MORPHOMETRIC, PHOTOSYNTHETIC AND MOLECULAR MARKERS BASED GENETIC DIVERSITY STUDY OF THE AFRICAN WILD RICE (*Oryza longistaminata* Chev. et Roher) FROM ETHIOPIA

GETACHEW MELAKU ADANE

Addis Ababa University
Addis Ababa, Ethiopia
June, 2018

MORPHOMETRIC, PHOTOSYNTHETIC AND MOLECULAR MARKERS BASED GENETIC DIVERSITY STUDY OF THE AFRICAN WILD RICE (*Oryza longistaminata* Chev. et Roher) FROM ETHIOPIA

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES ADDIS ABABA UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PhD) IN BIOLOGY (APPLIED GENETICS)

BY
GETACHEW MELAKU ADANE

Addis Ababa
June, 2018

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

Morphometric, Photosynthetic and Molecular Markers Based Genetic Diversity Study of the African Wild Rice (*Oryza longistaminata Chev.* Et Roher) from Ethiopia

By

Getachew Melaku Adane

A Thesis Presented to the School of Graduate Studies of the Addis Ababa University in Partial Fulfillment of the Requirements for the PhD in Biology (Applied Genetics)

Approved by Examining Board:

1. Dr. Kassahun Tesfaye (Advisor) ____________
2. Dr. T/Haimanot H/Selsassie (Advisor) ____________
3. Prof. Edward George Mamati (External Examiner) ____________
4. Dr. Kifle Dagne (Internal Examiner) ____________
5. Dr. Gurja Belay (Chairman) ____________

Department Of Microbial Cellular Molecular Biology

Declaration

I, the undersigned, declare that this Dissertation is my original work and its composition has never been submitted elsewhere for any other award. All sources of materials used for the Dissertation have been duly acknowledged.

Name: Getachew Melaku Adane Signature: __________________ Date: ________________

Acknowledgements

Above all, I thank the Almighty GOD for keeping me alive and giving me the courage during this study.

I would like to extend my most sincere and very special gratitude to my major advisor, Dr. Teklehaimanot Haileselassie. I have no words to thank him for all of his humanitarian and academic qualities. My progress since I joined Addis Ababa University for my MSc. study was the blueprint of his tremendous support. His technical guidance and advice on my carrier were priceless. Most importantly, he has been a great mentor and “esteemed friend”.

I highly appreciate the supervision I got from Professor Fengyi Hu, School of Agriculture, Yunnan University, China. As a PhD candidate, I learned a lot from his unwavering mentorship, guidance, knowledge and experience. His joy and enthusiasm towards a scientific research was highly communicable to me. I am also thankful for his selflessly availed support during my two years stay in his laboratory. Without the use of his funding sources, successful completion of this study would not have been possible.

I would also like to thank the co-advisor, Dr. Kassahun Tesfaye, for his professional supervision and encouragement.

I am grateful to my sponsor, Adigrat University, and appreciate the support by Dr. Alem Mebratu, Academic Vice President of Adigrat University. Equally importantly, I extend my sincere gratitude to the Department of Microbial, Cellular, and Molecular Biology of Addis Ababa University and its staff for the multidisciplinary support and budget allocated for an on time sample collection and field survey.

This study was further supported by the Yunnan Provincial Sciences and Technology Department (CN) grant to Prof. Fengyi Hu.

I appreciate the research scholarship and logistic support of Yunnan University as well.

I am very thankful to my research team mates Guangfu Huang, Liyu Huang, Zhang Shilai and Zhang Jing who fulfilled my appeal without any reservation.

I am greatly indebted to the highly committed support from Fekadu Gadissa (PhD. Candidate of the Microbial Cellular and Molecular Biology Department of Addis Ababa University), Tesfahun Bufebo, Fiseha Gidey and Gabriel Temesgen from Adigrat University.

I am very thankful to the technical assistance and guidance from Dr. Erik Sacks and Marlee Labroo both from the Department of Crop Sciences, University of Illinois, U.S.A.

There are no words to convey the lifelong sacrifice of my beloved parents, Ato Melaku Adane and W/ro Zewdenesh Fanta. I will never forget their love and integrity. It is unthinkable to pay back but I always wish them a happy and long life.

My thanks further extend to my dear sisters Meron Melaku, Sentayehu Melaku, Weyneshet Melaku and Rediet Melaku along with their family members; Dr. Sisay Kebede, Menilik Biru, Ephrata Sisay, Ephrem Sisay, Veronica Menilik, Miklol Sisay and Haset Menilik.

At last my deepest gratitude goes to my beloved wife, Tsige Mebratu, and my cute boy Alazar Getachew. Without their presence, love, and care my life might go differently.

Table of contents

List of tables

List of tables continued

Page

List of tables continued

Page

List of figures

Page

List of figures continued

Appendix

List of Acronyms/abbrevations

AMOVA – Analysis of Molecular Variance

BIC – Bayesian information criterion

DAPC – Discriminant Analysis of Principal Components

Df – Degrees of Freedom

Est. Var. – estimated variance

F – Fixation Index

F_{is} – Wrights fixation index

F_{it} –Total inbreeding coefficient

F_{st} –genetic differentiation

He –Expected heterozygosity

Ho –Observed heterozygosity

MS – Mean squares

MYA – Million years of age

N – Sample size

Na – Actual number of alleles

Ne – Effective number of alleles

PCA – Principal Component Analysis

PIC – Polymorphic Information Content

List of Acronyms/abbrevations continued

PCR – Polymerase Chain Reaction

RADSeq – Restriction Associated DNA Sequencing

RM – Rice Marker

SE– Standard Error

SS – Sum of squares

SSR – Simple Sequence Repeats

UPGMA –Unweighted Pair Group Method with Arithmetic Mean

Abstract

MORPHOMETRIC, PHOTOSYNTHETIC AND MOLECULAR MARKERS BASED GENETIC DIVERSITY STUDY OF THE AFRICAN WILD RICE (*Oryza longistaminata* Chev. et Roher) FROM ETHIOPIA

Getachew Melaku Adane, PhD Dissertation

Addis Ababa University, June 2018

Rice which feeds more than half of the world human population is affected by environmental changes. Therefore, its productivity should be supplemented by the genetic resources from other wild relatives. The African wild rice, *Oryza longistaminata* Chev. et. Roher, which possess AA genome as the cultivated rice, is hypothesized as a good candidate for widening the cultivated rice gene pool. However, its status of genetic diversity and relationship with other *Oryza* members is still unresolved. So far, the genetic resource from this African species is underutilized for the breeding of cultivated rice. Even, characterization and evaluation of the different agro-morphological and photosynthetic traits of *O. longistaminata* are still untapped. To this end, assessing the morphometric, photosynthetic and molecular markers based genetic diversity and regional differentiation of the African wild rice (*O. longistaminata*) in Ethiopia is very important for efficient breeding and conservation schemes. This study therefore, employed 67 SSR markers over 360 *O. longistaminata* accessions from 12 populations of Ethiopia, one *O. longistaminata* accession from Niger, 35 varieties of the *Indica* and *Japonica* subspecies of *O. sativa*, 25 accessions of four AA genome wild *Oryza* species, 8 accessions of five CC genome wild rice species and one *Japonica* weedy type for the assessment of genetic diversity and relatedness. In this study, evaluation of morphological differences and characterization of leaf gas exchange traits were done from the three replicates of 62 Ethiopian *O. longistaminata* accessions. The Restriction Associated DNA Sequencing data from 87 *O. longistaminata* accessions of Ethiopia was also analyzed for their phylogenetic relationships and region based differentiation.

The eight morphometric and seven leaf gas exchange traits assessed in this study revealed a significant variation. Correlations between the evaluated morphological traits of the assessed accessions showed highly significant pairwise associations. The multivariate analysis from the eight morphological traits implicated population effect. In particular, the inter-population distances among the six populations of origin showed distinctness of the Amhara from the rest five Gambella populations. The Pearson pairwise correlation matrix for the seven leaf gas exchange traits revealed their inter-relationship. For instance, the rate of photosynthesis was positively and significantly ($P < 0.001$) correlated with all other traits except the intercellular CO_2 concentration. Moreover, the accessions showed high photosynthetic rate with a minimum intercellular CO_2 concentration, transpiration rate and stomatal conductance.

When the 12 *O. longistaminata* populations from Ethiopia were pooled as one, the 67 SSR markers showed (Na = 4.35, Ne = 1.91, Ho = 0.15, He = 0.38, F = 0.53 and PIC = 83.82). Such diversity indices implicated higher diversity than the rest AA and CC genome rice types. A highly significant ($P> 0.001$) genetic differentiation (F_{st}) = 0.299 was also observed among the 12 *O. longistaminata* populations of Ethiopia. However, the Analysis of Molecular Variance (AMOVA) showed the highest variance (38%) among accessions with in a population. PCA of the whole 430 accessions showed five main clusters and a clear geographic pattern was emerged for the 12 populations of *O. longistaminata* from Ethiopia. Population structuring and clustering through UPGMA and hierarchical clustering clearly separated the 12 Ethiopian *O. longistaminata* populations from other wild and cultivated rice types. Moreover, structure analysis from the SSR data indicated genetic distinctness in between the Amhara and Gambella regions. Dendrogram, DAPC, and population divergence

(*Fst*) analysis from the 30,256 Single Nucleotide Polymorphisms (SNPs) generated by the Restriction Associated DNA Sequencing also showed genetic differentiation among the two regions.

In conclusion, the high diversity status of the Ethiopian *O. longistaminata* populations and their genetic distinctness to other *Oryza* species or in particular to *O. sativa* is an ideal opportunity for their utilization in diverse rice improvement programs. Moreover, the low transpiration rate and stomatal conductance records with a maximum Carboxylation Efficiency of the *O. longistaminata* accessions showed their environmental adaptability and high water use efficiency. The high Photosynthetic rate observed in the assessed accessions implicated that *O. longistaminata* could be used as a genetic resource for the photosynthetic improvement of rice cultivars. Therefore, the very efficient carbon balance mechanism and adaptation potential of *O. longistaminata* can be utilized for developing rice varieties and hybrids with higher yield even under harsh environment conditions.

1 Introduction

1.1 Background and justification

The genus *Oryza* is an economically important plant group, comprising the cultivated rice (*Oryza sativa*) (Khush 1997). Though cultivated rice supports more than half of the world's human population, the narrow gene pool of modern rice is affecting its cultivation under diverse climatic regimes (Zhu *et al*., 2007). Therefore, enhancing rice productivity, profitability and sustainability has become a major concern of diverse rice breeding programs (Haritha *et al*., 2017).

Undomesticated species of the genus *Oryza*, which are distributed globally, are valuable repository of useful genetic variation (Chouhan *et al*., 2014; Giuliani *et al*., 2013). The diverse and agronomically useful traits of such wild rice can serve for the improvement of rice yield (Brar and Khush, 1997). However, the vast majority of wild rice genes remain untapped to date (Ren *et al*., 2003). Therefore, precise genetic manipulation for the desirable traits of wild rice should be started from morphometric and physiological characterization along with a prior and thorough assessment of genetic diversity, population structure and kinship relatedness (Nachimuthu *et al*., 2015).

Of the different wild relatives, the African wild rice (*O. longistaminata*) is identified as a potential donor of perennial traits and a useful resource for the development of permanent ground cover rice cultivars (Sacks *et al*., 2003). The strong rhizome of *O. longistaminata* provides nitrogen use efficiency, high biomass production and minimal soil erosion (Song *et al*., 1995; Yang *et al*., 2010). This allogamus rice species is resistant to bacterial blight, the nematode *Meloidogyne graminicola*, and drought (Second *et al*., 1977; Taillebois 1983; Khush *et al*., 1990; WARDA 1997; Jones *et al*., 1996).

According to Wambugu *et al*., (2013), the vast range of phenotypes exhibited by the undomesticated rice species was not fully exploited. However, high yield potential of rice is associated with cultivars possessing effective plant morphology and higher leaf photosynthetic rate (Horie *et al*., 2003; Takai *et al*., 2006).

These days, characterization, evaluation and classification of rice genetic resources for several agronomic and physiological importances is exemplified by multivariate methods (Peeters and Martinelli, 1989). The information generated from such multivariate analysis of morphological traits can be useful for revealing patterns of variation in the collections along with the identification and crossing of groups of accessions with a trait of agronomic interest (Brown, 1991; Perry and McIntosh, 1991; Souza and Sorrells, 1991; Cowen and Frey, 1987; Camussi *et al*., 1985).

According to the suggestion from Cao *et al*. (2001) and Ohno (1976), selecting accessions with a high photosynthetic rate (Pn) from significantly different rice varieties is one of the means to come up with high-yielding varieties. Modifying the primary photosynthetic pathway and the process of biomass production under changing climate are therefore important for sustaining global food production (Kajala *et al*., 2011; Zhu *et al*., 2008). Better performance of wild rice species under adverse condition can thus renewed further utilization of their photosynthesis related genetic resources (Ding *et al*., 2014).

More interestingly, the African wild rice accessions showed a superior photosynthetic rate (Kondamudi *et al*., 2016). Therefore, the extremely efficient carbon balance system by the *O. longistaminata* accessions must be assessed with the corresponding flag leaf size and shape and physiological traits (Yue *et al*., 2006), stomatal conductance (Ohsumi *et al*., 2007a; Maruyama and Tajima, 1990), transpiration rate (Teng *et al*., 2004) and immense pool of useful genes (Cissé and Khouma, 2016).

In Ethiopia, the different agro-ecologies with unique environmental and climatic features resulted in a wider genepool of *O. longistaminata* (Melaku, 2011). Genetic diversity assessment using five SSR markers and 320 *O. longistaminata* accessions of Ethiopia was found to be high (Melaku *et al*., 2013). However, the diversified ecological threats speculated this African species as an important focus of conservation priority (Wambugu *et al*., 2013). Thus, adequate number of accessions and SSR primers along with Genome Based Sequencing data such as RADSeq must be utilized for a reliable genetic structure analysis and better resolution of the two ecologically distinct (Amhara and Gambella) regions of Ethiopia.

In a general sense, Ethiopia, which is a center of genetic diversity for many plants (Mengesha, 1975; Harlan, 1969; Vavilov, 1951), is a potential hot spot for the highly diversified gene pool of *O. longistaminata* (Melaku *et al*., 2013). However, the morphological and molecular markers based genetic status and photosynthetic related traits of the Ethiopian *O. longistaminata* accessions are not well characterized. The aim of the present study was therefore, to characterize and evaluate the morphological and photosynthetic features of the Ethiopian *O. longistaminata* accessions and assess their genetic diversity and relatedness with other worldwide *Oryza* collections.

1.2 Hypotheses

The study was carried out with the hypotheses:

1. The African wild rice which is adapted to the different geographic and climatic regions of Ethiopia might show different patterns of morphological diversity,
2. Genetic diversity would differ within and among populations of the Ethiopian *Oryza longistaminata*,
3. *Oryza longistaminata* collections from the Amhara region could be genetically differentiated from the Gambella regional state of Ethiopia,
4. The Ethiopian *Oryza longistaminata* populations could genetically be distinct from other wild and cultivated *Oryza* species elsewhere, and
5. *Oryza longistaminata* collections from Ethiopia would have a high photosynthetic rate with maximum water use efficiency.

1.3 Objectives

1.3.1 General objective

To characterize the morphometric and photosynthetic traits of the Ethiopian *O. longistaminata* accessions and assess their genetic diversity, regional differentiation and genetic relatedness with wild and cultivated rice collections

1.3.2 Specific objectives

i. Determine morphological variability of *O. longistaminata* collections from Ethiopia

ii. Assess the different photosynthetic and leaf gas-exchange traits of the Ethiopian *O. longistaminata* accessions for their use in future rice breeding programs

iii. Evaluate the genetic diversity within and among the Ethiopian *O. longistaminata* populations and their genetic relatedness with other *Oryza* species using SSR markers, and

iv. Investigate differentiation between the Amhara and Gambella *O. longistaminata* collections through both SSR markers and RADSeq data analysis.

