed polymorphism over the populations. Although there was a considerable variation among the 8 populations, the number of alleles per locus varied from 11 to 16 with an average of 14 alleles (Table 6). Allele frequencies obtained from all the five microsatellite loci, indicated RM-44 as the least polymorphic and RM-263 the most polymorphic and there was no population having all allels of all the loci assayed.

Table 6. Mean allele frequencies of all wild populations examined in this study.

Locus	Allel	Population							
		EG1	EG2	EG3	EG4	BD1	BD2	BD3	BD4
RM-44	p: 90	0.029	0.028	0	0.032	0.04	0.036	0.15	0.167
	p: 92	0	0	0.081	0.032	0	0.036	0	0
	p: 102	0.103	0.25	0.216	0.161	0.12	0.071	0.25	0.111
	p: 103	0	0	0	0	0	0.036	0	0
	p: 104	0.735	0.611	0.486	0.581	0.68	0.536	0.35	0.556
	p: 105	0.029	0	0.054	0.065	0	0.071	0.05	0.056
	p: 106	0.044	0.083	0	0	0.12	0.036	0.05	0.056
	p: 108	0.029	0	0	0.032	0	0.036	0	0
	p: 110	0.029	0.028	0.081	0.065	0.04	0.071	0.05	0
	p: 112	0	0	0.081	0.032	0	0.071	0.1	0.056
RM-180	p: 90	0	0	0	0	0	0.028	0	0
	p: 92	0	0	0	0	0	0.056	0	0
	p: 93	0.026	0.014	0.079	0.057	0.054	0	0.031	0
	p: 96	0	0	0	0	0	0	0.031	0.036
	p: 99	0.079	0	0.276	0.243	0.473	0.181	0.203	0.196
	p: 102	0.224	0.375	0.408	0.286	0.324	0.597	0.406	0.554
	p: 105	0.632	0.5	0.145	0.286	0.081	0.056	0	0
	p: 108	0	0.056	0.053	0.1	0.054	0	0.078	0.036
	p: 111	0.013	0.028	0	0	0.014	0	0.188	0.036
	p: 114	0	0	0	0	0	0.028	0.063	0.107
	p: 117	0	0	0	0	0	0.056	0	0.036
	p: 120	0	0.028	0.013	0	0	0	0	0
	p: 122	0	0	0.026	0.029	0	0	0	0
	p: 130	0.026	0	0	0	0	0	0	0

Table 6 – Continued

RM-234	p: 99	0.026	0	0	0	0	0	0	0

	p: 108	0	0	0.074	0	0	0	0	0
	p: 118	0	0	0	0	0	0.045	0	0
	p: 120	0.026	0.028	0.037	0.235	0	0.023	0.083	0.094
	p: 122	0.053	0.236	0.241	0.324	0.15	0.182	0.083	0.141
	p: 124	0.263	0.514	0.241	0	0.15	0.136	0.083	0.125
	p: 125	0	0	0	0	0	0.045	0	0
	p: 126	0.079	0.014	0.111	0.265	0.35	0.227	0.021	0.047
	p: 128	0.355	0.194	0.241	0	0	0.023	0	0.016
	p: 130	0.171	0.014	0.019	0.118	0.35	0.318	0.729	0.578
	p: 148	0.013	0	0	0	0	0	0	0
	p: 158	0	0	0.037	0	0	0	0	0
	p: 162	0.013	0	0	0	0	0	0	0
	p: 166	0	0	0	0.059	0	0	0	0
RM-263	p: 124	0.04	0.04	0	0	0	0	0	0
	p: 148	0.06	0.08	0.145	0	0	0.038	0.133	0.088
	p: 152	0.18	0.02	0.065	0.069	0	0.115	0.067	0.118
	p: 154	0	0	0.032	0	0	0	0	0
	p: 156	0	0	0.065	0	0	0	0	0
	p: 158	0.08	0.12	0.21	0.017	0	0.192	0.067	0.059
	p: 160	0.04	0.06	0.065	0.052	0	0.038	0.033	0.088
	p: 162	0.08	0.02	0.097	0.034	0	0.077	0.033	0.147
	p: 164	0.1	0.08	0.097	0.052	0.091	0.077	0.133	0.118
	p: 166	0.26	0.38	0.129	0.448	0.364	0.231	0.467	0.265
	p: 168	0.16	0.2	0.097	0.293	0.545	0.231	0.067	0.118
	p: 170	0	0	0	0.034	0	0	0	0
RM-280	p: 148	0	0.033	0	0	0	0.02	0	0
	p: 150	0.2	0	0	0	0.029	0.12	0	0.12
	p: 152	0.05	0	0.034	0	0.029	0	0.043	0.08
	p: 154	0.05	0.067	0.034	0	0	0.04	0	0.08
	p: 156	0	0	0.034	0.017	0.029	0	0.043	0
	p: 158	0	0	0.172	0.017	0.043	0.04	0.087	0
	p: 160	0.025	0	0.121	0.067	0.186	0.06	0.109	0.1
	p: 162	0.075	0.1	0.069	0.033	0.114	0.06	0.13	0.06
	p: 164	0	0.033	0.155	0.133	0.029	0	0.043	0
	p: 165	0	0	0.017	0	0	0	0	0
	p: 166	0.125	0.2	0.017	0.133	0.3	0.3	0.174	0.22
	p: 168	0.475	0.567	0.345	0.4	0.214	0.36	0.37	0.34
	p: 170	0	0	0	0.1	0.029	0	0	0
	p: 172	0	0	0	0.1	0	0	0	0

5.1 Levels of genetic diversity

The allelic frequency in all *O. longistaminata* populations of all loci showed the presence of genetic diversity due to the presence and absence of allels, the presence of rare allel and most frequent allel (Table 6). The observed number of allels (Na), the effective number of allels (Ne), mean observed and expected proportion of heterozygous loci and Wright's fixation index (Fis) were calculated for each population (Table 7).

At the population level, the observed number of allels (Na) ranged from 5.6 (BD1) to 8.2 (highest in EG3 and BD2) and the effective number of allels (Ne) ranged from 4.57 (EG4) to 6.57 (BD4). expected heterozygosity (He) was in a range of between 0.66 for EG2 and 0.8 for EG3. When all populations were pooled as one, there were 7.13 observed number of allels (Na), 5.93 effective number of allels (Ne) with the total mean observed (Ho) and expected heterozygosity (He) level of 0.163 and 0.73 respectively.

In this study, the observed hetrozygosity (Ho) is actually less than expected heterozygosity (He). The degree of differentiation (Fst) range of the five loci was between 0.008 for RM-44 and 0.153 of RM-234.

Table 7. Comparisons of genetic variations of the eight *O. longistaminata* populations by different parameters.

Population ID	Population parameters				
	Na	Ne	Ho	He	Fis
EG1	7.6	5.65	0.245	0.68	0.69
EG2	6.4	6.55	0.318	0.66	0.52
EG3	8.2	5.67	0.216	0.8	0.7
EG4	7.2	4.57	0.092	0.75	0.88
BD1	5.6	6.49	0.127	0.69	0.85
BD2	8.2	5.98	0.105	0.76	0.84
BD3	7	5.92	0.083	0.73	0.83
BD4	6.8	6.57	0.117	0.74	0.83
Mean	7.13	5.93	0.163	0.73	0.77

N.B. Na= Actual No of allels, Ne= Effective no of allels, Ho= Observed heterozygosity, He= expected heterozygosity, Fis= Wrights fixation index.

From the five microsatellite loci assayed in this study, significant deviation from Hardy-Weinberg expectations was revealed for all of them as they show the Wright's fixation index (Fis) range from 0.52 in EG2 to 0.88 in EG4 which indicated significant heterozygote deficiency. The mean degree of differentiation (Fst) and the Wright's fixation index (Fis) values were 0.062 and 0.75 correspondingly (Table 8).

Locus	PIC	Ho	He	Fis	Fit	Fst

RM-44	0.6194	0.0260	0.68	0.965	0.966	0.008
RM-180	0.7226	0.1187	0.661	0.826	0.845	0.113
RM-234	0.7817	0.3532	0.638	0.5	0.576	0.153
RM263	0.8057	0.2754	0.89	0.676	0.683	0.019
RM280	0.7901	0.1626	0.823	0.783	0.787	0.018
Mean	0.7439	0.1872	0.7384	0.75	0.7714	0.0622

Table 8. Summary of P, observed and expected heterozygotes and F-statstics for all loci pooling all populations together.

N.B. PIC= Polymorphic Information Containet, Ho= Observed heterozygosity, He= expected heterozygosity, Fis= Wrights fixation index, Fit= Total inbreeding coefficient and Fst =Genetic differentiation.

5.2 Genetic distance and genetic identity

Genetic identity and genetic distance were determined for each pair of the eight populations using the methods of Nei (1972) (Table 9). The distance between populations ranged from 0.0828 to 0.5219. The highest distance (0.5219) was between populations of EG1 and BD3 and the smallest genetic distance (0.0828) was observed between the two Amhara populations (BD2 and BD4).

A dendrogram constructed based on Nei's (1972) unbiased genetic distance (Figure 6) comprises two clusters where the three Gambella populations (EG1, EG2 and EG3) were confined in the first cluster and the second cluster was partitioned in to two clusters consisting of the three Amhara populations (BD2, BD3 and BD4) in one cluster and BD1 and EG4 in the other.

Table 9. Nei's Unbiased measures of genetic distance determined for each pair of the eight *O. longistaminata* populations.

	BD1	BD2	BD3	BD4	EG1	EG2	EG3	EG4
BD1	0.0000							
BD2	0.1839	0.0000						
BD3	0.3248	0.2161	0.0000					
BD4	0.2439	0.0828	0.0886	0.0000				
EG1	0.4369	0.3289	0.5219	0.3623	0.0000			
EG2	0.4255	0.2914	0.4928	0.3842	0.1115	0.0000		
EG3	0.3604	0.2354	0.4525	0.3229	0.2644	0.2086	0.0000	
EG4	0.1837	0.2232	0.3523	0.3006	0.3062	0.2889	0.2765	0.0000

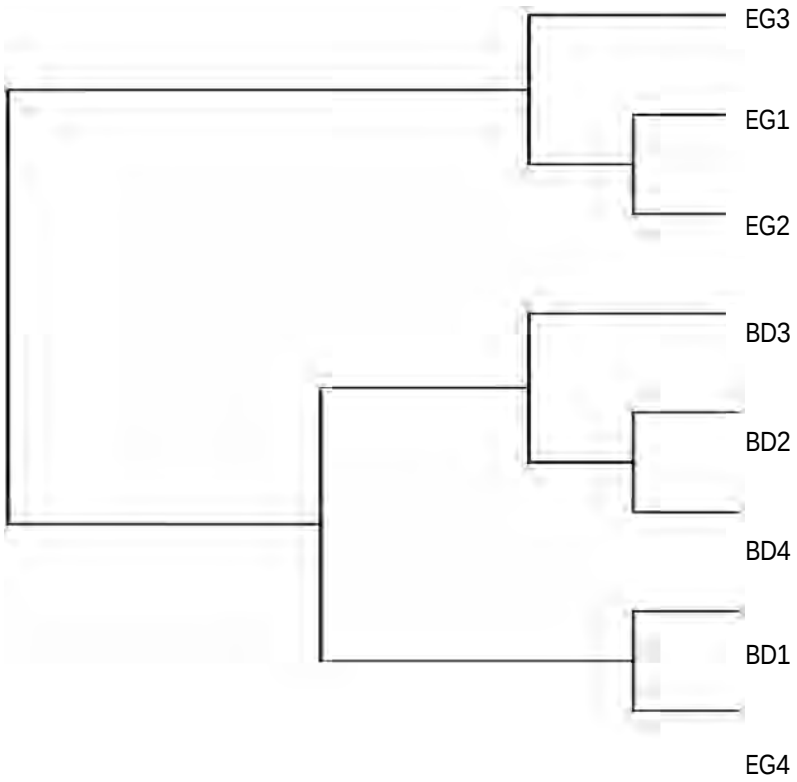

Figure 6. Dendrogram based on Nei's (1978) unbiased genetic identity and genetic distance.