2 Literature review

2.1 Origin, domestication and cultivation of rice

Extensive archeological evidence reported rice cultivation before 10,000 years ago in the middle Yangtze and upper Huai rivers of China (Snyderand and Slaton, 2001; Nayar, 2014a). In the next 2,000 years, rice cultivation spread down these rivers and/or to different parts of Asia (GRiSP, 2013). During 344-324 B.C., expedition of members of Alexander the Great to India allowed introduction of rice to Greece and the neighboring areas of the Mediterranean (Nayar, 2014a). From the center in Greece and Sicily, the Europe’s great Age of Exploration promoted spreading of rice throughout southern Europe, Central and South America, and few areas of North Africa (GRiSP, 2013).

In the 12th century, rice cultivation spread steadily to Egypt (Nayar, 2014a). Though less attention is given to the historical aspects of rice introduction to east Africa, some archaeologists and historians reported recovery of rice from the Red Sea port of Berenike and Myos Hormos (Egypt) (Cappers, 2007). In the first century, this port was a trans-shipment and trading port for goods emanating from eastern Africa (Nayar, 2014a).

After all these historical events of distribution, rice become widely adapted to diverse and very wide types of environmental conditions (Nayar, 2014b) and grown in latitudes ranging from 53° North to 35° South (Lu and Chang, 1980) and altitudes ranging from 4.0 meters below sea level to about 2,600 meters above sea level (Figure 1). Rice is produced under a variety of climatic conditions like the high slopes of Nepal’s mountains, mountains of Himalayas and hot deserts of Egypt (Kondamudi *et al.*, 2016). This wide adaptive feature of rice allows its cultivation in 177 countries of the world (Nayar, 2014b).

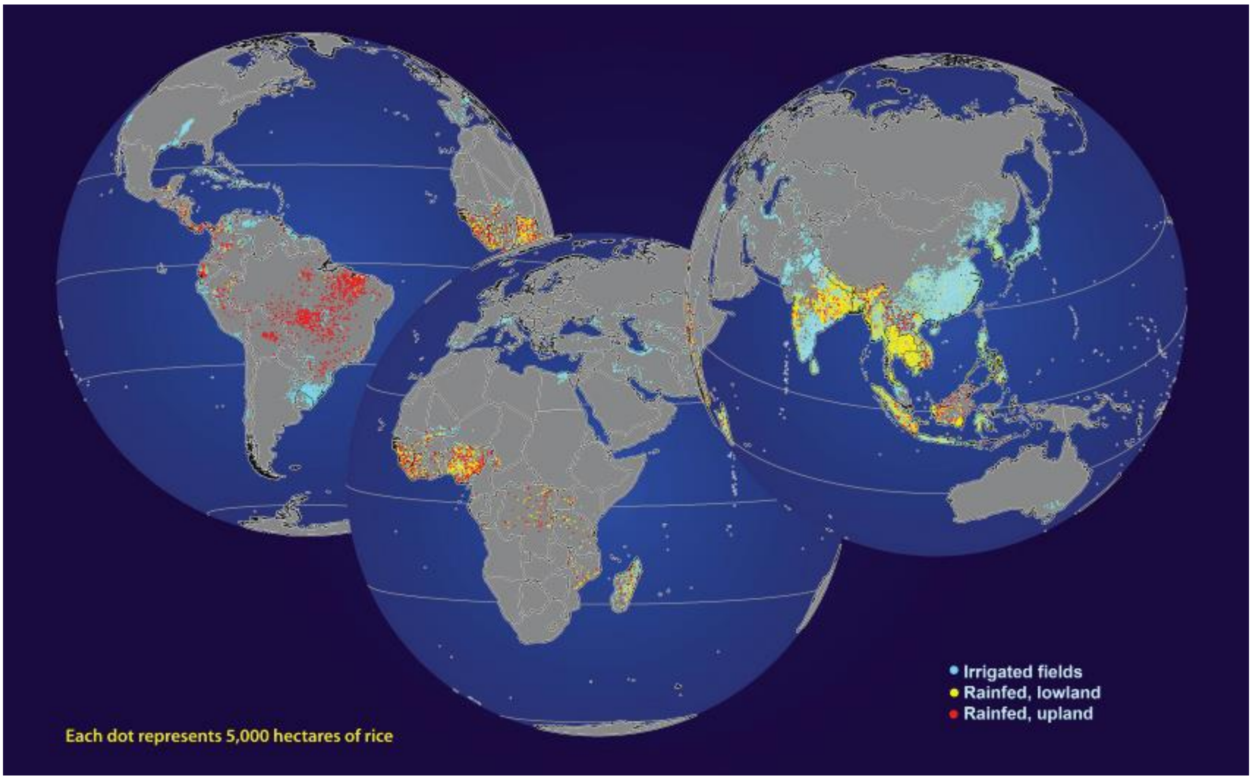

Figure 1. Major global rice-growing areas and ecosystems (Source: Nyara, 2014a)

In terms of cultivation area and production, rice is taken as the most important food crop of humankind (FAO, 2011). In 2010, around 696 million ton of rice was produced from 11% of the world's total arable land (Guimara˜es, 2009). Asia accounts for 90% of the world's production and consumption of rice because of its favorable warm and humid climate (Figure 2). In 2007, Africa, America, Europe and Oceania had a total rice production of 3.2%, 5.1%, 0.5% and 0.02%, respectively (Nayar, 2014a).

Though rice production is steadily increasing in Africa, it still ranked only fifth in terms of area under cultivation (Nayar, 2014a). In the continent, rice provides 8.0% of the calories and occupies 4.4% of the net cropped area (Nayar, 2010). Egypt, the west and east Africa are the major rice cultivation regions of Africa (FAOSTAT, 2010; Nayar, 2010). However, the 13.9 million ton annual production from Africa is only enough for half of the annual requirements in the continent (Nayar, 2014a).

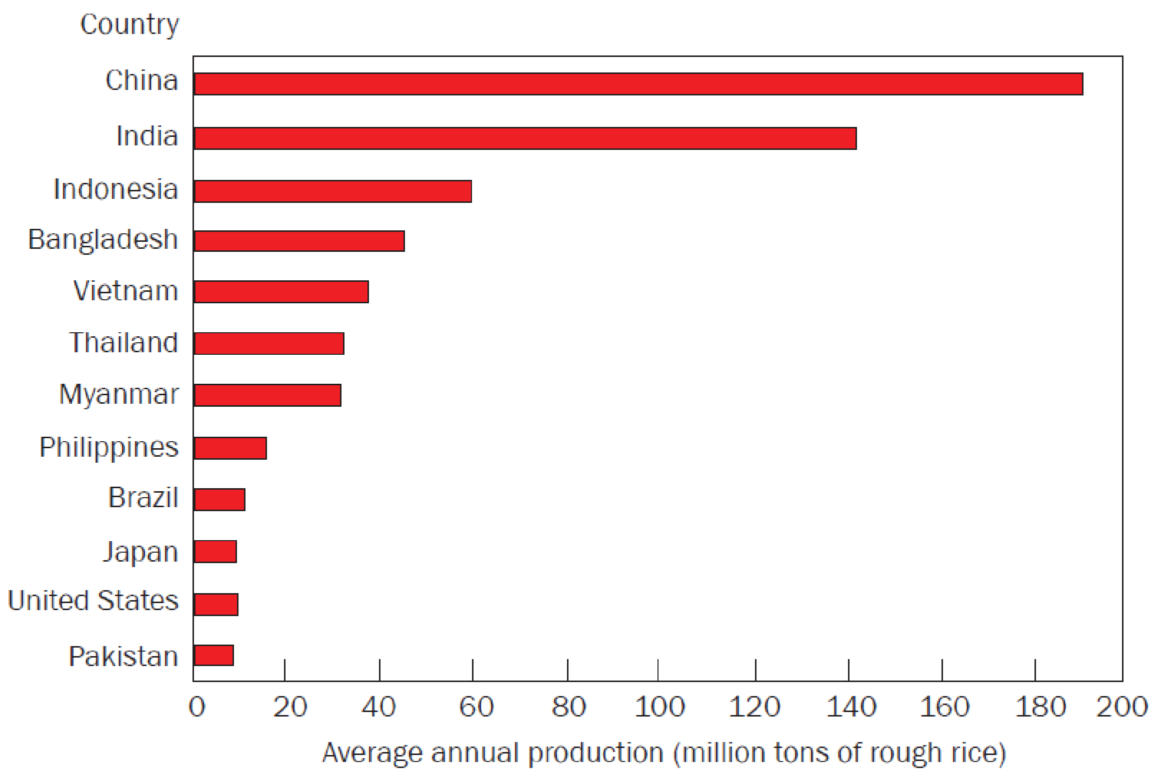

Figure 2. Leading rice producers in the world, 2006-2010 (Source: GRiSP, 2013)

The east African country, Ethiopia, has the major rice ecosystem types (Claudia and Knauer. 2013). Due to its agro-ecological diversity, Ethiopia can have a great potential for rice cultivation (Shahi, 1985). Based on the water regime, drainage, temperature, soil type and topography, rice growing ecosystems can be classified in to four major farming systems called rainfed upland, rainfed lowland, irrigated and deep water rice ecosystems (Claudia and Knauer, 2013). As indicated in Table 1, three of the major rice growing ecosystems (rain fed upland, rain fed lowland and irrigated) are found in Ethiopia.

Table 1. Characterisation of rice growing ecosystems, cropping seasons and production constraints in Ethiopia (Source: MoA, 2010).

Rice Ecosystem	**Rain fed low land production system**	**Rain fed upland production system**	**Irrigated production system**
Location	Parts of South Gondar, Pawe, Mai-Tsebri, Kamashi, Gojeb, Chewaka	Parts of South Gondar, Pwe, Assosa (Kamashi), Tepi, Jimma, Gondar, Chewaka, Guraferda, Gabmela	Werer, Mehoni, Gode, Dupti
Cropping seasons	Meher (June-December)	Meher (June-December)	Meher, Belg (June-December),
Cropping system	Mixed farming System(crop-livestock)	System(crop-livestock) System(crop-livestock)	Pastoralist and agro – pastoralist mixed farming System(crop-livestock)
Rice production constraints	Biotic stress (blast, Brown spot, RYMV, Sheath blight); Abiotic stress (Cold, drought), improved variety, availability of quality seeds and moisture stress	Biotic stress (Termite, blast, Brown spot, RYMV, Sheath blight), Abiotic stress (Cold and drought), lack of quality seeds and moisture stress	Abiotic stress (Salinity, acidity, heat), lack of improved rice variety, irrigation facility

In Ethiopia, rice is produced under rain-fed and irrigation condition in all regions except Dire Dawa (MoA, 2010). GIS data from the rainfall, slope, soil texture, altitude, and temperature of the country indicate the potential of Ethiopia for a rain fed rice cultivation (Figure 3). According to the report of MoA (2010), the potential area for rain fed rice production was estimated to be about 30 million hectares, of which more than 5.6 million ha was highly suitable. Besides, 3.8 million hectare of land is also suitable for irrigation (Figure 3).

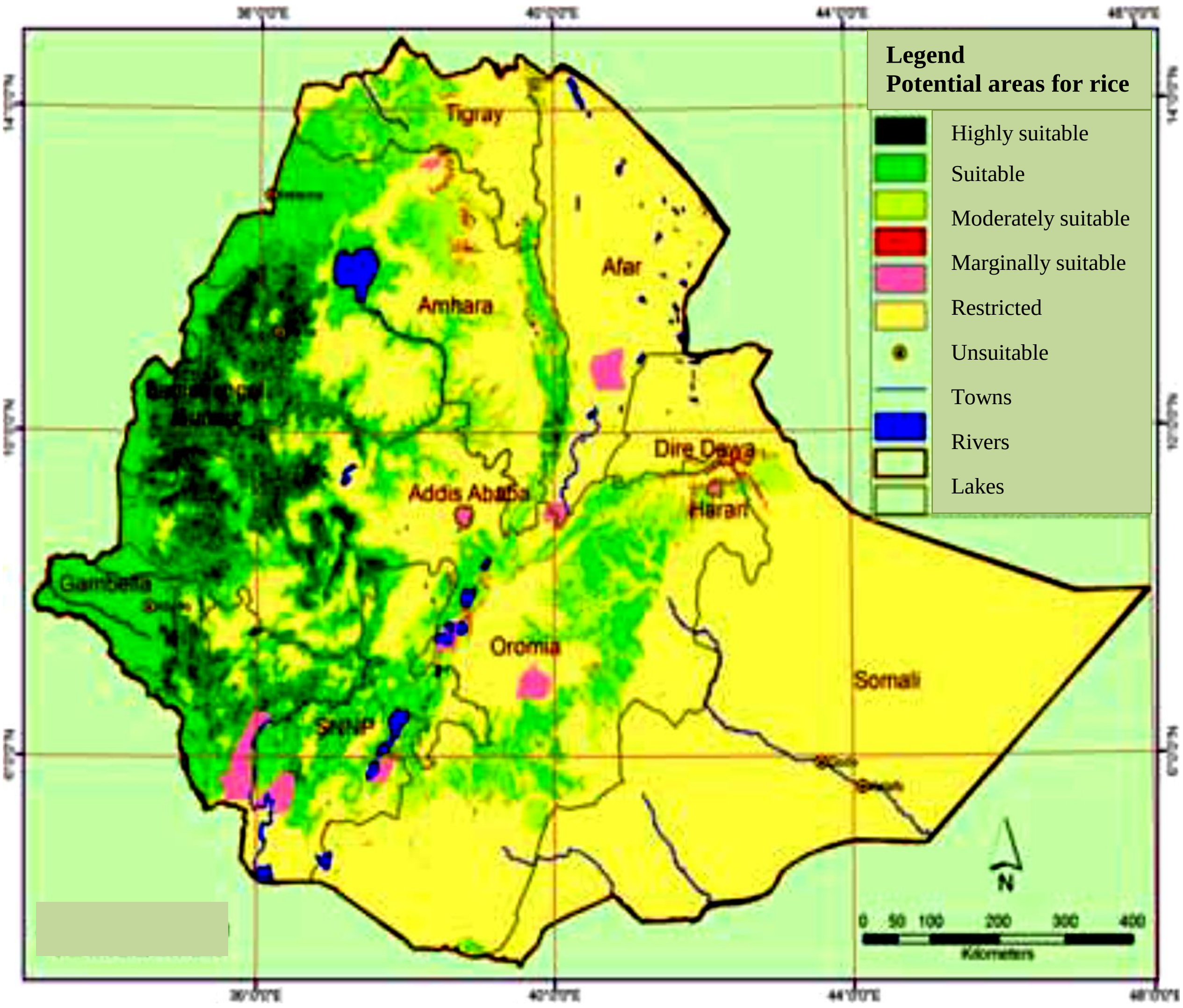

Figure 3. Suitability map of rain-fed rice production in Ethiopia (Source: MoA, 2010)

From the very recent rice cultivation history of Ethiopia, a general increment was observed for grain production, area coverage, productivity and number of households engaging in rice farming (Fig. 4 and Fig. 5). Unless the current production especially, the area covered under rice cultivation gets a rapid increase, the country will be dependent on imported rice from abroad and fail to export rice in the near future (MoA, 2010).