5.3. PCO Analysis

All the data obtained using 5 SSR primers were used in PCO analysis using Jaccard's coefficients of similarity. Individuals of the eight populations using three coordinates observed to form separation through the contribution of the 3 axes PC1 (8.7%), PC2 (7.8%) and PC3 (7.1%) (Figure 7).

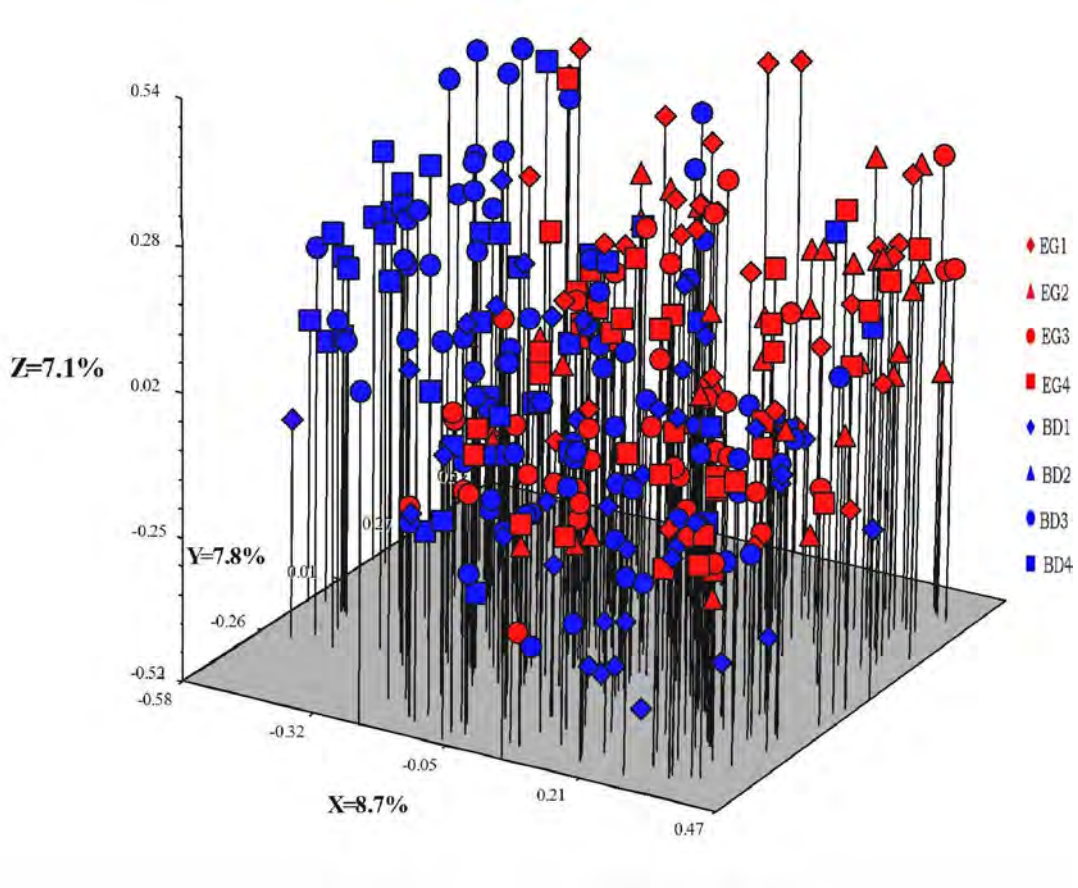

Figure 7. Three dimensional representations of individuals for the 8 *O. longistaminata* populations based on Jaccard's similarity coefficients.

5.4 Intergeographical relationship of the populations

As shown in Table 10, the Gambela and Amhara populations were pooled separately and populations collected from Gambela had Na=11.2, Ne=10.27, Ho=0.672 and He=0.75. While, the Amhara populations had shown Na=9.8, Ne=9.34, Ho=0.223 and He= 0.74 with the population differentiation (Fst) among the regions of 0.064 and Fis values of 0.826 and 0.696 for the Amhara and the Gambela populations respectively.

Table 10. Comparisons of genetic variation of the Gambella and Amhara populations of *O. longistaminata*.

Regions	Na	Ne	Ho	He	Fis
Gambella	11.2	10.27	0.672	0.75	0.696
Amhara	9.8	9.34	0.223	0.74	0.826

6. Discussion

6.1. Genetic diversity of *O. longistaminata* populations

The microsatellite marker system in this study revealed the presence of very large amount of genetic variability (He= 0.73, Na = 7.13, Ne =5.93 and average P = 0.7513) within the eight *O. longistaminata* populations. This is not surprising since high levels of variations are associated with abundance of the species in the different ecological regions (high altitude, low altitude and in different ecological adaptations) for long period of time (Kiambi *et al*., 2008).

The AFLP based genetic diversity analysis of *O. longistaminata* populations reported by Kiambi *et al*., (2005) and the *O. glumaepatula* populations of Marines *et al*., (2007) showed low level of diversity than this work. The high level of differentiation in all *O. longistaminata* individuals in comparison to the result of Kiambi *et al*., (2005) could be attributed to the microsatellite markers used in the study (Beaumont *et al*., 2001). These hypervariable molecular markers have a high possibility to be segregated in to a high number of alleles. Thus, high degree of genetic diversity and higher number of detected alleles are effected (Marines *et al*., 2007). These large number of different alleles can also be generated over a relatively short time or period (Freeland *et al*., 2000). Actually, the variable periods of genetic isolation throughout the evolutionary history of the populations studied (Marines *et al*., 2007) might be the reason for the variability in the numbers of unique alleles as observed in this study. But the tendency of the non-adaptive feature and high mutation rate of microsatellite loci are the major agents to enhance genetic variability among individuals (Freeland *et al*., 2000).

Even if Marines *et al*., (2007) used microsatellite markers, the report showed a relatively low average genetic diversity indices (Na = 3.091, PIC = 0.773 and Ne = 0.393) for the Brazilian *O. glumaepatula* populations that share the same AA genome with *O. longistaminata* (Kiambi *et al*., 2005). In fact, this may be expected as *O. glumaepatula* is an inbreeding species. Whereas,

this study was done on the allogamous and partially self incompatible *O. longistaminata* species which develops reproductive organs especially, anther and stigma, in a manner favouring cross pollination (Chu *et al*., 1969). The larger sample size in the study which is about twice of the size for the Brazilian *O. glumaepatula* populations also allows detection of large number of alleles for better estimation on numbers of shared alleles among populations (Freeland *et al*., 2000).

Based on the allelic frequency, the presence and absence of allels, the presence of rare allels and most frequent allels, genetic variability is high in all populations of all the loci. In this regard, genetic variability is higher in RM-234 than all the other markers used whereas half of the populations (EG1, EG3, EG4 and BD2) had their own unique allels.

Using AFLP, Kiambi *et al*., (2005), found within population diversity of (He = 0.1755) and between population diversity of (Fst = 0.1688) of *O. longistaminata* populations. Similarly, Buso *et al*., (1998) reported within population diversity (He= 0.081) and higher levels of between population diversity (Fst= 0.29) in *O. glumaepatula*. Here, both results coincide with the low levels of between population diversity (Fst) in comparison to the highest average diversity (Fst = 0.200) over all the populations of Bangladesh rice landraces (Parsons *et al*., 1999). The total mean genetic diversity among populations and the within genetic difference (Fst = 0.13485 and He = 0.716) in the *O. longistaminata* populations of this study, is quite different from any of the three. This relatively high levels of mean diversity over all the loci and the higher within (0.716) than between (0.1348) population diversity may be attributed to the wild nature and mating system of the plant (Girma, 2010), type of marker used and sampling technique (Maki and Horie, 1999).

6.2. Distinction of *O. longistaminata* populations

The pairwise Nei's genetic distance between populations indicated the least genetic distance (0.0828) from the two Amhara populations and the highest genetic distance (0.5219) between populations of EG1 from Gambela and BD3 from Amhara populations (Table 9). Similarly, the UPGMA dendrogram (Figure 5) showed that all populations from the same regions except BD1 and EG4 form clusters based on proximity. This deviation of BD1 and EG4 from the exact correlation between the genetic and geographic distances might be because of long distance gene flow (Girma *et al*., 2010) and the presence of high number of private alleles in four of the eight populations (Thingsgaard, 2001). For instance, the five unique alleles in BD2 and EG3 populations were high as compared to the within populations mean number of alleles ranged from 5.8 to 8.2 (Table 7) or because of possible gene flow between wild rice populations in the surrounding areas (Girma *et al*., 2010).

6.3 Comparison of the Gambela and Amhara populations

On the basis of geographic locations, there are not much pronounced differences in diversity. But the Gambela region showed slightly higher level of genetic diversity than the Amhara populations (Table 10). According to Freeland *et al*., (2000), it is difficult to suggest the responsible causes for the current patterns of diversity. As, both historical and ongoing events can regulate the responsible factors. Some of such determinant factors are; mating systems, rate of mutation, migration and dispersal mechanisms, biotic and abiotic selection intensities (Nevo *et al*., 1979; Frankel, 1984; Parsons *et al*., 1999; Kark *et al*., 1999).

The intergeographical relationship which involved all the *O. longistaminata* populations assessed in this study showed populations from the Gambela region to be slightly more diverse than the Amhara region. This is because of the highest number of allels and a high number of rare alleles recorded (Kiambi *et al*., 2005). The close association of *O. longistaminata* genetic diversity with the appropriate rainfall intensity as observed in this study

(Nevo *et al*., 1979) and the optimum vegetation type also favour the region (Kiambi *et al*., 2008).

Due to the altitudinal range of the region, populations of the Amhara region also indicated large genetic variability difference (Kiambi *et al*., 2000). Furthermore, environmental heterogeneity as reported by IUCN (1990) and Kiambi *et al*., (2001) and scarcity of water level which promotes self sterility (Kiambi *et al*., 2005) play a great role to the high genetic diversity in the region.

When all the sampled populations were pooled together, there was large difference between the observed heterozygosity (Ho =0.22462) and gene diversity (He= 0.7116). Departure from Hardy-Weinberg in all proportions is also high since, there is significant Wright's fixation index (Fis = 0.743). But as indicated in Table 10, the Amhara populations having a mean observed heterozygosity (Ho =0.223) and gene diversity (He=0.74) and a highly significant Wright's fixation index (Fis = 0.82) took the large contribution for the departure from Hardy-Weinberg proportions. Basically, male sterility, drift or Wahlund's effect and selection against heterozygotes might be the possible causes for the low level of heterozygosity detected (Ho<He) in the Amhara region (Murphy *et al*., 1996). But this study was based on microsatellite markers which are not adaptive and hence the genetic drift effect is random throughout the evolutionary time (Pemberton *et al*., 2000). But magnificent level of overgrazing, agricultural expansion and population pressure inhibit the plant from attaining maturity. The influence of such factors also let the plant to have prevailing intrapopulation inbreeding mating system through its rhizomes (Kiambi *et al*., 2005).