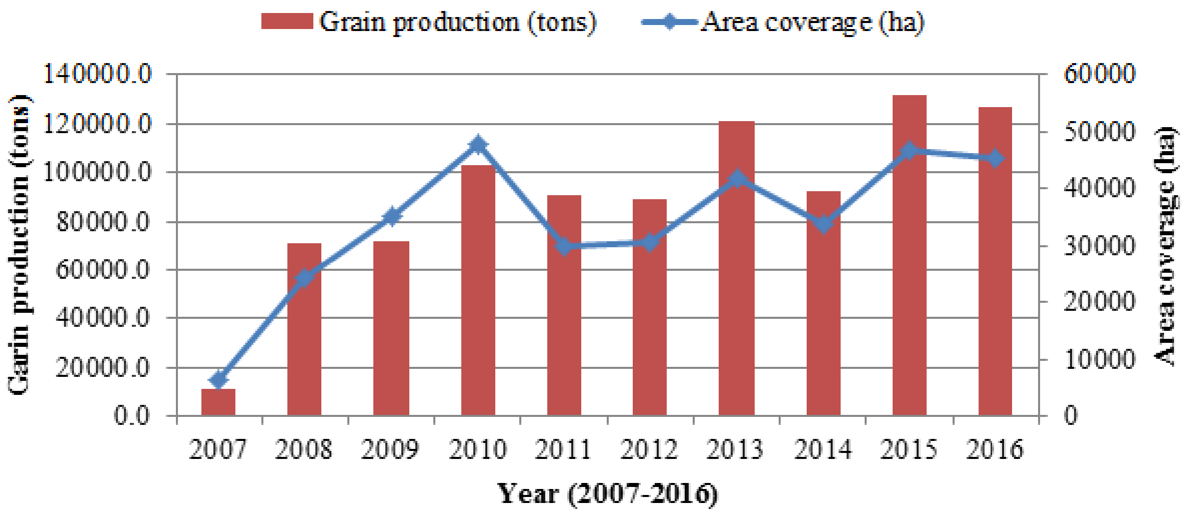

Figure 4. Annual rice grain production and area coverage in Ethiopia during 2007-2016 (Source: MoA, 2017).

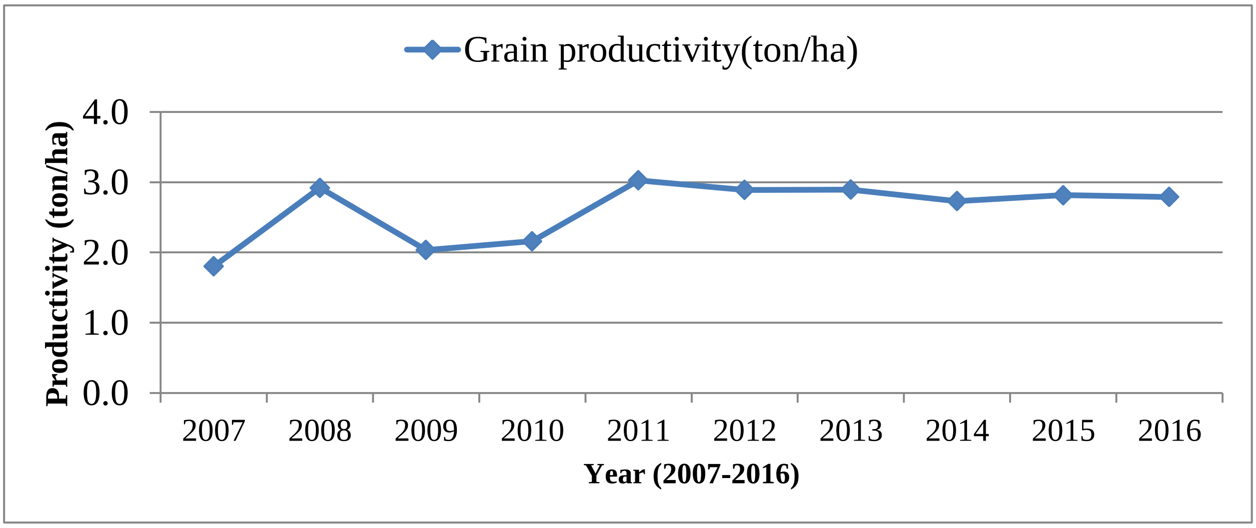

Figure 5. Trend in annual gain productivity of rice in Ethiopia from 2007 to 2016 (Source: MoA, 2017).

Farmers in different parts of Ethiopia have shown interest to get involved in rice production and this will provide them with the opportunity contributing to food security (MoA, 2010). Ethiopia can also earn hard currency by exporting the surplus rice (CSA, 2007-2016). However, production, productivity and expansion of rice farming has been constrained by lack of improved and adaptive varieties, poor crop management practices customized for different rice ecosystems, lack of pre and postharvest management technologies, lack of awareness on its utilization, and other biotic and abiotic stresses (Tesfaye *et al*., 2005). Therefore, the currently available rice cultivars must be tested in different agroecologies of Ethiopia and their genetic relatedness with other wild rice types such as *O. longistaminata* should be assessed for future breeding programs.

2.2 Divergence in the tribe *Oryzeae*

The tribe Oryzeae is the largest tribe in the subfamily Ehrhartoideae (Kellogg, 2001). Currently, the tribe is distributed in the tropical and temperate regions of the world and represented by 11 genera (Tzvelev, 1989; Vaughan, 1994). Five genera of Oryzeae are found in America, Africa and Asia and three of its genera; *Leersia*, *Luziola* and *Oryza* have five or fewer extant species in each continent (Vaughan *et al*., 2008).

Before 60 million years ago, plant species resembling members of the tribe Oryzeae were present in India (Prasad *et al*., 2005). For decades, lack of adequate fossil records and molecular evidences contradicted the origin and diversification of the tribe (Vaughan and Morishima, 2003). From the very few isozyme studies, earlier taxonomists classified the tribe Oryzeae in to two groups or subtribes called Oryzinae and Zizaniinae (Ge *et al*., 2002; Guo and Ge, 2005).

However, most of the current cytological, RFLP based, and DNA sequencing studies concluded the tribe Oryzeae to be a monophyletic group (Marathi *et al*., 2015). Phylogenetic study from the combined data sets of *mat*K, *trn*L-*trn*F, *nad*1 and *Adh*2 and *GPA*1 over 12 genera of the tribe *Oryzae*, also showed the monophyletic divergence of the tribe Oryzeae and inclusion of the three genera, *Leersia*, *Oryza*, and *Porteresia* in one clade (Ge *et al*., 2005).

The synonymous molecular clocks of *mat*K and *Adh*2 on the rice tribe roughly estimated 35 Million Years of Age (MYA) for origination of the tribe, 20 MYA for the separation of *Oryza* and *Leersia* from the rest of the tribe and 14 MYA from each other (Gaut, 2002). With regard to evolutionary terms, the genus *Oryza* is the most remarkably successful genus of all the other genera of Oryzeae (Vaughan *et al*., 2008) and age of the deepest split between the most basal G genome and the rest genome types was estimated to be 8 MYA (Ge *et al*., 2004).

2.3 The genus *Oryza*

The name *Oryza* was given to the rice genus by Linnaeus in 1753 (Vaughan *et al*., 2003). Rudimentary sterile lemmas, bisexual spikelets, narrow, linear and herbaceous leaves with scabrous margins are the basic morphological features of *Oryza* species (Launert 1965; Tateoka 1964). *Oryza* is a genus under the subfamily Oryzoideae and in the family Poaceae (Graminae) (Chang and Vaughan, 1989; Duistermaat, 1987; Second, 1985b). *Oryza* is a small genus containing the two cultivated rice species (*O. sativa* and *O. glaberrima*) and 21 wild species (Table 2).

Species identification and classification in the genus *Oryza* are attributed to the spikelet structure (Vaughan, 1988). Since spikelet morphology has less power in distinguishing closely related wild species, most of the current taxonomic and phylogenetic studies are based on advanced cytological and molecular approaches (Nayar, 1973; Duvall *et al*., 1993; Ge *et al*., 2002). Knowledge about genome structure of *Oryza* members is of great value in understanding homologous and homeologous relationship between different genomes (Marathi *et al*., 2015).

Cytological studies in the genus *Oryza* identified 11 genomes for representing an enormous gene pool of both the cultivated and wild rice species (Lu *et al*., 2009). Those genome types across the whole genus are also used as keys for classifying *Oryza* species into basic groups or species complexes (Table 2).

Table 2. Chromosomal number, genome group and habitat types of *Oryza* species (Source: Vaughan *et al*., 2003)

Species/species complex	Chromosome number	Genome group	Usual habitat
***Oryza sativa* complex**			
O. sativa L.	24	AA	Upland to deep water
O. rufipogon sensu lacto	24	AA	Seasonally deep water, annual and wet year round
O. glaberima Steud	24	AA	Upland to deep water
O. barthi A. Chev	24	AA	Seasonally dry
O. longistaminata Chev. et Roehr.	24	AA	Seasonally dry to deep water
O. meridionalis Ng	24	AA	Seasonally dry
O. glumaepatula steud.	24	AA	Inundated areas that become seasonally dry
***Oryza officinalis* complex**			
O. officinalis Wall ex Watt	24	CC	Seasonally dry
O. minuta JS Presl. ex CB Presl.	48	BBCC	Stream sides, semi shade
O. rhizomatis Vaughan	24	CC	Seasonally dry
O. eichingeri Peter	24	CC	Stream sides, forest floor, semi shade
O. malapuzhaensis Krishnaswamy and Chandrasakaran	48	BBCC	Seasonally dry, forest pools, shade
O. punctata Kotschy ex Steud.	24 48	BB, BBCC	Seasonally dry (diploid) and forest floor, semi shade (Tetraploid)
O. latifolia Desv.	48	CCDD	Seasonally dry
O. alta Swallen	48	CCDD	Seasonally inundated
O. grandiglumis (Doell.) Prod.	48	CCDD	Seasonally inundated
O. australiensis Domin	24	EE	Seasonally dry
Ridleyanae Tateoka			
O. schlechteri Pillger	48	Unknown	River banks
***Oryza ridleyi* complex**			
O. ridleyi Hook.	48	HHJJ	Seasonally inundated, forest floor, shade
O. longiglumis Jansen	48	HHJJ	Seasonally inundated, forest floor, shade
***Oryza granulata* complex**			
O. granulata Nees et Arn ex Watt	24	GG	forest floor, shade
O. meyeriana (Zoll. et Mor. ex Steud.) Baill.	24	GG	forest floor, shade
Brachyantha B.R. Lu			
O. brachyantha Chev. et Roehr.	24	FF	Rock pools

2.3.1 The *Oryza* members and their species complexes

Species complex refers to a group of species devoid of good taxonomic key characters (Vaughan *et al*., 2008). From ecological, genetic and morphological perspectives, all *Oryza* species except *O. schlechteri* and *O. brachyantha*, fall into four clearly defined groups called the *O. granulata*, *O. ridleyi*, *O. officinalis* and *O. sativa* species complexes (Table 2).

One of the species with no obvious relative in the genus *Oryza* is the *Oryza brachyantha* (Table 2). This species has a small, narrow spikelet with a long awn (Vaughan, 1988). Of all members of the genus *Oryza*, *O. brachyantha* is the most closely related species to the genus *Leersia* (Launert, 1965; Vaughan, 1988). However, the coriaceous and rigid structured awn with a single vascular bundle of own let it under the genus *Oryza* (Launert, 1965). The genome of this diploid species is designated as FF and it is totally different from all other species of the genus *Oryza* (Li *et al*., 1961; Yang *et al*., 1965).

The other taxa which is not categorized under a species complex is *Oryza schlechteri* (Tateoka, 1962b). This species is a tufted and 1m tall perennial having an erect 4 to 5cm long panicle with a small and unawned spikelets (Duisternmaat, 1987). According to Vaughan (1988), *O. schlechteri* is the least studied species with no herbarium and germplasm collections. This diminutive species which is believed to be extinct could spread horizontally or climb over other vegetation through its prolific stolon (Vaughan *et al*., 2008). Though its complete genomic information and genetic relationship to other *Oryza* species is unknown (Vaughan, 1988), molecular studies over part of the genome suggested *O. schlechteri* as HHKK tetraploid species (Ge *et al*., 1999).

From the four species complex of the genus *Oryza*, *O. granulata* complex is the one consisting of taxonomically confusing species (Vaughan, 1988). The spikelet length which could be varied as far as many samples are considered is the basic taxonomic key for categorizing species under this complex (Tateoka, 1962b). Other than the highest weight given to spikelet length, poor representations of the species in herbaria and less germplasm collections complicate the task of classification and taxonomy in the complex (Vaughan, 1988). Such weak taxonomic tools and methods were the cause for the publication of five *Oryza* species; *O. abromeitiana*, *O. granulata*, *O. indandamanica*, *O. meyeriana* and *O. neocaledonica* in the species complex (Vaughan *et al*., 2008).

However, most taxonomists consider *O. abromeitiana*, *O. indandamanica* and *O. neocaledonica* as synonymous or local variants of the two *Oryza* species, *O. granulata* and *O. meyeriana* (Chang 1985; Duistermaat 1987). Consequently, the two recently accepted and widely distributed species in the complex are *O. granulata* and *O. meyeriana* (Vaughan *et al*., 2008). These two species which differ in their panicle length have no defined cytological relationship with other species in the genus (Vaughan, 1988). The most common species of the *O. meyeriana* complex called *Oryza granulata* Nees *et* Arn.*ex* Watt, has spikelets shorter than 6.4 mm and the other species *Oryza meyeriana* (Zoll. *et* Mor. *ex* Steud.) Baill., is recognized by its 6.4-9 mm long spikelet (Tateoka, 1962b).

The second species complex in the genus *Oryza*, called the *O. ridleyi* complex, comprises *O. longiglumis* Jansen and the *O. ridleyi* Hook. f. (Table 2). Unlike members of the *O. granulata* complex, the two species in the *O. ridleyi* complex have no controversial naming (Tateoka, 1962b). These two species are clearly distinguished from other species in the genus *Oryza* (Vaughan, 1988). Both species of the *O. ridleyi* complex have stolons and looking alike to each other (Vaughan *et al*., 2008). However, *O. longiglumis* has much longer sterile lemmas than *O. ridleyi* (Vaughan, 1988). Even if both *O. longiglumis* and *O. ridleyi* are tetraploid and possess HHJJ genome type, *O. ridleyi* can specifically be crossed with *O. sativa*, *O. minuta* and *O. officinalis* (Katayama and Onizuka, 1979; Katayama *et al*., 1981).

The largest complex in the genus *Oryza* is the *O. officinalis* or *O. latifilia* species complex (Tateoka, 1962a; 1965). This group comprises diploid or tetraploid B, C, D or E genome containing species (Vaughan *et al*., 2008). Other than the diploid CC types such as; *O. offcinalis*, *O. rhyzomatis* and *O. eichengeri*, CC genome is a key genomic constituent of six allotetraploid *Oryza* species (Table 2). This polyploidy event can be explained by the fusion between the CC and either BB or DD genome types and giving BBCC or CCDD tetraploid species of the complex (Vaughan *et al*., 2003).

The two African species of the *O. officinalis* complex named, *O. punctata* Kotschy *ex* Steud. and *O. eichingeri* Peter. are reported as species having both diploid and tetraploid forms (Hu 1970; Tateoka 1965). The diploid forms of these species have different genome designations like, BB for *O. punctata* and CC for *O. eichingeri* (Kihara, 1963; Nayar, 1973). Morphological and cytological assessments from the large field collections by Tateoka (1964a), clearly distinguish *O. eichingeri* from *O. punctata*. A study by Sano (1980), listed fundamental characteristics used to distinguish the annual and diploid form of *O. Punctata* from its perennial, allotetraploid form. On the basis of those distinct morphology and genome based races of *O. punctata*, taxonomists proposed the need for subspecific ranking (Vaughan, 1988). For instance, Sharmia and Sampath (1985) used the scientific name *"Oryza schweinfurthiana* Prod." for the tetraploid form of *O. punctata*.

Among the BBCC tetraploid species, *O. minuta* and *O. malampuzhaensis* have cytoplasm from the diploid BB genome of *O. punctata* (Kihara, 1963). However, the tetraploid form of *O. punctata* (BBCC) is believed to have a cytoplasm from a CC genome species (Dally and Second 1990; Kanno and Hirai 1992). Distinct morphology (Li *et al*., 2001), current distribution (Vaughan, 1994) and molecular analyses (Shcherban *et al*., 2000) implicated the different polyploid events of *O. minuta* and *O. malampuzhaensis* (Vaughan *et al*., 2003).

Cytological evidences suggested *O. minuta* as an allopolyploid species from the fusion of CC genome *O. officinalis* and the BB genome *O. punctata* (Kihara, 1963). However, current distribution of *O. punciata* doesn't show its role on the evolution of *O. minuta* (Vaughan, 1988). Some studies such as that of Duistermaat (1987), considered the *O. officinalis* and *O. minuta* as similar species. However, the systematical comparison by Tateoka and Pancho (1963) reported *O. officinalis* and *O. minuta* as clearly distinguishable and independent species. Besides, morphological characterization of these two species recorded smaller spikelet, panicle, and plant size of the tetraploid, *O. minuta* than the diploid *O. officilalis* (Sharma and Sampath, 1985; Tateoka and Pancho, 1963).

The other BBCC tetraploid species called *Oryza malampuzhaensis* also thought to be distinguished from *O. officinalis* (Vaughan, 1988). Despite the two species share similar habitat, they showed a slight spikelets size difference (Tateoka, 1963b). Like other taxa in the genus, *O. malampuzhaensis* is poorly represented in the gene banks and its relationship with other species in the complex is less understood (Vaughan, 1988).

The second allotetraploid group in the *O. officinalis* species complex is explained by the CCDD genomic designation (Kihara, 1963). The three American species called *O. latifolia* Desv., *O. alta* Swallen and *O. grandiglumis* (Doell) Prod. share the same CCDD genome type (Vaughan, 1988). However, these three *Oryza* species are very polymorphic to morphological features such as; leaf width, spikelet morphology, and spikelet dimensions (Chang, 1985; Sharma and Sampath, 1985; Tateoka, 1963b). To the contrary Vaughan *et al*., (2003), suggested those key characteristics as ambiguous for species identification and defined the three species as closely related but diversified due to different ecological conditions. Such field observations were also supported by molecular and cross-compatibility data and interpreted the CCDD genome species as one species complex with different ecotypes (Jena and Kochert, 1991; Nayar, 1973).

Absence of any DD genome *Oryza* species accessed research interest on the three CCDD genome species of Latin America (Vaughan *et al*., 2003). One of the most frequently proposed hypotheses for the origin of DD genome is modification from CC genome (Wu, 1992). However, current cytological studies don't support the concept of genome modification and rather suggest distinctness of the DD genome (Li *et al*., 2001). Despite the diploid DD genome species is still undiscovered, at least discovering its progenitor from an extinct or extant species is still on the way (Vaughan *et al*., 2003).

In the *O. officinalis* complex, the only *Oryza* species with unique EE genome type is the *Oryza australiensis* Domin. (Ii *et al*., 1963). The unique EE genome is supposed to be the origin of DD genome (Ge *et al*., 1999). This Australian *Oryza* species is also different from other species of the genus by its strongly rhizomatous habit (Vaughan, 1988). Despite *O. australiensis* is genetically the most divergent species of the *O. officinalis* complex, it is the least studied species (Ii *et al*., 1963). However, this species contrasts with the diploid and single AA genome containing species of the *O. sativa* complex (Vaughan *et al*., 2008).

2.3.2 Phylogeny of the Genus *Oryza*

Since the moment of separation between *Oryza* and *Leersia* (14.2 MYA) (Guo and Ge, 2005), *Oryza* species became more diversified in New Guinea than any other geographic region (Vaughan, 1991). Besides, wide distribution of the closest grass relatives such as; Flagellariaceae, Joinvilleaceae and Ecdeiocoleaceae to the primitive species complex made the Australasian region as center of Origin for *Oryza* (Watson and Dallwitz, 1992). Distribution and diversification of current *Oryza* species and their grass relatives implicated Australia as the most probable region for the evolution of *O. schlechteri* and other diploid H, J and K *Oryza* species (Vaughan, 1991). Distribution, morphology and habitat studies suggested the more likely descendant of *O. schlechteri* from an early lineage in the genus *Oryza* (Vaughan *et al*., 2008).

Many morphology based taxonomic studies and other molecular techniques suggested *O. brachyantha* as a distinct and most closely related *Oryza* species to the genus *Leersia* (Zhang and Kochert, 1998; Aggarwal *et al*., 1999, Joshi *et al*., 2000; Vaughan 1988; Launert, 1965). However, common existence of a tourist element class (Tourist-olo9) in the FF genome *O. brachyantha* and AA genome *O. longistamminata* was interpreted as a close relatedness between the two genomes (Zhang and Kochert, 1998). The other two ancient species complexes; *O. ridleyi* and *O. granulata* are also grouped to the 'primitive' or 'ancestral' section of the genus called *Padia* (Sharma, 2003). Recently, molecular clock data suggested divergence of the *O. granulata* complex from other *Oryza* at about 8 MYA (Guo and Ge, 2005).

In the other section, taxonomists categorized the two most recent lineages of *Oryza* called *O. officinalis* and *O. sativa* complexes (Sharma, 2003). Some DNA sequence analysis explain evolution in the genus as expected from other types of biosystematics (Ishii and McCouch, 2000; Kanazawa *et al*., 2000). For instance, RFLP map comparison in between *O. sativa* and *O. officinalis* revealed high homosequentiality between the AA and CC genomes of rice (Jena *et al*., 1994). High degree of genome segment conservation between the CCDD genome of *O. latifolia* and AA genome of *O. sativa* was also reported based on comparative RFLP mapping (Huang and Kochert, 1994).

2.3.3 Biogeography of *Oryza* species

In the tribe Oryzeae, the three genera; *Leersia*, *Luziola* and *Oryza* have distinct pattern of biogeographic distribution (Vaughan *et al*., 2008). In particular, the genus *Oryza* is specifically diversified in Australasia (Vaughan *et al*., 2005a). The antiquity of *Oryza* can be reflected by the distribution of all the four *Oryza* species complexes (*O. granulata, O. ridleyi*, *O. officinalis* and *O. sativa*) in between the two sides of the biogeographical boundary called the Wallace Line (Table 3). The ancient taxonomic section *Padia* which contains the *O. granulate* and *O. ridleyi* complex is distributed in Asia, Australia, New Caledonia and Papua New Guinea (Vaughan *et al*., 2008). The few existing herbarium specimens also recorded the northeast New Guinea as a home for *O. schlechteri* (Vaughan, 1988). The phylogenetic distribution of functional traits suggested shaded, disturbed-forest understory and/or shady stream sides as ideal ecological settings of early angiosperms (Feild *et al*., 2003). Such localities are precisely alike with the habitat where *O. schlechteri* grown or the Finisterre Mountains of Papua New Guinea (Vaughan *et al*., 2008). Before its extinction, *O. schlechteri* were highly abundant as a pioneer community in shady stream sides (Hovius *et al*., 1998). The other taxa with no species complex called *O. brachyantha*, is widely distributed across the continent Africa (Vaughan, 1988). *O. brachyantha* grows well in the Sahel zone (Vaughan *et al*., 2008). This species shares the small and temporary pools of East Africa with *O. barthii* (Chevalier, 1932).

The very confusing species of the group *O. granulata* complex are adapted to mountainous and non-flooded forests (Watt, 1891). The species complex is diversified across the insular Southeast Asia, particularly Java and Indonesia (Vaughan *et al*., 2008). *O. granulata* has been collected from Nepal, Northern Thailand and South India (Vaughan, 1988). The two members of the *O. granulata* species complex grow in the shade or partial shade areas of the degraded, primary or well-established secondary forests (Watt, 1891). *Oryza meyeriana* (Zoll. *et* Mor. *ex* Steud.) Baill., is specifically found in Southeast Asia. However, the most common and widely spread species of the complex called *Oryza granulata* Nees *et* Arn.*ex* Watt, grows in South Asia, Southeast Asia, and southwest China (Tateoka, 1962b).

Like the *O. granulata* group, the *O. ridleyi* complex is distributed across Southeast Asia and New Guinea (Table 3). The stolon possessed by species in the *O. ridleyi* complex helps their adaptation to swampy forest habitats (Vaughan *et al*., 2008). Both *O. longiglumis* and *O. ridleyi* usually grow in shaded areas near rivers, streams and pools (Vaughan, 1988). *O. longiglumis* Jansen is recorded only on few sites along the Koembe River, Irian, Jaya, Indonesia, and in Papua New Guinea (Duistermaat, 1987; Katayama, 1961). However, the species *O. ridleyi* distributes from the Southeast Asia to Papua New Guinea (Vaughan, 1988).

Table 3. Genome group, usual habitats and life cycles of the four *Oryza* species complexes (Source: Vaughan *et al*., 2008).

Characteristics	***O. granulate* complex**	***O. ridleyi* complex**	***O. officinalis* complex**	***O. sativa* complex**
Genome (s)	GG	HHJJ	BB, CC, BBCC, CCDD, EE	AA
Life cycle	Perennial	Perennial	Perennial	Both annual and perennial
Usual habitat conditions:				
(a) Elevation	Tropical uplands and lowlands	Tropical lowlands	Tropical lowlands	Tropical lowlands
(b) Associated main vegetation types	Tropical deciduous forests	Tropical rain-forests	Tropical deciduous forests and grasslands	
(c) Basic water requirements	Dry land	Seasonal wetlands	Seasonal wetlands	Seasonal and permanent wetlands
(d) Light requirements	Full to partial shade	Full to partial shade	Full or partial shade or full sun	Full sun

The most diversified *O. officinalis* complex has species groups in Asia, Africa and Latin America (Vaughan, 1988). The most common Asian species called *O. officinalis* is widely distributed in the South and Southeast Asia (Katayama, 1961, 1963). *O. officinalis* thrives in partial shade or fully bright condition (Vaughan, 1988). In Philippines, *O. officinalis* is found in sympatry with *O. minuta* (Katayama, 1961; Katayama, 1963). *O. minuta* usually grows in shade or partial shade edges of a stream (Vaughan, 1980). From its very few populations, the distribution of the third Asian species of the complex called *O. malampuzhaensis* is localized to South India (Krishnaswamy and Chandrasekharan, 1957). However, Vaughan and Muralidharan (1989) further reported presence of the *O. malampuzhaensis* populations in marshy area of a partially shaded forest floor of tea plantation in Latin America.

From the three American species of the *O. officinalis* complex, *O. latifolia* is widely distributed and growing throughout the Central America, South America and Caribbean islands (Vaughan, 1988). Whereas, the other two species called *O. alta* and *O. grandiglumis* are restricted only to South America in particular to the Amazon basin (Oka, 1961).

The African continent is reported as the home for the two species of the *O. officinalis* complex called *O. punctata* and *O. eichingeri* (Hu, 1970; Tateoka, 1965). However, *O. australiensis* is the only Australian species in the complex (Tateoka, 1962a; Second, 1985a). The species *O. australiensis* inhabits the northern tropical zone of Queensland, Northern Territory, and Western Australia alongside with the other AA genome *O. meridionalis* (Vaughan, 1988). Such coexistence of *O. australiensis* and *O. meridionalis* verified Australia as center of origin of both *O. officinalis* and *O. sativa* species complex (Vaughan *et al*., 2008). Climate change and isolation after distribution resulted in the currently existing and diversified ecogeographic races of both the *O. sativa* and *O. officinalis* complexes (Second, 1991).

2.4 The *Oryza sativa* species complex

The eight taxa of *Oryza sativa* species complex have the same AA genome (Table 2). However, there are genomic differences among species inhabiting the three continents (IRRI, 1964). Some taxonomists symbolize such genomic differences using letters like; g, l, m and gp as a superscript to the AA designation (Vaughan, 1988).

Africa consists of the West African cultivated rice, *O. glaberrima* (Oka, 1977), its annual wild relative called *O. barthii* Chev. (Clayton, 1968) and the strongly rhizomatous and perennial *O. longistaminata* (Cope, 2008). The African cultivated rice, *O. glaberrima* is distinguished from its Asian counterpart, *O. sativa* by its short and rounded ligule, panicle devoid of secondary branches and smooth lemma and palea (Vaughan, 1988). The sub-genomic differences among *O. glaberrima* and *O. sativa* cause hybrid sterility (Nayar, 1973). However, the two African wild species, *O. longistaminata* and *O. barthii* can be hybridized with cultivated rice species, *O. glaberrima* and *O. sativa* (Bezancon *et al* 1977; Ghesquiere, 1985).

In Asia, *O. sativa* and its AA genome wild relatives, *O. nivara* and *O. rufipogon* were grouped in an independent species complex (Morishimia *et al*., 1961; Oka, 1956). For instance, Tateoka (1963a) used the name *O. rufipogon* to refer to any Asian and America wild relatives of rice. However, Sharma and Shastry (1965) split *O. rufipogon* of Asia into a new annual species, *Oryza nivara* Sharma *et* Shastry and perennial wild rice, *O. rufipogon*. Wild rice collections from the cost of Orissa, west Bengal and from the Deccan plateau of India also indicated the two distinct but close relatives of rice (Vaughan, 1988). The difference among these two wild species of rice can be explained by the high seed production of *O. nivara* and the low seed production by *O. rufipogon* (Harper, 1977).

The *O. sativa* complex also includes two other species called the *O. meredionalis* Ng and *O. australiensis* (Table 2). The Australian *O. meredionalis* has many morphological similarities with that of *O. nivara* (Sharma and Shastry, 1965). However *O. meredionalis* possess longer awns, narrower spikelets and a more compact panicle (Vaughan, 1988). To the contrary, the only AA genome Latin American species, *O. glumaepatula* Steud., is closely related to the other wild rice *O. rufipogon* (Vaughan, 1988). Second (1985b) defined *O. glumaepatula* as a recent introduction or an escaped form of the cultivated rice. However, *O. glumaepatula* is perennial not annual as *O sativa* (Vaughan, 1988).

2.4.1 Phylogenetic relatedness among the AA genome *Oryza* species

Evidence showed that the *O. sativa* complex is one of the most recently diverged lineages within *Oryza* (Sharma, 2003). Geographical isolation among AA genome *Oryza* species by an open ocean implies contribution of oceanic dispersal for their evolution and divergence (Ge *et al.*, 2005). Presence of the two cultivated rice and their domestication from wild species in the *O. sativa* complex makes evolution of the group as a considerable research topic (Sotowa *et al.*, 2013). Therefore, evolutionary relationships among AA genome *Oryza* species have been extensively investigated using morphological characteristics (Morishima and Oka, 1960), isozymes (Ishii *et al.*, 1996) and molecular markers (Doi *et al.*, 2000; Lu *et al.*, 2002). However, the information generated from those phylogenetic studies is still limited and sometimes contradictory (Duan *et al.*, 2007). In particular, such methods used in the phylogenetic reconstruction can't resolve the positions and relationships of *O. glumaepatula*, *O. longistaminata* and *O. meridionalis* in the AA-genome group (Ishii *et al.*, 1996). Moreover, low sequence divergence or recent radiation of the AA genome *Oryza* disabled the resolution capacity of the commonly used ITS and cpDNA sequences in the species complex (Ge *et al.*, 2005).

However, there are several studies reporting similar dates for the divergence of AA genome *Oryza* species (Vaughan *et al*., 2008). These molecular clock approaches estimate the first divergence between ancestors of the Australian AA genome species, *O. meridionalis*, and other AA genome *O. longistaminata* about 2 MYA (Ge *et al*., 2005). The same dates, 2–3 MYA, was also suggested for the first divergence of Asian–African AA genome species or between ancestors of *O. longistaminata* and Asian AA genome *Oryza* (Vitte and Panaud, 2003). Based on nuclear DNA sequence data, Ma and Bennetzen (2004) estimated 0.64 MYA as a date for the second divergence between Asian and African AA genome species or divergence between ancestors of *O. glaberrima* and *O. sativa*. Intron sequence data from three nuclear genes also reported a very similar dates (approximately 0.7 MYA) for the divergence between the Asian cultivated rice (*O. sativa*) and the African cultivated rice (*O. glaberrima*) (Zhu and Ge, 2005). Evidences also implicated separation of the two subspecies of *O. sativa* (*indica* and *japonica*) at approximately 0.4 MYA (Ge *et al*., 2005).

2.4.2 Biogeography of the *Oryza sativa* complex

Members of the *Oryza sativa* complex exhibit a remarkable eco-geographic differentiation (Vaughan *et al*., 2003). Distribution of such species in large geographic region is likely from both biotic and abiotic factors (Tiffney, 2004). Studies on AA genome species from a wide geographic area suggested Australasian region as their origin followed by the spreading and diversification to other regions (Vaughan *et al*., 2008).