Continuity on the current climatic change, human threats along with the geographical mobility resulted from erosion-deposition in the natural community of *O. longistaminata* populations might cause loss of genetic

diversity and rapid cycles of population extinction and colonization. Therefore, ecologists, plant breeders and other responsible bodies should give a high concern to the gene pool of this species as it is a novel gene source of perennial, disease resistant, high pollen producing, long stigma and drought tolerant traits for the cultivated rice (*Oryza sativa*).

7. Conclusions

The genetic parameters estimated from the microsatellites data indicated that there are high levels of genetic diversity in the eight populations of *O.*

longistaminata studied and this diversity is higher within than between populations. The high level of genetic diversity in the sampled populations indicates that the species has a great potential for breeding programs which aim on agronomically important gene transfer to the cultivated rice (*Oryza sativa*) and the African rice *Oryza glaberima* or creating a hybrid that can ensure food security to Ethiopia and even to the world.

The total mean genetic diversity for all the loci in all populations (Fst = 0.13485) shows that there is low among populations genetic diversity compared to the average expected level of the self compatible plant species. But, still this result is indicative of the presence of moderate differentiation among populations.

The different natural and artificial threats on the Amhara populations are the possible causes for the very high level of heterozygosity deficient in the region.

Besides, the study has clearly demonstrated the usefulness of microsatellites in studying diversity in wild rice populations. This is due to their resolution power in discriminating individuals, estimating diversity between and within populations and generating significant understanding on the population structure and genetic diversity of *O. longistaminata* in Ethiopia.

8. Recommendation

This study does not include populations from Illubabor. Hence, more survey and sample collection has to be carried out in this area so as to have a complete record from the whole populations of *O. longistaminata* in Ethiopia. Moreover, due attention should be given to Amhara populations since they showed heterozygous deficiency and for EG3 (from Gambela) as it is showed highest level of genetic diversity.

Generally, genetic diversity of wild rice populations including *O. longistaminata* in east and southern Africa is being eroded by the expansion of agricultural activities and overgrazing. Intensive efforts are therefore, required to collect and conserve this diversity for future use in rice improvement programmes.

Molecular phylogenetics and cytological analysis of *O. longistaminata* in Ethiopia should be carried out to realize its distribution and evolution.

Further study on the occurrence and distribution of other *Oryza* members like *O. barthi* in Ethiopia is also very crucial.

Evolutionary relationships of the two or more *Oryza* species coupled with crossability study, should be conducted to better understand the evolution of the related species and make use of *O. barthi* and other wilds for cultivated rice improvement.

9. References

Abdelkrim, J., Robertson, B.C., Stanton, J.A. and Gemmell, N.J. (2009). Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* 46: 185–192.

Akagi, H. Yokozeki, Y., Inagaki, A. and Fujimura, T. (1996). Microsatellite DNA Markers for rice chromosomes. *Theor. Appl. Genet.* 93: 1071–1077.

Ammiraju, J.S., Luo, M. and Goicoechea, J.L. (2006). The *Oryza* bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large insert BAC libraries that represent the 10 genome types of the genus *Oryza*. *Genome Res*. 16: 140–147.

Andersen, J.R. and Liberstedt, T. (2003). Functional markers in plants. *Trends Plant Sci.* 8: 554–560.

Balkenhol, N., Gugerli, F. and Cushman, S.A. (2009). Identifying future research needs in landscape genetics: where to from here? *Landscape Ecol*. 24: 455–463.

Beaumont, M., Barratt, E. M., Gottelli, D., Kitchener, A. C., Daniels, M. J., Pritchard, J. K. and Bruford, M. W. (2001). Genetic diversity and introgression in the Scottish wildcat. *Mol. Ecol.* 10: 319–336.

Boder, P., Deak, T., Bacso, R., Velich, I., Bisztray, G.D., Fascar, G. and Gyulai, P. (2006). Morphological and genetic investigation of medieval grape seeds. *Acta Horticulture (ISHS)* 713–718.

Bonin, A., Ehrich, D. and Manel, S. (2007). Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Mol*. *Ecol*. 16: 3737–3758.

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

Brar, D.S. and Khush, G.S. (1997). Alien introgression in rice. *Plant Mol*. *Biol*. 35: 35–47.

Bretting, P.K. and Widrlechner, M.P. (1995). Genetic markers and plant genetic resource management. *Plant Breed. Rev.* 13: 11–86.

Buso, G.C., Rangel, P.H. and Ferreira, M.E. (1998). Analysis of genetic variability of South American wild rice populations (*Oryza glumaepatula*) with Isozymes and RAPDs markers. *Mol. Ecol.* 7: 107-117.

Chamberlain, J.R. (1998). Isozyme variation in *Calliandra calothyrsus* (leguminosae): its implications for species delimitations and conservation. *Amer. J. Bot*. 85: 1-23.

Chev, A. and Roehr, C. (1914). *Oryza longistaminata*. *Sci*. *Terre*. 159: 561.

Choi, I.Y., Hyten, D.L., Matukumalli, L.K., Song, Q., Chaky, J.M., Quigley, C.V., Chase, K., Lark, K.G., Reiter, R.S. and Yoon, M.S. (2007). A soybean transcript map: Gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genet*. 176: 685–696.

Chu, Y.E., Morishima, H. and Oka, H.I. (1969). Partial self-incompatibility found in *Oryza perrennis*. subsp. Barthii. Jpn. *J. Genet.* 44: 225-229.

Clerc, V.L., F Bazante, F., Baril, C., Guiard, G. and Zhang, D. (2005). Assessing temporal changes in genetic diversity of maize varieties using microsatellite markers. *Theor Appl Genet.* 110: 294–302.

Condit, R. and Hubbell, S.P. (1991). Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome*. 34: 66–71.

Cope, T.A. (2008). In: Flora of Somalia, Vol 4

Crawford, D.J. (1990). *Plant Molecular Systematics*. John Wiley and Sons,

D.M., Moritz, C.and Mable. B.K., eds). Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts U.S.A.

Doi, K., Nonomura, M.N., Yoshimura, A., Iwata, N. and Vaughan, D.A. (2000). RFLP relationships of A-genome species in the genus *Oryza*. *J. Faculty Agric. Kyushu Univ*. 45: 83–98.