Distinct ecogeographic races within the *O. sativa* complexes largely arose from the climatic change and isolation after distribution (Second, 1991). Currently, these species are widely distributed in Asia, Africa, Latin America and Oceania (Figure 6). For more details, *O. nivara* and *O. rufipogon* exist in Asia; Africa comprises the African cultivated rice, *Oryza* *glaberrima*, *Oryza barthii*, and *Oryza longistaminata*, Australia has *Oryza meridionalis*, and the *Oryza glumaepatula* exist in Latin America (Vaughan *et al*., 2005b).

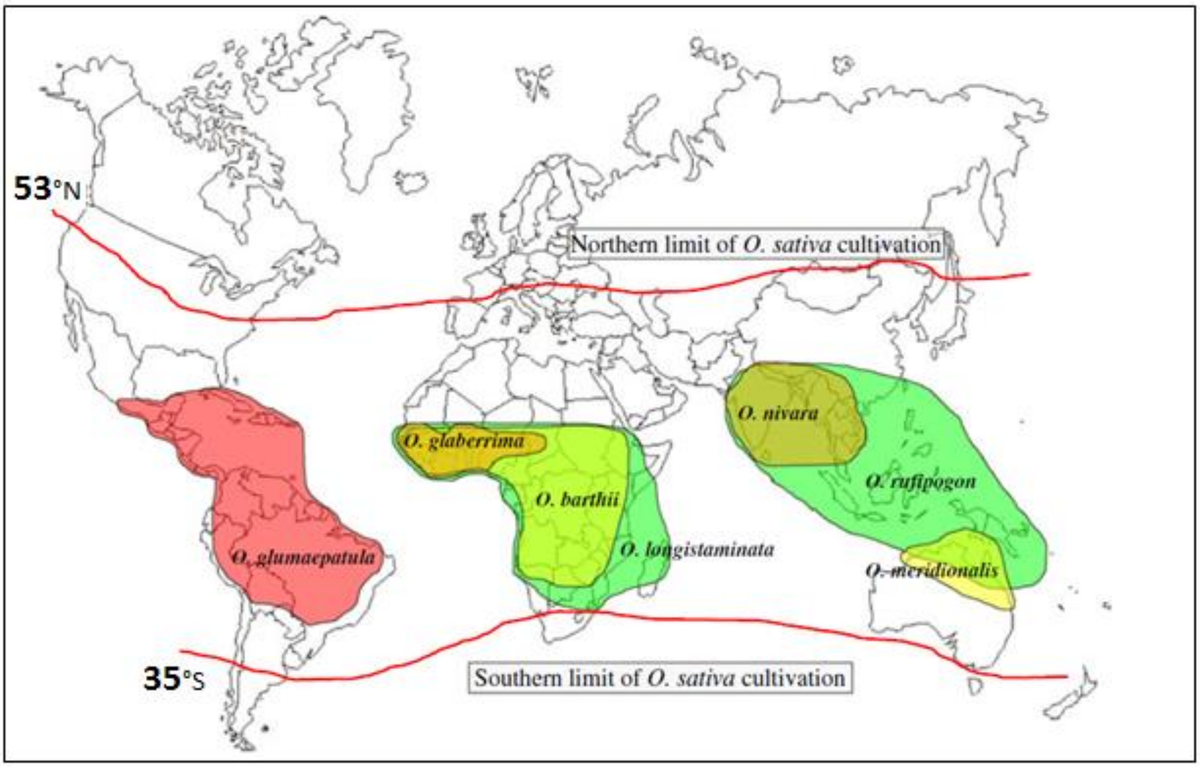

Figure 6. Worldwide geographic distribution of AA genome wild *Oryza* species (Source: Lu *et al*., 2010).

Biogeographic studies indicated presence of all AA genome *Oryza* species in East Africa (Vaughan *et al*., 2005b). However, environmental changes resulted in distinct ecogeographic races within the *O. sativa* complex (Second 1991). For instance, the African cultivated rice, *O. glaberrima*, is now mainly cultivated in the local agricultural ecosystems of West Africa (Chang, 1985; Khush, 1997). Besides, *O. barthii*, which is presumed as an annual ancestor of *O. glaberrima*, was widely distributed along the entire southern fringe of the Sahara Desert from Ethiopia to Senegal (Vaughan, 1994). However, this wild progenitor is currently restricted to West Africa (Morishima, 2000). Unlike the two African AA genome species, the

African wild rice *O. longistaminata* is looking adaptable to the climatic change and maintain its wide distribution from Ethiopia in the East to South Africa (Sotowa *et al*., 2013).

According to Akimoto *et al*. (1997), distribution of *O. glumaepatula* in the New World probably occurred later than in other areas. This regionally distinct Latin American form of the complex is widely distributed in a geographic boundery from 23° N in Cuba to 23° S in Brazil (Vaughan *et al.*, 2003). The Brazilian *O. glumaepatula* populations are genetically more diverse and widely distributed in the extensive river basins of the Amazon and Pantanal Matogrossense (Oliveira, 1994; Brondani *et al*., 2005).

The most diverged Australasian AA genome species, *O. meridionalis* is endemic to Oceania including New Guinea and Australia (Ng *et al*., 1981; Vaughan 1994; Lu, 1999). Haplotype studies of *O. rufipogon* suggested that this species is most ancient in India and Indochina (Londo *et al*., 2006). In New Guinea, highly diverse lineages of *O. meridionalis* are found with *O. rufipogon* accessions (Doi *et al*., 2000). However, distribution of *O. rufipogon* is not restricted and endemic to Oceania (Vaughan *et al.*, 2003). Rather, the perennial *O. rufipogon* is distributed from tropical Asia to Australia (Chang, 1976, Vaughan, 1994). Whereas, *O. nivara*, which is sometimes referred to as the annual form of *O. rufipogon* spreads from the Deccan Peninsula to the Gangetic valley (Sotowa *et al*., 2013). The current distribution of Australian *O. rufipogon* is quite broad as that of *O. meridionalis* (Henry *et al*., 2010). The Asian domesticated rice *O. sativa* is related to *O. rufipogon* and can exchange its genes by spontaneous hybridization (Oka and Chang 1959). Few records briefly indicated domestication of *O. sativa* in Australia (Waters *et al*., 2012). However, it is a common knowledge that Asian cultivated rice was domesticated in South and Southeast Asia and now cultivated worldwide (Duan *et al*., 2007).

2.5 Genetic resources of wild *Oryza* species

Though rice is cultivated worldwide under various agro-climatic conditions, its production in recent years has been seriously affected by many biotic and abiotic stresses (Normile, 2008). During the course of domestication from wild relatives, the cultivated rice was left with a significant loss of genetic diversity (Brar and Khush, 2003). Surprisingly, selection in favor of desirable agronomic traits reduced 50–60% alleles of the wild rice (Sun *et al*., 2001). Thus, wild species of *Oryza* must be exploited both to broaden the existing narrow genetic base and enrich the current varieties with desired agronomically important traits (Joshi *et al*., 2000).

In terms of agronomic traits, wild *Oryza* species are in fact phenotypically inferior (Ali *et al*., 2010). However, they exhibit tremendous diversity in morphological traits and adaptation to diverse habitats (Sanchez *et al*., 2014). Moreover, wild species of *Oryza* contains numerous genes of economic importance and are being used as alternate sources of resistance or tolerance to biotic and abiotic stresses (Table 4). Despite effective transfer of valuable qualitative and quantitative traits from the wild *Oryza* species would broaden the gene pool and improve productivity of the cultivated rice, most wild species have limited success of useful genes transfer under natural or biological means (Brar and Khush, 1986; Brar and Khush, 1997). Based on the ease of gene transfer from wild species into cultivated rice, gene pools of wild *Oryza* species were classified into three main groups called the primary or *O. sativa* complex, secondary or *O. officinalis* complex and tertiary *O. meyeriana* and *O. ridleyi* complex (Khush, 1997).

For many years, wide hybridization has been used to introduce qualitative characters such as disease and insect resistance, and male sterility from wild species into elite breeding materials (Brar and Khush, 1997; Dalmacio *et al*., 1995). However, several incompatibility barriers such as; low crossing success, increased sterility and limited recombination between the chromosomes of wild and cultivated species seriously hampered the task of useful genes transfer (Sitch, 1990; Brar and Khush, 1986; Brar and Khush, 1997). Recent advances in tissue culture, molecular marker technology, genomics and fluorescence *in situ* hybridization have actually opened new opportunities to tap alien genetic variability from distant *Oryza* genomes through interspecific hybridization (Sanchez *et al*., 2014).

Species in both *O. meyeriana* and *O. ridleyi* complex are highly cross incompatible with the cultivated species, *O. sativa* (Normile, 2008). However, such wild species possess very valuable genes such as; tolerance to shade and drought, salinity resistance, immune to bacterial blight and resistance to brown planthopper (Qian *et al*., 2006). In the near future, many important genes from such distant wild species are expected to be transferred into cultivated rice through the latest tools of Molecular Genetics and Biotechnology (Normile, 2008).

Table 4. Potentially useful traits of *Oryza* species (Source: Lu *et al*., 2010).

Species	**Potentially useful traits**
O. sativa	Cultigen
O. nivara	CMS, hybrid breakdown locus, pollen sterility locus and resistance to grassy stunt virus, blast, sheath blight, stem borer, whorl maggot and drought
O. rufipogon	Resistance to BB, stem rot, tungro virus, blast, stem borer and WBPH; elongation ability, tolerance to Aluminum and soil acidity, source of CMS, improved yield, salinity tolerance, fertility restoration ability
O. glaberima	Cultigen, resistance to yellow mottle virus, rice stripe necrosis virus and nematodes; tolerance to Aluminum, soil acidity and Iron toxicity; drought avoidance and source of CMS
O. barthi	Source of CMS, drought avoidance, Resistance to BB, BLS, blast, brown spot, sheath blight and GLH
O. glumaepatula	Elongation ability, source of CMS
O. longistaminata	Resistance to BB, blast, BPH, nematodes and yellow stem borer; drought avoidance, pollen/spikelet fertility
O. meridionalis	Elongation ability, drought avoidance and sheath blight
O. punctata	Resistance to BPH, zigzag leafhopper
O. minuta	Resistance to sheath blight, blast, BB, BPH and GLH; grain size, awn length, heading date, panicle number, spikelet number

Table 4 continued

O. eichingeri	Resistance to yellow mottle virus, BPH, WBPH and GLH
O. officinalis	Resistance to thrips, BPH, GLH and WBPH
O. rhizomatis	Drought avoidance, rhizomatous
O. alta	Resistance to striped stem borer, high biomass production
O. grandiglumis	High biomass production
O. latifolia	Resistance to BPH, GLH and WBPH
O. australiensis	Drought avoidance, Resistance to blast, BPH
O. brachyantha	Resistance to yellow stem borer, leaf-folder, whorl maggot and BB; tolerance to laterite soil
O. granulate	Shade tolerance, adaptation to aerobic soil
O. meyeriana	Shade tolerance, adaptation to aerobic soil
O. longiglumis	Resistance to blast, BB
O. ridleyi	Resistance to stem borer, whorl maggot, blast, BB

Where, BPH - brown plant hopper, GLH - green leaf hopper, WBPH - white-backed plant hopper, BB - bacterial blight; BLS - bacterial leaf streak, CMS - cytoplasmic male sterility

Likewise, members of the *O. officinalis* complex are also cross incompatible with the cultivated rice and showed non-homologous chromosome pairing (Sanchez *et al*., 2014). Even if it is not easy, genes for resistance to brown plant hopper, bacterial blight and blast have also been introgressed from such distant species of the *O. officinalis* complex through embryo rescue technique (Brar and Khush, 1997). Following the embryo rescue, several interspecific hybrids with a desirable trait have been produced (Brar *et al*., 1991; Jena and Khush, 1984; Shin and Katayama, 1979). For instance, Jena and Khush (1990) produced several introgression lines with brown planthopper (BPH), white backed plant hopper (WBPH) and bacterial blight (BB) resistance from the cross between *O. sativa* and *O. officinalis*. More recently, successful introgression of traits was possible from *O. minuta* and *O. australiensis* for the transfer of blast resistance to *O. sativa* (Brar and Khush, 2002; Fu *et al*., 2008; Jena and Khush, 2000). Several introgression lines from *O. latifolia* was also reported for resistance to BPH, WBPH and BB and other yield enhancing traits such as growth duration and purple pigmentation (Sanchez *et al*., 2014). The other diploid and CC-genome species, *O. eichingeri* has also been used in a rice breeding program as a source of resistance to brown plant hopper (Yan *et al*., 1997).

Of the wild relatives, species containing the AA genome are taken as part of the primary gene pool of rice (Ge *et al*., 2005). Those members of the *Oryza sativa* species complex are diploid, cross compatible and show homologous chromosome pairing with the cultivated rice (Jena, 2010). Under conventional breeding methods, wild AA genome *Oryza* species can be easily crossed with the cultivated rice and transfer valuable genes for rice improvement (Lu *et al*., 2000). Such AA genome wild *Oryza* species can be sources of cytoplasmic male sterility and provide useful genes such as resistance to grassy stunt, tungro viruses and bacterial blight (Brar and Khush, 1997). Some of the breeding lines with genes introgressed from wild AA genome *Oryza* species have been released as varieties (Sanchez *et al*., 2014).

The first report on transfer of an agronomically important gene from a wild species to cultivated rice was the introgression of grassy stunt virus resistance from *O*. *nivara* (Khush *et al*., 1977). One of the closest relatives of cultivated rice called *O*. *rufipogon* was also found to contain genes that can substantially increase the yield of rice (Xiao *et al*., 1996). Moreover, broad spectrum of blast resistance were transferred from *O*. *rufipogon* and released as a variety (Ram *et al*., 2007a; Ram *et al*., 2007b). In particular to the *Indica* subspecies, genes for tungro virus resistance and tolerance to acid sulfate soil conditions have been transferred from *O*. *rufipogon* (Sanchez *et al*., 2014).

Although *O*. *glaberrima* is a low yielding cultivated species, it bears numerous interesting genes for the improvement of *O*. *sativa* (Ali *et al*., 2010). Some of the desirable traits of this African cultivated rice include resistance to nematodes, insects, rice yellow mottle virus, rice stripe necrosis virus and tolerance to drought acidity and iron toxicity (Sanchez *et al*., 2014). Therefore, combining the high productivity of *O*. *sativa* with the strong biotic and abiotic stress tolerance of *O*. *glaberrima* offers a tremendous potential in rice breeding (Ali *et al*., 2010). The AA genome species *O*. *glumaepatula*, is the source of cytoplasmic male sterility

(CMS) and development of CMS lines for commercial scale hybrid rice production (Lin and Yuan, 1980).

In addition, the first cloned rice disease resistance genes *Xa-21*, which encodes a protein with unusual leucine rich repeat (LRR)-kinase domains (Song *et al*., 1995), was introgressed from the African wild rice (*Oryza longistaminata* Chev. et Roehr.) (Qian *et al*., 2006). This gene which confers broad-spectrum of resistance to rice bacterial blight disease has been widely applied to rice breeding and introduced into many rice cultivars (Khush *et al*., 1990). Using BAC library and flourescence *in-situ* hybridization, the *Xa-21* gene has been cloned and physically mapped on chromosome 11 of rice (Brar and Khush, 1997).

2.6 The African wild rice (*O. longistaminata*)

The African wild rice (*O. longistaminata*) is a perennial and allogamous African *Oryza* species (Chev and Roher, 1914; Khush, 1997; Sacks *et al*., 2006). This undomesticated rice species is morphologically characterized by the very long ligule with an erect, soft and spongy stem (Cope, 2008). The very long, branched and extensive rhizome at the lower nodes of *O. longistaminata* is the means for its perenniallity (Tao *et al*., 2000). The strongly perennial and vigorously rhizomatous feature of *O. longistaminata* enables its extensive distribution throughout Africa (Figure 1). As reported by Matsuo *et al*. (1997), *O. longistaminata* inhabited regions from northern Uganda 3^{0}38'N; 15^{0}34'E to South Africa 26^{0}17'S; 48^{0}58'E with an altitudinal range from 400m a.s.l. to 1800 m a.s.l. However, the distribution further extends to Ethiopia (N 11°97’86’’ and 37°70’72’’E) (Melaku *et al*., 2013).

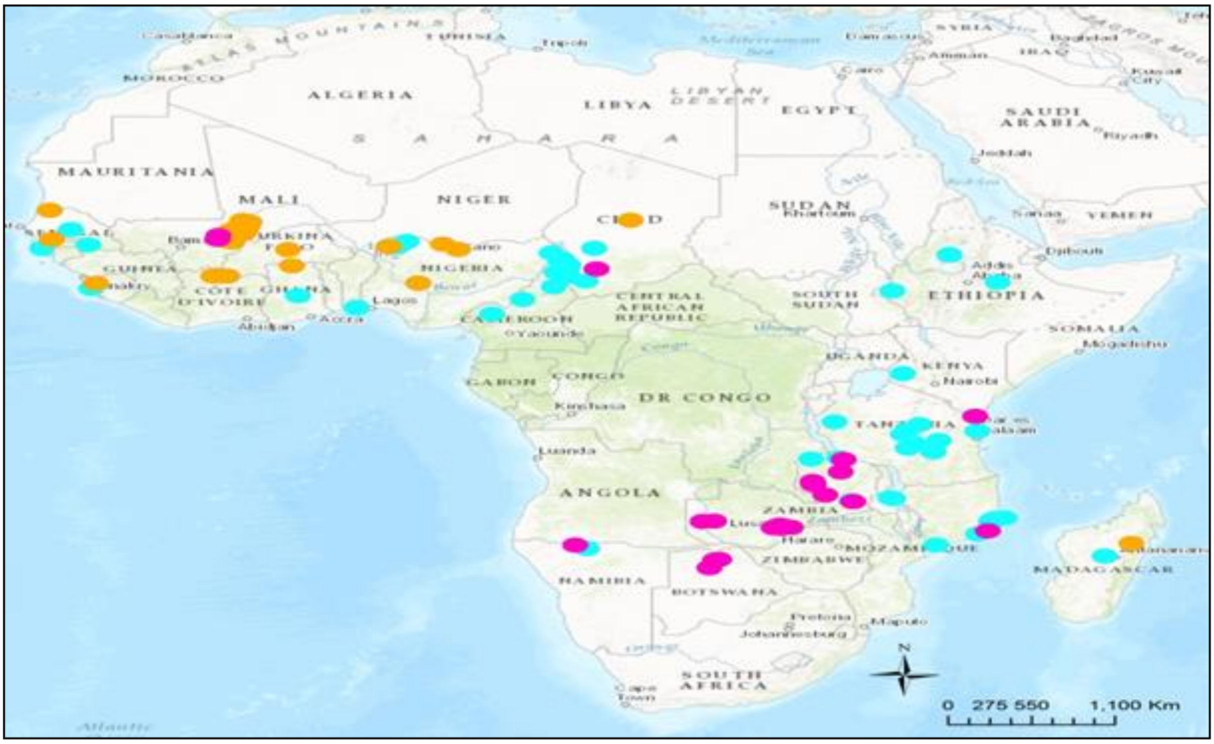

Figure 7. Collection sites of *O. longistaminata* accessions that are conserved at IRRI (Source: Marlee, 2016).

In the Niger Valley of Mali, *O. longistaminata* was sometimes used to supplement cultivated rice as a food grain (Nyoka, 1983). In other African countries such as Sudan and Ethiopia this undomesticated rice species was used as famine food (Vaughan and Sitch, 1991). Other than serving as cereals for human food and beverage, *O. longistaminata* is a good source of forage (Hu *et al*., 2003). Its strong rhizome provides high nitrogen use efficiency, high biomass production under poor soil condition and prevention of soil erosion (Song *et al*., 1995; Yang *et al*., 2010). Therefore, *O. longistaminata* can be the best donor of perennial traits and a useful resource for the development of permanent ground cover and perennial rice cultivars (Sacks *et al*., 2003). One potential strategy for the arousal of perennial rice is introgression of genes for perennial growth and rhizome production from *O. longistaminata* into the current annual rice cultivars (Sacks *et al*., 2006).

Accessions of *O. longistaminata* were reported as immune to rice yellow mottle sobemo virus, resistant to bacterial blight (*Xanthomonas oryzae* pv. *oryzae*) and rice blast (*Magnaporthe grisea*) (Vales, 1985; Thottapilly and Rossel, 1993; Zhang *et al*., 1994). Along with the disease resistance gene, *Xa-21*, high pollen production, long stigmas and drought tolerant traits of *O. longistaminata* were transferred to the cultivated rice (Tao *et al*., 2000; Khush *et al*., 1990). Such useful gene transfer tasks from *O. longistaminata* to *O. sativa* can be accomplished through conventional hybridization, selection and backcrossing procedures (Brar and Khush, 1997). However, effective introgression is typically hampered by embryo abortion at the initial cross and in the early backcross generations (Oka, 1988). The problem associated with hybrid sterility is, therefore, the major cause for the less utilization of *O. longistaminata* in different rice improvement programs (Chen *et al*., 2009).

Despite of the excellent gene pool from *O. longistaminata*, its presence in a rice farm bears a competitive effect on the cultivated rice and serves as an alternate host for some important rice pests and pathogens including rice yellow mottle sobemovirus (John *et al*., 1984) and rice bacterial blight (*Xanthomonas oryzae* pv. oryzae) (Buddenhagen, 1982). In some cases, its seed may be contaminated with the commercial rice grain and reduce quality (Parker and Dean, 1976).

Wide range of wetland areas of Gambella, Amhara and Illubabor regions of Ethiopia are habitats for the highly diversified *O. longistaminata* (Phillips, 1995). According to Melaku *et al*., (2013), such Ethiopian *O. longistaminata* accessions are rich in genetic diversity. The study from Marlee (2016), implicated the genetic distinctness of the Ethiopian *O. longistaminata* accessions. However, very few collections are found in gene banks like the International Rice Research Institute (IRRI) (Figure 2). Besides, local consideration of *O. longistaminata* as noxious weed in a rice farm and wetland areas, climatic change, overgrazing and human population pressure are affecting its genetic diversity in Ethiopia (Khush, 1997; Melaku *et al*., 2013). Therefore, country wise germplasm collection of *O. longistaminata* and assessment of their morphological features, photosynthetic potential, extent of genetic diversity, genetic relationship with other *Oryza* species and population structure will contribute a lot for their conservation and future uses in a multidisciplinary rice breeding efforts.

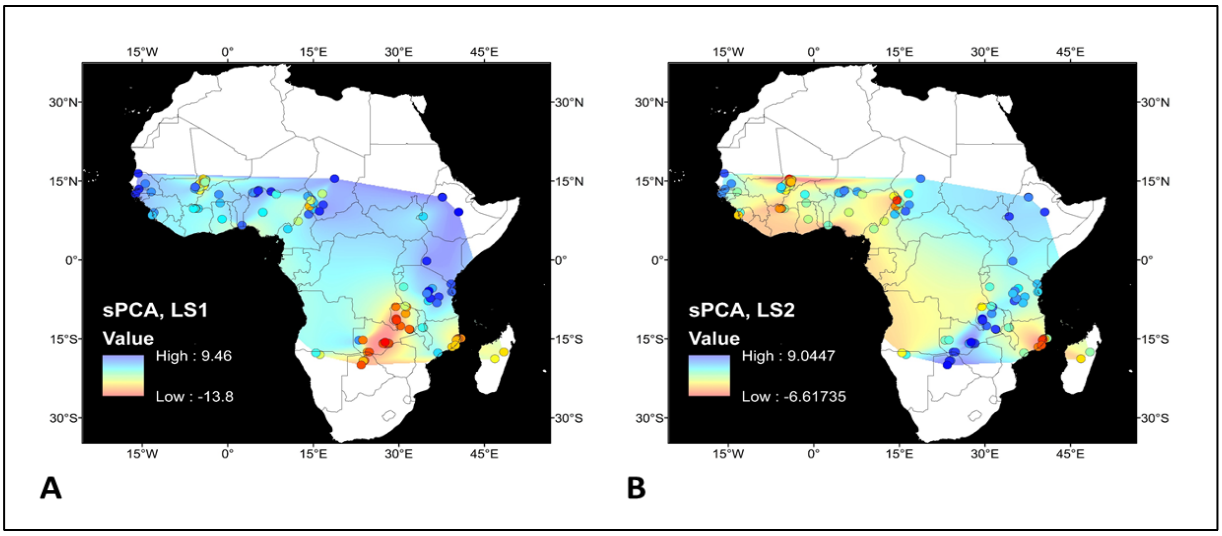

Figure 8. Spatial principal component analysis depicting genetic distinctness of the Ethiopian *O. longistaminata* samples (Source: Marlee, 2016).

2.7 Morphological and photosynthetic characterization of wild rice

Undomesticated species of the genus *Oryza* are valuable repository of useful genetic resources and might improve rice yield in diverse environmental conditions (Chouhan *et al*., 2014; Giuliani *et al*., 2013). Such undomesticated rice species further exhibited a wide range of phenotypes (Voughan, 1994; Vaughan *et al*., 2008; Brar and Singh, 2011; Jagadish *et al*., 2011). The highly diverse morphological characteristics (Tanksley and McCouch, 1997; Swamy and Sarla, 2008) and biomass level (Voughan, 1994) made both undomesticated and cultivated rice species as useful subjects of morphological diversity, photosynthesis and water use efficiency studies (Giuliani *et al*., 2013).

Traditionally, both qualitative and quantitative agro-morphological traits have been used to estimate genetic relationships among crops (Goodman, 1972). In particular, rice species were mainly described based on their morphological characteristics (Chang and Bardenas, 1965) and some biochemical traits (Glaszmann, 1987). Currently, high yield potential of rice is associated with cultivars commonly possessing effective plant morphology and higher leaf photosynthetic rate (Pn) (Horie *et al*., 2003; Takai *et al*., 2006).

Multivariate methods for handling morphological variation in rice germplasm collections have been demonstrated by Kanwal *et al*. (1983). Such methods are actually useful for the characterization, evaluation and classification of genetic resources for several agronomic and physiological importance (Peeters and Martinelli, 1989). The information generated from such multivariate analysis of morphological traits can be useful for revealing patterns of variation in the germplasm collections along with the identification and crossing of groups of accessions with a trait of agronomic interest (Brown, 1991; Perry and McIntosh, 1991; Souza and Sorrells, 1991; Cowen and Frey, 1987; Camussi *et al*., 1985).

According to the suggestions of Ohno (1976) and Cao *et al*. (2001), selecting accessions with a high photosynthetic rate (Pn) from significantly different rice varieties is one of the means to come up with high-yielding varieties. Modifying the primary photosynthetic pathway and the process of biomass production under a changing climate are, therefore, important for sustaining global food production (Kajala *et al*., 2011; Zhu *et al*., 2008).

Unlike their cultivated counterparts, wild rice species are known for their high biomass formation at a less nutrient requirement (Vaughan, 1989). This realization has renewed interest of utilizing photosynthetic genetic resources of wild species (Ding *et al*., 2014).

However, the mechanism of variation in photosynthetic rate is complex and associated with different structural and functional parameters of leaf (Ohsumi *et al.*, 2007). For instance, photosynthetic rate under current atmospheric CO_2 concentration is limited by carboxylation efficiency which further relies on the amount of Ribulose-1,5-bisphosphate and its kinetics (Farquhar and Sharkey 1982; Makino *et al.*, 1985). Pn is significantly correlated with stomatal conductance (Hubart *et al.*, 2007; Hirasawa *et al.*, 2010; Ohsumi *et al.*, 2007b). The correlation between Pn and gs is always higher than that between Pn and CE (Giuliani *et al.*, 2013). However, it is unclear that which trait would be more important for breeding crops with high photosynthetic rate (Ding *et al.*, 2014).

Many studies on historical varieties of different crop species revealed the influence of Pn on yield potential (Hubart *et al.*, 1966; Peng *et al.*, 2000; Zheng *et al.*, 2011; Gutierrez-Rodriguez *et al.*, 2000; Peng *et al.*, 2012). A study by Kondamudi *et al.* (2016), indicated that almost 80 % of the investigated undomesticated rice accessions had higher mean with a wider range of Pn than cultivated rice. This observation perhaps creates a scope of enhancing photosynthetic rate of cultivated rice using its undomesticated relatives (Ding *et al.*, 2014). Many other previous studies further reported the existence of large genotypic variation on the photosynthetic rate of rice and its relatives (Horie *et al.*, 2003; Cook and Evans, 1983; Yeo, 1994). For instance, Zhao *et al.* (2003) showed feasibility of selecting lines from interspecific crosses of cultivated and undomesticated rice for high Pn. The wild species *O. australiensis*, *O. rufipogon* and *O. nivara* in particular were reported as potential sources of enhanced Pn via introgression into the cultivated rice gene pool (Zhao *et al.*, 2003; Zhao *et al.*, 2010; Kiran *et al.*, 2013).

More interestingly, African wild rice (*O. longistaminata*) accessions showed superior photosynthetic rates and a 1.3 fold of the cultivars photosynthetic rate to dark respiration (Rd) difference (Kondamudi *et al*., 2016). Such an extremely efficient carbon balance system in *O. longistaminata* accessions must be assessed with the corresponding morphological traits like; size and shape of flag leaf and physiological factors (Yue *et al*., 2006), stomatal conductance (Maruyama and Tajima, 1990; Ohsumi *et al*., 2007a), transpiration rate (Teng *et al*., 2004) and an immense pool of useful genes (Cissé and Khouma, 2016).

Ethiopia, which is a center of genetic diversity for many plants (Mengesha *et al*., 1975; Harlan, 1969; Vavilov, 1951), is a potential hot spot of diversity in the *O. longistaminata* gene pool (Melaku *et al*., 2013). However, the morphological variability and photosynthetic related traits of *O. longistaminata* in Ethiopia and the world collections in general is not well characterized. Future role of *O. longistaminata* in rice breeding programs for the development of high yielding rice cultivar needs prior morphological and photosynthetic characterization and assessment of its genetic diversity and relatedness with other *Oryza* species.

2.8 Genetic diversity of *O. longistaminata* and its relatedness with other rice types

These days, the limited genetic variation for stress-related traits in cultivated rice affects its productivity (Ali *et al*., 2010). Thus, there must be an urgent need to explore diverse sources of genes for tolerance to various stress factors (Joshi *et al*., 2000). Wild *Oryza* species are important sources of agronomically useful traits and could be used extensively in rice improvement programs (Brar and Khush, 1997). In spite of their rich genetic diversity, few of the wild *Oryza* species have been utilized in rice breeding programs (Das et al. 2013). The vast majority of wild rice genes remain untapped to date, as it is often difficult to transfer these genes into the cultivated rice (Ren et al. 2003). Therefore, precise genetic manipulation of desirable traits should start from a prior and thorough assessment of genetic diversity, population structure and kinship relatedness of both cultivated rice and its wild relatives (Nachimuthu *et al*., 2015).

The African wild rice (*O. longistaminata* A. Chev. et Roehr.) which shares the same AA genome as *O. sativa* is the most accessible genetic resource for rice improvement (Vaughan *et al*., 2005; Lu *et al*., 2000; Ren *et al*., 2003). In addition to its high rate of self-incompatibility, morphological features such as high pollen production and long, exserted stigma presumably facilitate outcrossing (Ghesquiere, 1987; Jones *et al*., 1996). Therefore, genetically diverse *O. longistaminata* accessions in each natural population are expected from the high gene flow (Melaku *et al*., 2013).

In Ethiopia, the different agro-ecologies with unique environmental and climatic features resulted in wide genepool of *O. longistaminata* (Melaku, 2011). However, the different ecological threats such as population pressure, climatic change, expansion of rice farming and overgrazing made this African species as an important focus for conservation priority (Melaku *et al*., 2013; Wambugu *et al*., 2013). Though understanding the extent of genetic diversity is very basic for the task of conservation, adequate attention is not given (Aliyu *et al*., 2013). Therefore, assessing the genetic variability within and among *O. longistaminata* populations is important for its conservation and further utilization in rice breeding.