Donini, P., Koebner, R.D., Elias, M.L. and Bougourd, S.M. (1997). AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. *Genome* 40: 521–526.

Drescher, A. and Graner, A. (2002). PCR-genotyping of barley seedlings using DNA samples from tissue prints. *Plant Breeding*. 121: 228-231.

Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. *Nat. Rev. Genet.* 5: 435–445.

Frankel, O.H. (1984). Genetic perspectives of germplasm conservation. In: Genetic Manipulation Impact on Man and Society Aber W, Llimensee K, Peacock WJ, Starlinger P (eds). Cambridge University Press, Cambridge, UK.

Freeland, J.R., Romualdi, C. and Okamura, B. (2000). Gene flow and genetic diversity: a comparison of freshwater bryozoan populations in Europe and North America. *Heredity* 85: 498–508.

Girma, G., Tesfaye, K. and Bekele, E. (2010). Inter Simple Sequence Repeat (ISSR) analysis of wild and cultivated rice species from Ethiopia. *African Journal of Biotechnology,* 9: 5048-5059.

Gottlieb, L.D. (1977). Electrophoretic evidence and plant systematics. *Ann. Missouri Bot. Gard.* 64:161-180.

Goudet, J. (2001). Fstat, a program to estimate and test gene diversities and fixation indices (version 2.9.3). http://www. Unil. Ch/izea/softwares Fstat.

Hammer, O., Harper, D.A.T. and Ryan, P.D. (2001).PAST : Paleontological statistics software package for education and data analysis. *Palaeontologia electronic* 4:9, http//palaeo-electronica.org/2001_1/past/issue1-01.htm.

Hashimoto, K., Hashimoto, M., Mishiro, S. and Oota, Y. (2004). Method of detecting nucleic acid relating to disease. *EU Patent App.* 13: 56-72.

Hauge, X. Y. and Litt, M. (1993). A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum. Mol. Genet.* 2: 411–415.

Iannone, M.A., Taylor, J.D., Chen, J., Li, M.S., Rivers, P., Slentz-Kesler, K.A. and Weiner, M.P. (2000). Multiplexed single nucleotide polymorphism genotyping by oligonucleotide ligation and flow cytometry. *Cytometry* 39:131–140.

Islam, M.N., Molla, M.R. and Rahman, L. (2007). Microsatellite allele size profiling to identify and distinguish soybean cultivars in Bangladesh. *Progress. Agric.* 18: 9-17.

IUCN (1990). Biodiversity in sub-Saharan Africa and its Islands: conservation, management and sustainable use. IUCN – The World Conservation Union, Gland, Switzerland.

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

Jeffreys, A.J., Murray, J. and Neumann, R. (1998). High-resolution mapping of crossovers in human sperm defines a minisatellite-associated recombination hotspot. *Mol. Cell* 2: 267–273.

Jian-Hong, X., Nori, K., Masahiro, A., Hisako, O. and Eiichi, O. (2005). Identification and characterization of Australian wild rice

strains of *Oryza meridionalis* and *Oryza rufipogon* by SINE insertion polymorphism. *Genes Genet. Syst*. 80: 129-134.

Jones, M.P., Dingkuhn, M., Johnson, D.E. and Fagade, S.O. (1996). Inter-specific hybridization: progress and prospects. In: Proceedings of workshop on: Africa/Asia joint research on interspecific hybridization between the African and Asian rice species (*O. glaberrima* and *O. sativa*) (WARDA ed.). West Africa Rice Develop. Ass. Bouake, Cote D'Ivoire, pp. 67-75.

Kark, S., Alkon, P.U., Safriel, U.N. and Randi, E. (1999). Conservation priorities for chukar partridge in Israel based on genetic diversity across an ecological gradient. *Conser*. *Biol*. 13: 542-552.

Khush, G.S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant*. *Mol*. *Biol*. 35: 25–34.

Khush, G.S., Balacangco, E. and Ogawa, T. (1990). A new gene for resistance to bacterial blight from *Oryza longistaminata*. *Rice Genet*. *Newsl*. 7: 121-122.

Kiambi, D.K, Newbury, H.J., Ford-Lloyd, B.V. and Dawson, I. (2005). Contrasting genetic diversity among *Oryza longistaminata* (A. Chev et Roehr) populations from different geographic origins using AFLP. *African Journal of Biotech*. 4: 308-317.

Kiambi, D.K., Ford-Lloyd, B.V., Mwangi, M., Guarino, L., Newbury, H.J., Maxted, N. and Jarvis, A. (2000). The ecogeographic distribution of five wild rice (*Oryza* spp.) in East and Southern Africa. PhD. Thesis, University of Birmingham. U.K.

Kiambi, D.K., Ford-Lloyd, B.V., Newbury, H.J. and Maxted, N. (2001). Mapping the geographic distribution of molecular diversity of AFLPs using *O. longistaminata* as case study. PhD. Thesis, University of Birmingham.

Kiambi, D.K., Newbury, H.J., Maxted, N. and Ford-Lloyd, B.V. (2008). Molecular genetic variation in the African wild rice *Oryza longistaminata* A. Chev. et Roehr. and its association with environmental variables. *Afri. J. of Biotech*. 7: 1446-1460.

Kim, N.S., Park, N.I., Kim, S.H., Kim, S.T., Han, S.S. and Kang, K.Y. (2000). Isolation of TC/AG repeat microsatellite sequences for fingerprinting rice blast fungus and their possible horizontal transfer to plant species. *Mol. Cells* 10: 127–134.

Liang, T., Xin-hui, Z., Gaston, A., Christina, P., Gerard, S., Da-yong, Z and Song, G. (2010). Phylogeny and biogeography of the rice tribe (Oryzeae): Evidence from combined analysis of 20 chloroplast fragments. *Mol. Phyl*. and *Evol*. 54: 266–277.