Most DNA marker based genetic relationship studies of the genus *Oryza* or taxa in the AA gene pool supported the morphological based *Oryza* classification (Duan *et al*., 2007). However, the different marker types used in the relationship studies of the whole *Oryza* species revealed controversial findings (Ge *et al*., 2004). In particular, AA-genome *Oryza* species relationship studies are limited and sometimes ambiguous (Duan *et al*., 2007). The self-incompatible, rhizomatous, and unique ligular characteristics made the African *O. longistaminata* different in its AA-genome group (Ghesquiere, 1985; Vaughan, 1994).

However, the genetic relationship between *O. longistaminata* and other AA-genome *Oryza* species is still unresolved (Ren *et al*., 2003).

The more complicated genetic variation observed in the AA-genome *Oryza* and their weak interspecific relationships could be associated with the marker type, sampling methods of the taxa and polytony in the clade (Duan *et al*., 2007). Thus, adequate number of accessions and SSR primers along with Genome Based Sequencing (GBS) markers must be utilized for a reliable genetic structure analysis and better resolution over related species (Doyle *et al*., 1998).

2.9 Molecular markers

Understanding the extent of genetic variations and relatedness among wild and cultivated rice species are important considerations to design effective breeding programmes (Judith, 2015). Molecular marker analysis is one of the most useful methods of investigating the genetic diversity of wild rice (Wang *et al*., 2008; Zhou *et al*., 2003; Gao, 2004). In wild rice populations, molecular markers have become fundamental tools for assessing genetic variations, establishing phylogenetics, understanding population structure, tagging desirable genes, resolving cultivar identities and study of genome organization (Dinesh and Kiran, 2010; Sun and Yang, 2009).

Broad ranges of molecular markers are available and used in a variety of ways (Rahman *et al*., 2010). Detection of genetic variations at molecular level can be facilitated by Random Amplified Polymorphic DNA (RAPD) (Park *et al*., 2009), Amplified Fragment Length Polymorphism (AFLP) (Maheswaran, 2004), Restriction Fragment Length Polymorphism (RFLP), Inter Simple Sequence Repeats (ISSR), Simple Sequence Repeats (SSR) (Judith, 2015), and next generation sequencing approaches (Ren *et al*., 2003).

2.9.1 Microsatellite (SSR) markers

Microsatellites or Simple Sequence Repeats are short tandem repeats which are interspersed throughout the genome (Park *et al.*, 2009). SSRs are highly informative, abundant, co-dominant, cost effective and highly reproducible marker types exhibiting high degree of allelic variation at a locus (Panaud *et al.*, 1996; Temnykh *et al.*, 2000). SSRs are the best molecular marker systems for many types of genetic analyses including germplasm surveys, linkage mapping and phylogenetic studies (Ghneim *et al.*, 2008). The extent of genetic diversity and relatedness in different wild rice populations are commonly estimated by microsatellite markers (Ngu *et al.*, 2010). In rice, about 50% of the genome consists of repetitive DNA sequences (Deshpande and Ranjekar, 1980). Since establishment of the first molecular map of rice genome in 1988, McCouch *et al.*, (1997, 2002) identified and mapped 2740 SSRs on the 12 rice chromosomes.

2.9.2 RAD-Seq based genome wide association studies

Next-generation sequencing technology provides large scale molecular marker data for resolving genetic diversity and refined population structure analyses (Ren *et al.*, 2003). A form of genome reduction strategy called restriction site associated DNA sequencing (RADSeq) creates a reduced representation of the genome and identify thousands of genetic markers (Baird *et al.*, 2008; Miller *et al.*, 2007). Despite RADSeq targets subset of a genome, its advantage over the whole-genome sequencing lies on providing greater depth of coverage per locus and generating sequences of higher numbers of samples for a given budget (Emerson *et al.*, 2010). RADseq does not require any prior genetic information and development of genomic resources (Andrews *et al.*, 2016). Rather, its system often identifies SNPs through the disruption of restriction endonuclease recognition sites (Botstein *et al.*, 1980, Vos *et al.*, 1995; Wenzi *et al.*, 2004).

Using diverse Bioinformatic tools like Stacks, the massively parallel and multiplexed sequences of RAD tag libraries facilitate the rapid discovery of thousands of SNPs (Catchen *et al*., 2011; Baird *et al*., 2008; Catchen *et al*., 2013a). These days, those genome-wide SNPs are successfully applied in population genomics, phylogenetics and speciation studies (Pan *et al*., 2016).

3. Materials and Methods

3.1 Plant Material and Sample Collection

From a field survey conducted in October 2015, both seed and leaf materials used in this study were sampled from the different populations of both Amhara and Gambella regions of Ethiopia (Figure 9). Sampling of the Ethiopian *O. longistaminata* individuals here after called accessions was done randomly with an approximate distance of 10 m from each other. Till the transportation to Yunnan University, Kunming, China, both germplasms and Silica gel dried leaves were stored at room temperature in the Plant Molecular Biology Laboratory of Addis Ababa University.

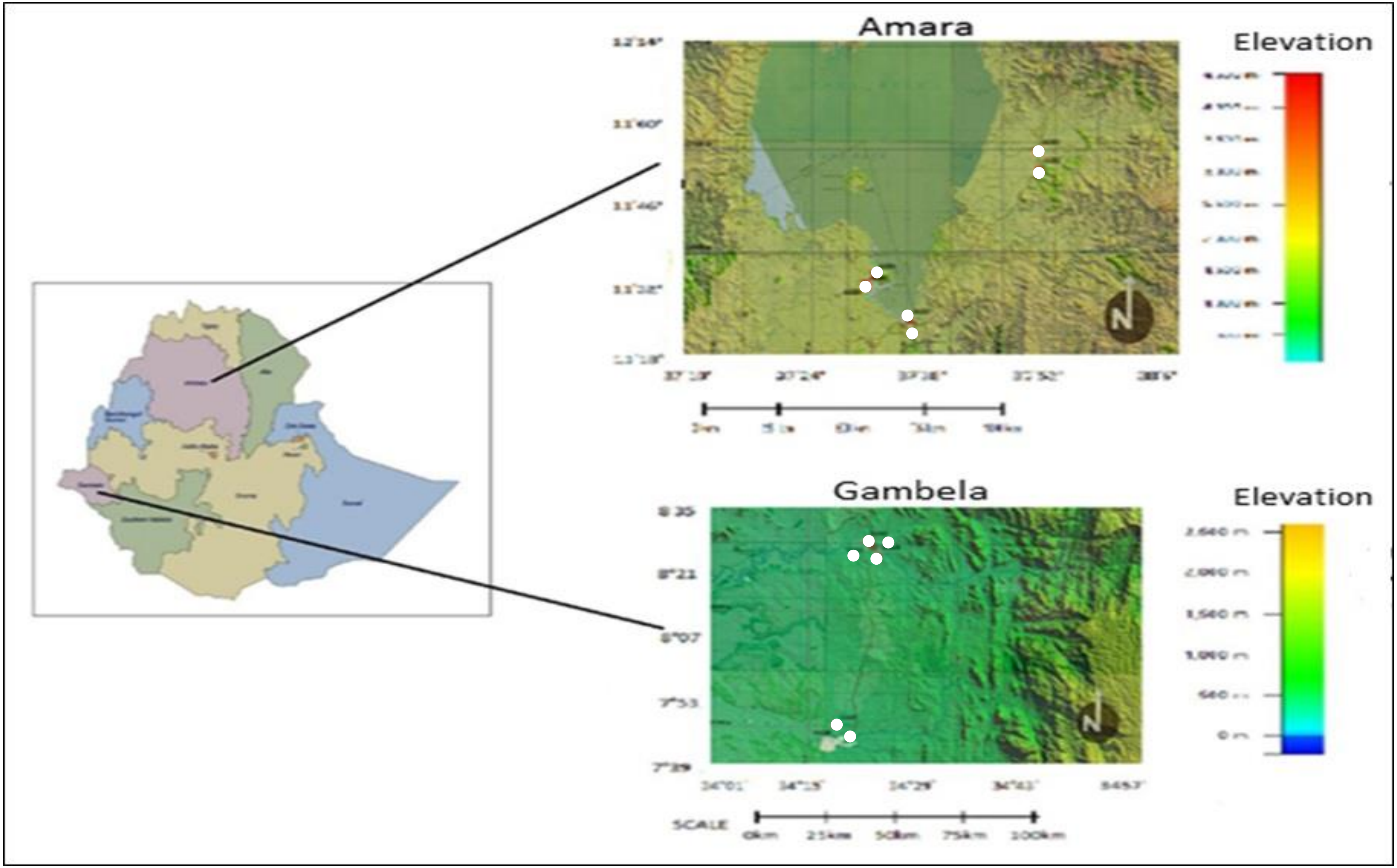

Figure 9. Site maps for the Ethiopian *O. longistaminata* collections (White dots designated locations of each population)

3.2 Morphological and photosynthetic characterization of the Ethiopian *O. longistaminata* accessions

3.2.1 Planting materials and experimental design

Seeds from 62 *O. longistaminata* accessions were collected from six populations of Ethiopia (Table 5). Dehulled and surface sterilized seeds were grown in MS media at the School of Agriculture, Yunnan University (YNU) as described by Boissot *et al*., (1990). In the wet season of June 2016, the seedlings were transplanted into a field with a loam soil at JingHong Rice Breeding Station, Gasha Town, Yunnan Province, China. The accessions were laid out in a randomized complete block design with three replications. A total of five seedlings (plants) per plot with a spacing of 40cm between rows and 20cm between plants in a row were grown. All replications were fertilized with 4 g/m^2 of nitrogen and 12 g/m^2 of phosphorus and potassium basally and top-dressed with 2 g/m^2 of nitrogen every 20 days (Ohsumi *et al*., 2007b).

Table 5. Populations of the 62 *O. longistaminata* accessions with their code and GPS locations

Population	Population code	No of accessions/population	Latitude	Longitude	Altitude (m)
Gambella-1	G1	11	8°25’29’’	34°57’60’’	437
Gambella-2	G2	11	8°25’04’’	34°57’82’’	438
Gambella-3	G3	14	8°24’47’’	34°58’39’’	442
Gambella-4	G4	13	8°22’50’’	34°59’79’’	443
Gambella-5	G5	9	7°89’39’’	34°52’12’’	460
Amhara-3	A3	4	11°97’86’’	37°70’72’’	1,801

3.2.2 Morphological data collection and gas exchange traits measurement

At the heading time, Plant Height (PH), Panicle Length (PL), first Leaf Length (LL1), second Leaf Length (LL2), third Leaf Length (LL3), first Leaf Width (LW1), second Leaf Width (LW2) and third Leaf Width (LW3) were measured in centimeters (cm) as described by IRRI-IBPGR (1980).

During the panicle initiation stage, the leaf photosynthetic characteristics such as; Photosynthesis rate (Pn), stomatal conductance (gs), Transpiration rate (E) and intercellular CO_2 concentration (Ci) for all the three replicates of the 62 *O. longistaminata* accessions were measured from fully expanded flag leaves by the portable photosynthesis measuring system (*LI6400XT* LI-COR, Lincoln, NE, USA) which was connected to *Leaf Chamber Fluorometer* (*6400-40, LI-COR*, USA). All variables were measured on sunny and windless days from 10:00 a.m. to 1:00 p.m.

Carboxylation efficiency (CE), water use efficiency (WUE) and intrinsic water use efficiency (iWUE) were calculated according to Giuliani *et al*., (2013) and Haritha *et al*., (2017) from the respective quotients; Pn/Ci, Pn/Tr, and Pn/gs.

3.2.3 Morphometric and photosynthetic data Analysis

Descriptive statistics for both morphological and leaf gas exchange traits were computed in IBM SPSS version 21 (IBM, 2012). For each trait, Analysis of Variance (ANOVA) among accessions was conducted using PROC GLM with Type III sums of squares in SAS 9.2. PROC CORR in SAS was also used to calculate pairwise Pearson correlation coefficients among the eight morphological and between the seven photosynthetic traits. Principal components analysis (PCA) was conducted for all morphological traits in MINITAB® Release 14.13 and communalities along with test of appropriateness for factor analysis was determined in SPSS using the Kaiser-Meyer-Olkin (KMO) and Bartlett tests.

3.2.4 Multivariate analysis

A series of multivariate analyses were done through appropriate procedures of the SAS computer program (SAS Institute, 1993). Using SAS, *O. longistaminata* accessions were grouped into their homogenous classes. For the sake of subsequent statistical analysis, number of clusters was determined by taking the maximum non-monogenic cluster number (Amsalu and Endashaw, 1999).

Genetic distance among clusters was calculated using the generalized Mahalanobis's D^2 statistics (Mahalanobis, 1936). Relationships among populations of origin were also assessed by measuring the inter-cluster distances using Mahalanobis distance (D^2) (Amsalu and Endashaw, 1999). Based on Euclidean distance from the 8 quantitative characters, cluster analysis for the populations of origin was run using MINITAB® Release 14.13. The clustering pattern of the six populations was further illustrated by UPGMA.

3.2.5 Molecular grouping

For the UPGMA analysis, DNA was extracted from leaves of 186 individuals representing each replicates of the 62 *O. longistaminata* accessions using a CTAB protocol (Doyle and Doyle 1987). Polymerase chain reaction (PCR) was done by using 7 SSR markers (Table 6) at a 10 µl reaction mixture in a 96-well plate. Each reaction mixture contained 4 µl of genomic DNA, 0.5 µl of each of the two primers (at a concentration of 10mM), 1.75 µl of a 10X Taq buffer, 0.5 µl of a 2.5 mM dNTP mixture, 0.1 µl of Taq DNA polymerase and 3 µl of double distilled water. The temperature profile used for PCR amplification consisted of a 94°C and 5 minutes preheating followed by 34 cycles of denaturation at 94°C for 45s, annealing at 55°C for 30 s, extension at 72°C for 30 s, and a final extension at 72°C for 7 min. The PCR products were run on PAGE and subsequently silver stained. From the scored bands, Poptree2 program computed genetic distance and UPGMA was drawn by tree view.

Table 6. Primers used, their sequence, repeat motif's, linkage group and annealing temperature (Source: Panaud *et al*., 1996; Chen *et al*., 1997 and Temnykh *et al*., 2000).

No	Primer	5'-3' Sequence of F/R Primer	Repeat motif	Chromosomal location	Annealing temperature
1	RM44	ACGGGCAATCCGAACAACC TCGGGAAAACCTACCCTACC	(GA)16	8	55^0C
2	RM225	TGCCCATATGGTCTGGATG GAAAGTGGATCAGGAAGGC	(CT)18	6	55^0C
3	RM31	GATCACGATCCACTGGAGCT AAGTCCATTACTCTCCTCCC	(GA)15	5	55^0C
4	RM209	ATATGAGTTGCTGTCGTGCG CAACTTGCATCCTCCCCTCC	(CT)18	11	55^0C
5	RM105	GTCGTCGACCCATCGGAGCCAC TGGTCGAGGTGGGGATCGGGTC	(CCT)6	10	55^0C
6	RM224	ATCGATCGATCTTCACGAGG TGCTATAAAAGGCATTCGGG	(AAG)8(AG)13G (AG)	11	55^0C
7	RM1	GCGAAAACACAATGCAAAAA GCGTTGGTTGGACCTGAC	(GA)26	1	55^0C

3.3 SSR marker based population study

3.3.1 Plant materials and their origin

A total of 430 accessions comprising 360 *O. longistaminata* accessions from 12 locations in Ethiopia (Table 7) along with 35 *O. sativa* accessions (*Indica* and *Japonica* cultivars), 25 accessions of 5 AA genome species, 5 accessions from CC genome, 1 accession from CCDD genome, 2 accessions from EE genome and one Japonica weedy type (Table 8) were included in this study.