Lin, J.J., Fleming, R., Kuo, J., Mathews, B.F. and Saunders, J.A. (2000). Detection of plant genes using a rapid, nonorganic DNA purification method. *Biotech*. 28: 346-350.

Liu, J. (2005). Power marker version 3.2.5.

Lu, B.R., Song, Z.P. and Chen, J.K. (2003). Can transgenic rice cause ecological risks through transgene escape? *Progress in Nat. Sci*. 13: 17-24.

Lyamichev, V., Mast, A.L., Hall, J.G., Prudent, J.R., Kaiser, M.W., Takova, T., Kwiatkowski, R.W., Sander, T.J., de Arruda, M. and Arco, D.A. (1999). Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotl.* 17: 292–296.

Ma, J. and Bennetzen, J. L. (2004). Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci*. 101: 12404–12410.

Maki, M. and Horie, S. (1999). Random amplified polymorphic DNA (RAPD) markers reveal less genetic variation in the endangered plant *Cerastium fischerianum* var. molle than in the widespread conspecific *C. fischerianum* var. *fischerianum* (*Caryophyllaceae*). *Mol. Ecol*. 8: 145-150.

Manen, J.F., Bouby, L., Dalnoki, O., Marinval, P., Turgay, M. and Schlumbaum, A. (2003). Microsatellites from archaeological *Vitis vinifera* seeds allow a tentative assignment of the geographical origin of ancient cultivars. *J. Archaeol. Sci.* 30: 721–729.

Marines, M.G. Karasawa, I., Roland, V., Silva, C.M., Maria, I.Z., Giancarlo, C.X., Oliveira, I. and Elizabeth, A. V. (2007). Genetic structure of Brazilian wild rice (*Oryza glumaepatula* Steud., Poaceae) populations analyzed using microsatellite markers. *Genet. and Mol. Biol.* 30: 400-410.

Matsuo, T., Futsuhara, Y., Kikuchi, M. and Yamaguchi, H. (1997). (Eds) Science of the rice plant. Volume three: Genetics. Food and Agricultural Policy Research Centre, Tokyo, Japan.

Molla, M.R., Islam, M.N. and Rahman, L. (2007). DNA fingerprinting of maize (*Zea mays* Linn.) cultivars of Bangladesh using SSR markers. *Bangladesh J. Crop Sci.* 18: 63-72.

Moustacchi, E. and Williamson, D.H. (1966). Physiological variations in satellite components of yeast DNA detected by density gradient centrifugation. *Biochem. Biophys. Res. Commun.* 23: 56–61.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986). Specific enzymatic amplification Of DNA *in vitro*: The polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51: 263–273.

Murphy, R.W., Sites, J.W., Buth, D.G. and Haufler, C.H. (1996). Proteins: isozyme electrophoresis. In: *Molecular Systematics*, pp. 51-121, (Hills, D.M., Moritz, C.and Mable. B.K., eds). Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts U.S.A.

Murphy, R.W., Sites, J.W., Buth, D.G. and Haufler, C.H. (1996). Proteins: isozyme electrophoresis. In: *Molecular Systematics*, pp. 51-121, (Hills,

Nei, M. (1972). Genetic distance between populations. *Am. Natur.* 106: 283-292.

Nevo, E., Zohary, D., Brown, A.H. and Haber, H. (1979). Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum* in Israel. *Evol*. 33: 815-833.

Pardue, M.L. and Gall, J.G. (1970). Chromosomal localization of mouse satellite DNA. *Sci*. 168: 1356–1358.

Park, Y., Lee, J.K. and Kim, N. (2009). Simple Sequence Repeat Polymorphisms (SSRPs) for Evaluation of Molecular Diversity and Germplasm Classification of Minor Crops. *Molecules*. 14: 4546-4569.

Parsons, B.J., Newbury, H.J., Jackson, M.T. and Ford-Lloyd, B.V. (1999). The genetic structure and conservation of aus, aman and boro rices from Bangladesh. *Gen*. *Res*. *and Crop Evol*. 46: 587-598.

Pemberton, J.M., Coltman, C., Coulson, T.N. and Slade, J. (2000). Using microsatellites to measure the fitness consequence of inbreeding and outbreeding. In: Microsatellites Evolution and Applications Goldstein D.B. and Schlotterer C. (eds). Oxford University Press, New York, pp 151-164.

Phillips, S. (1995). Poaceae (Gramineae). In: Flora of Ethiopia and Eritrea (Hedberg, I., Edwards, S., eds.). Volume7.Upsala, Sweden.

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* 5: 94–100.

Rahman, M.S., Sohag, M.H. and Rahman, L. (2010). Microsatellite based DNA fingerprinting of 28 local rice (*Oryza sativa* L.) varieties of Bangladesh. *J. Bangladesh Agril. Univ*. 8: 7–17.

Randhawa,G. J., Verma, D. D., Bhalla, S., Hota, M., Chalam, V. C. and Tyagi, V. (2006). *Document on Biology of Rrice (Oryza sativa L.) in India.* National Bureau of Plant Genetic Resources, New Delhi. 88pp.

Ren, F., Lu, B., Li, S., Huang, J. and Zhu, Y. (2003). A comparative study of genetic relationships among the AA-genome *Oryza* species using RAPD and SSR markers. *Theor. Appl. Genet.* 108:113–120.

Richard, G.F., Kerrest, A. and Dujon, B. (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. Rev.* 72: 686–727.

Ronning, C.M., Stegalkina, S.S., Ascenzi, R.A., Bougri, O., Hart, A.L., Utterbach, T.R., Vanaken, S.E., Riedmuller, S.B., White, J.A. and Cho, J. (2003). Comparative analyses of potato expressed sequence tag libraries. *Plant Physiol.* 131: 419–429.

Sacks, E.J., Roxasb, J.P. and Cruzb, M.T. (2003). Developing Perennial Upland Rice II Field Performance of S1 Families from an Intermediated *Oryza sativa*/O. *longistaminata* Population. *Cropsci*. 43: 129-134.

Scribner, K.T., Petersen, M.R. and Fields R.L. (2001). Sex-biased gene flow in spectacled eiders (Anatidae): inferences from molecular markers with contrasting modes of inheritance. *Evol*. 55: 2105–2115.

Second G. 1991. Molecular markers in rice systematics and the evaluation of genetic resources. In: *Rice Biotechnology in Agriculture and Forestry*.(Y. Bajaj, ed.). Springer-Verlag, Vol. 14, Berlin, Germany, pp 468-494.

Sharma, S. D. (2003). Species of the genus *Oryza* and their interrelationships. In: *Monograph on the Genus Oryza*.Nanda (J.S. Sharma, eds.). Science Publishers, Enfield, New Hampshire, pp 73–111.

Shinde, D., Lai, Y., Sun, F. and Arnheim, N. (2003). Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT) n and (A/T) n microsatellites. *Nucleic Acids Res.* 31: 974–980.

Slavov, G.T., Howe, G.T., Gyaourova, A.V., Birkes, D.S. and Adams, W.T. (2005). Estimating pollen flow using SSR markers and paternity exclusion: Accounting for mistyping. *Mol. Ecol.*, 14: 3109–3121.

Smouse, P.E. *(*2010). How many SNPs are enough? *Mol*. *Ecol*. 19: 1265–1266.

Smulders, M.M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P. and Vosman, B. (1997). Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theor. Appl. Genet.* 94: 264–272.

Song, Z.P., Xu, X., Wang, B., Chen, J.K. and Lu, B.R. (2003). Genetic diversity in the northernmost *Oryza rufipogon* populations estimated by SSR markers. *Theor Appl Genet*. 107: 1492-1499.

Strand, M., Prolla, T.A., Liskay, R.M. and Petes, T.D. (1998). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nat*. 365: 274–276.

Susan R. McCouch, S.R., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D. and Stein, L. (2002). Development and Mapping of 2240 New SSR Markers for Rice (*Oryza sativa* L.). *DNA Res*. 9: 257-279.

Sybanen, A.C. (1999). From gels to chips: "Minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum. Mutat.* 13: 1–10.

Takazaki, N. and Nei, M. (2008). Empirical tests of the reliability of phylogenetic trees constructed with microsatellite data. *Genet*. 178: 385–392.

Takeota, T. (1962). Taxonomic studies of *Oryza*. II. Several species complexes. *Bot*. *Mag*. *Tokyo* 75: 455-461.

Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989). RFLP mapping in plant breeding: New tools for an old science. *Nat. Biotech.* 7: 257–264.

Tao, D., Hu, F., Yang, Y., Xu, P., Li, J., Sacks, E., Nally, M. and Scripichitt, K. P. (2000). Rhizomatous individual was obtained from inerspecific BC1F1 progenies between *Oryza sativa* and *O. longistaminata*. In: *The abstract of International Rice Genetic Symposium*, International Rice Research Institute, Los Bnaos, Philippines, 22–27.

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

Thingsgaard, K. (2001). Population structure and genetic diversity of the amphiatlantic haploid peatmoss *Sphagnum affine* (Sphagnopsida). *Heredity* 87: 485–496.

Toth, G., Gaspari, Z. and Jurka, J. (2000). Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Res.* 10: 967–981.

Vaughan, D.A. (1989). The Genus *Oryza* L. Current status of taxonomy IRRI Research paper, series No. 138, Los Banos, Philippines.

Vaughan, D.A. and Morishima, H. (2003). Biosystematics of the genus *Oryza*. In: *Rice: Origin, History, Technology, and Production* (W. Smith, ed.). John Wiley and Sons, Inc. New York, N.Y. USA, pp 27-65.

Vaughan, D.A. and Sitch, L.A. (1991). Gene flow from the jungle to farmers. *BioSci.* 41: 22–28.

Vaughan, D.A. eds. (1994). The Wild Relatives of Rice. A *Genetic Resources Handbook*. International Rice Research Institute. Los Banos, Philippines, 112 PP.

Vaughan, D.A., Lu, B. and Tomooka, N. (2008). The evolving story of rice evolution. *Plant Sci*. 174: 394-408.

Vaughan, D.A., Morishima, H. and Kadowaki, K. (2005). Diversity in the genus *Oryza. Curr*. *Opi*. *in Plant Biol*. 6: 139–146.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van, D.T., Hornes, M., Frijters, A., Pot, J., Peleman, J. and Kuiper, M. (1995). AFLP: A new

technique for DNA fingerprinting. *Nucleic Acids Res.* 23: 4407–4414.

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

Wang, Z., Weber, J.L., Zhong, G. and Tanksley, S.D. (1994). Survey of plant short tandem DNA repeats. *Theor. Appl. Genet.* 88: 1–6.

Wang, Z.Y., Second, G. and Tanksley, S.D. (1992). Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theor*. *Appl*. *Genet*. 83: 565-581.

Weber, J.L. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* 44: 388–396.

Weeden, N.F. and Wendel, J.F. (1990). Isozyme analysis of plant mating. Visualization and interpretation of plant isozymes. In: *Isozymes in Plant Biology*, pp. 5-46, (Soltis, D.E. and Soltis, P.S., eds). Chapman and Hall, London.

Weir, B.S. and Cockerham, C.C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evol*. 38: 1358-1370.

Zeng, Y., Yang, S., Cui, H., Yang, X., Xu, L., Du, J., Pu, X., Li, Z., Cheng, Z. and Huang, X. (2009). QTLs of cold-related traits at the booting stage for NIL-RILs in rice revealed by SSR. *Genes Genom.* 31: 143–145.

Zhu, Q. and Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytol*. 167:249–265.

Declaration:

I, the undersigned, here by declare that this thesis is my original work and that all sources of materials used for the thesis have been duly acknowledged.

Name:_____

Signature: _____

Date:_____