Table 7. GIS information, ecology type and locality name of the Ethiopian *O. longistaminata* populations

Group code	**Site name**	**Latitude**	**Longitude**	**Elevation**	**No of accessions**	**Ecology type**
Gambella1 (G1)	Gambella Agricultural Research Institute	8°25’29’’ N	34°57’60’’ E	431M	30	In a small stream with no rice cultivation history
Gambella2 (G2)	Kera	8°25’04’’ N	34°57’82’’ E	438M	30	A wetland which is going to dry in a dry season without a rice cultivation practice
Gambella3 (G3)	Elai	8°24’47’’ N	34°58’39’’ E	442M	30	Baro river provides water and had rice cultivation trial in 2011
Gambella4 (G4)	Echewi	8°22’50’’ N	34°59’79’’ E	443M	30	A swampy area without any rice production practice
Gambella5 (G5)	Aleworo	7°89’39’’ N	34°52’12’’ E	460M	30	An isolated area and free from any rice cultivation
Gambella6 (G6)	Gambella Catholic Church	7°89’41’’ N	34°51’95’’ E	457M	30	A semi-wetland area without rice production history
Amhara1 (A1)	Tana Hotel	11°60’40’’ N	37°39’74’’ E	1,800M	30	Tana lake is the source of water and it is an isolated area with no rice production practice

Table 7 continued

Amhara 2 (A2)	Tewa Zakena	11°93’80’’ N	37°70’97’’ E	1,801M	30	A stream water near the road is the source and there is a rice farm at 100 Meter away from the collection site
Amhara 3 (A3)	Abwa Kokite School	11°97’86’’ N	37°70’72’’ E	1,801M	30	Many rice varieties by the Woreta Agricultural Institute had been grown at 50M distance
Amhara 4 (A4)	Gami Gote	11°70’37’’ N	37°31’37’’ E	1,790M	30	Had a trial to cultivate rice in 2011 and 2012 but the overflow of water from Tana lake makes it less likely
Amhara 5 (A5)	Wenjeta Kebelle	11°67’82’’ N	37°29’64’’ E	1,792M	30	A rice farm from 2009-2013. Though, it is abandoned now
Amhara 6 (A6)	Fessesa	11°68’70’’ N	37°29’33’’ E	1,798M	30	A rice farm

Table 8. Details on the accession code and genome type of 70 worldwide rice accessions representing 11 *Oryza* species and 1 weedy type from the Yunnan University and International Rice Research Institute gene banks

Sample ID	**Species/ Subspecies**	**Genome type**	**Accession code**	**Origin**
P01	*Japonica*	AA	Nippobrane	Yunnan University gene bank
P02	*Japonica*	AA	zhonghua11	Yunnan University gene bank
P03	*Japonica*	AA	Kitaake	Yunnan University gene bank
P04	*Indica*	AA	RD23	Yunnan University gene bank
P05	*Indica*	AA	minghui63	Yunnan University gene bank
P06	*Japonica*	AA	Koshihikari	Yunnan University gene bank
P07	*Japonica*	AA	yundao2	Yunnan University gene bank
P08	*Japonica*	AA	chujing27	Yunnan University gene bank
P09	*Indica*	AA	9311	Yunnan University gene bank
P10	*Indica*	AA	luyin46	Yunnan University gene bank
P11	*Indica*	AA	Huanghuazhan	Yunnan University gene bank
P12	*Japonica*	AA	IRAT104	Yunnan University gene bank
P13	*Indica*	AA	IR64	Yunnan University gene bank
P14	*Indica*	AA	guichao2	Yunnan University gene bank
P15	*Japonica*	AA	Xueheaizao	Yunnan University gene bank
P16	*Japonica*	AA	Hejiang19	Yunnan University gene bank
P17	*Japonica*	AA	Pokkali	Yunnan University gene bank
P18	*Indica*	AA	IR29	Yunnan University gene bank
P19	*Indica*	AA	Mengwanggu	Yunnan University gene bank
P20	*Indica*	AA	taiwanludao1	Yunnan University gene bank
P21	*Japonica*	AA	chujing28	Yunnan University gene bank
P22	*Japonica*	AA	dianjingyou5	Yunnan University gene bank
P23	*Indica*	AA	wendao12	Yunnan University gene bank
P24	*Japonica*	AA	hexi41	Yunnan University gene bank
P25	*Indica*	AA	yunhui290	Yunnan University gene bank
P26	*Indica*	AA	banna21	Yunnan University gene bank
P27	*Japonica*	AA	Akitakomati	Yunnan University gene bank
P28	*Japonica*	AA	Akihikari	Yunnan University gene bank
P29	*Indica*	AA	dianrui449	Yunnan University gene bank
P30	*Indica*	AA	diantun502	Yunnan University gene bank
P31	*Japonica*	AA	PR23	Yunnan University gene bank
P37	*Japonica*	AA	yunjing37	Yunnan University gene bank
P41	*Indica*	AA	mengwangmaxiangu	Yunnan University gene bank
P42	*Indica*	AA	Mengwangdabaigu	Yunnan University gene bank
P43	*Indica*	AA	mengwangxiaobaigu	Yunnan University gene bank
YSD16	*O.barthii*	AA	Acc.100921	Mali
WR23	*O.barthii*	AA	Acc.104061	Niger
WR26	*O.barthii*	AA	Acc.104284	Mali
YSD88	*O.glumaepatula*	AA	Acc.100894	Cuba
YSD90	*O.glumaepatula*	AA	Acc.103812	Brazil
YSD92	*O.glumaepatula*	AA	Acc.105661	Brazil
YSD95	*O.glumaepatula*	AA	Acc.105662	Brazil
WR88	*O.glumaepatula*	AA	Acc.100894	Cuba
WR90	*O.glumaepatula*	AA	Acc.103812	Brazil
YSD64	*O.nivara*	AA	Acc.80611	India
YSD66	*O.nivara*	AA	Acc.80677	India
WR60	*O.nivara*	AA	Acc.80432	India
WR65	*O.nivara*	AA	Acc.80625	India

Table 8 continued

WR77	*O.nivara*	AA	Acc.80696	India
WR78	*O.nivara*	AA	Acc.81855	India
YSD42	*O.rufipogon*	AA	Acc.82040	Thailand
YSD46	*O. rufipogon*	AA	Acc.106133	India
YSD52	*O.rufipogon*	AA	PCR9607	Thailand
WR40	*O. rufipogon*	AA	Acc.81994	Papua New Guinea
WR46	*O. rufipogon*	AA	Acc.106133	India
WR48	*O.rufipogon*	AA	Acc.106340	Myanmar
WR55	*O.rufipogon*	AA	Acc.105832	Bgo2,1998 gkanlaus
WR57	*O.rufipogon*	AA	Acc.106138	Bgo2,2001 gkanlaus
YSD55	*O.rufipogon*	AA	Acc.105832	Bgo2,1998 gkanlaus
YSD57	*O.rufipogon*	AA	Acc.106138	Bgo2,2001 gkanlaus
Lg	*O. longistaminata*	AA	PR1	Niger
YSD01	*O. officinalis*	CC	8712	Sukhothai
YSD05	*O.rhizomatis*	CC	W95018	IRRI
WR05	*O.rhizomatis*	CC	W95018	IRRI
WR03	*O.eichingeri*	CC	Acc.105181	Uganda
YSD03	*O.eichingeri*	CC	Acc.105181	Uganda
YSD08	*O. latifolia*	CCDD	Acc.100169	Costa Rice
WR09	*O.australiensis*	EE	Acc.101410	Australia
YSD09	*O.australiensis*	EE	Acc.101410	Australia
YSD58	*Japonica* weedy type		96036	Korea

3.3.2 Genomic DNA extraction and Polymerase chain reaction (PCR)

Total genomic DNA was extracted from silica gel dried leaves of the 360 Ethiopian *O. longistaminata* accessions and fresh leaves of the rest 70 rice accessions by using CTAB protocol as described by Doyle (1987). Quality of the extracted DNA was determined by electrophoresis in a 1% agarose gel and quantification was accomplished using spectrophotometer.

A total of 67 Simple Sequence Repeat (SSR) markers across the 12 rice chromosomes were amplified (Table 9). Such SSR markers were selected for PCR based on the reports from Panaud *et al.*, (1996), Chen *et al.*, (1997) and Temnykh *et al.*, (2000).

PCR was done in a 10 µl reaction mixture containing 4 µl of genomic DNA, 0.5 µl of each of the two primers (at a concentration of 10mM), 1.75 µl of a 10X Taq buffer, 0.5 µl of a 2.5

mMdNTP mixture, 0.1 µl of 2.5U/ µlTaq DNA polymerase and 3 µl of double distilled water. The temperature profile used for PCR amplification consisted of 5 minutes preheating at 94°C followed by 34 cycles of denaturation at 94°C for 45s, annealing at 55°C to 67°C (Table 9) for 30 s, extension at 72°C for 30 s, and a final extension at 72°C for 7 min.

Table 9. Primers from Gramene database with their sequence, repeat motifs, chromosomal location and annealing temperature

S. No	Primer	5’-3’ Sequence of FWD/REV Primer	SSR repeat motif	Chromosomal location	Annealing Temperature
1	RM1	FWD: GCGAAAACACAATGCAAAAA REV: GCGTTGGTTGGACCTGAC	(GA)26	1	55
2	RM5	FWD: TGCAACTTCTAGCTGCTCGA REV: GCATCCGATCTTGATGGG	(GA)14	1	55
3	RM6	FWD: TGCAACTTCTAGCTGCTCGA REV: TCGTCTACTGTTGGCTGCAC	(AG)16	2	55
4	RM11	FWD: TCTCCTCTTCCCCCGATC REV: ATAGCGGGCGAGGCTTAG	(GA)17	7	55
5	RM14	FWD: CCGAGGAGAGGAGTTCGAC REV: GTGCCAATTTCCTCGAAAAA	(GA)17	1	55
6	RM17	FWD: TGCCCTGTTATTTTCTTCTCTC REV: GGTGATCCTTTCCCATTTCA	(GA)21	12	55
7	RM19	FWD: CAAAAACAGAGCAGATGAC REV: CTCAAGATGGACGCCAAGA	(ATC)10	12	55
8	RM21	FWD: ACAGTATTCCGTAGGCACGG REV: GCTCCATGAGGGTGGTAGAG	(GA)21	11	55
9	RM22	FWD: GGTTTGGGAGCCCATAATCT REV: CTGGGCTTCTTTCACTCGTC	(GA)22	3	55
10	RM23	FWD: CATTGGAGTGGAGGCTGG REV: GTCAGGCTTCTGCCATTCTC	(GA)15	1	55
11	RM24	FWD: GAAGTGTGATCACTGTAACC REV: TACAGTGGACGGCGAAGTCG	(GA)29	1	55
12	RM25	FWD: GGAAAGAATGATCCTTTTCATGG REV: CTACCATCAAAACCAATGTTC	(GA)18	8	55
13	RM26	FWD: GAGTCGACGAGCGGCAGA REV: CTGCGAGCGACGGTAACA	(GA)15	5	55
14	RM29	FWD: CAGGGACCCACCTGTCATAC REV: AACGTTGGTCATATCGGTGG	(GA)7-18(GA)5(AG)4	2	55
15	RM30	FWD: GGTTAGGCATCGTCACGG REV: TCACCTCACCACACGACACG	(AG)9A(AGA)12	6	55
16	RM31	FWD: GATCACGATCCACTGGAGCT REV: AAGTCCATTACTCTCCTCCC	(GA)15	5	55
17	RM42	FWD: ATCCTACCGCTGACCATGAG REV: TTTGGTCTACGTGGCGTACA	(AG)5(GA)(AG)2T(GA)A(AG)5	8	55
18	RM44	FWD: ACGGGCAATCCGAACAACC REV: TCGGGAAAACCTACCCTACC	(GA)16	8	55

Table 9 continued

S. No	Primer	5’-3’ Sequence of FWD/REV Primer	SSR repeat motif	Chromosomal location	Annealing Temperature
19	RM49	FWD: TTCGGAAGTTGGTTACTGATCA REV: TTGGAGCGGATTCGGAGG	(GA)27	3	55
20	RM50	FWD: ACTGTACCGGTCGAAGACG REV: AAATTCCACGTCAGCCTCC	(CTAT)4(CT)15	6	55
21	RM55	FWD: CCGTCGCCGTAGTAGAGAAG REV: TCCCGGTTATTTTAAGGCG	(GA)17	3	55
22	RM60	FWD: AGTCCCATGTTCCACTTCCG REV: ATGGCTACTGCCTGTACTAC	(AATT)5AATCT(AATT)	3	55
23	RM80	FWD: TTGAAGGCGCTGAAGGAG REV: CATCAACCTCGTCTTCACCG	(TCT)25	8	55
24	RM84	FWD: TAAGGGTCCATCCACAAGATG REV: TTGCAAATGCAGCTAGAGTAC	(TCT)10	1	55
25	RM105	FWD: GTCGTCGACCCATCGGAGCCAC REV: TGGTCGAGGTGGGGATCGGGTC	(CCT)6	10	55
26	RM110	FWD: TCGAAGCCATCCACCAACGAAG REV: TCCGTACGCCGACGAGGTCGAG	(GA)15	2	55
27	RM118	FWD: CCAATCGGAGCCACCGGAGAGC REV: CACATCCTCCAGCGACGCCGAG	(GA)8	7	67
28	RM125	FWD: ATCAGCAGCCATGGCAGCGACC REV: AGGGGATCATGTGCCGAAGGCC	(GCT)8	7	55
29	RM126	FWD: CGCGTCCGCGATAAACACAGGG REV: TCGCACAGGTGAGGCCATGTCG	(GA)7	8	55
30	RM129	FWD: TCTCTCCGGAGCCAAGGCGAGG REV: CGAGCCACGACGCGATGTACCC	(CGG)8	1	55
31	RM133	FWD: TTGGATTGTTTTGCTGGCTCGC REV: GGAACACGGGGTCGGAAGCGAC	(CT)8	6	60
32	RM134	FWD: ACAAGGCCGCGAGAGGATTCCG REV: GCTCTCCGGTGGCTCCGATTGG	(CCA)7	7	55
33	RM136	FWD: GAGAGCTCAGCTGCTGCCTCTAGC REV: GAGGAGCGCCACGGTGTACGCC	(AGG)7	6	55
34	RM144	FWD: TGCCCTGGCGCAAATTTGATCC REV:GCTAGAGGAGATCAGATGGTAGTGCATG	(ATT)11	11	57
35	RM152	FWD: GAAACCACCACACCTCACCG REV: CCGTAGACCTTCTTGAAGTAG	(GGC)10	8	55
36	RM153	FWD: GCCTCGAGCATCATCATCAG REV: ATCAACCTGCACTTGCCTGG	(GAA)9	5	55
37	RM154	FWD: ACCCTCTCCGCCTCGCCTCCTC REV: CTCCTCCTCCTGCGACCGCTCC	(GA)21	2	61
38	RM159	FWD: GGGGCACTGGCAAGGGTGAAGG REV: GCTTGTGCTTCTCTCTCTCTCTCTCTCTC	(GA)19	8	55
39	RM161	FWD: TGCAGATGAGAAGCGGCGCCTC REV: TGTGTCATCAGACGGCGCTCCG	(AG)20	5	57
40	RM162	FWD: GCCAGCAAAACCAGGGATCCGG REV: CAAGGTCTTGTGCGGCTTGCGG	(AC)20	6	61
41	RM164	FWD: TCTTGCCCGTCACTGCAGATATCC REV: GCAGCCCTAATGCTACAATTCTTC	(GT)16TT(GT)4	6	55
42	RM167	FWD: GATCCAGCGTGAGGAACACGT REV: AGTCCGACCACAAGGTGCGTTGTC	GGAA(GA)16GGGG	11	55
43	RM171	FWD: AACGCGAGGACACGTACTTAC REV: ACGAGATACGTACGCCTTTG	(GATG)5	10	55

Table 9 continued

S. No	Primer	5’-3’ Sequence of FWD/REV Primer	SSR repeat motif	Chromosomal location	Annealing Temperature
44	RM180	FWD: CTACATCGGCTTAGGTGTAGCAACACG REV: ACTTGCTCTACTTGTGGTGAGGGACTG	(ATT)10	7	55
45	RM184	FWD: ATCCCATTCGCCAAAACCGGCC REV: TGACACTTGGAGAGCGGTGTGG	(CA)7	10	55
46	RM205	FWD: CTGGTTCTGTATGGGAGCAG REV: CTGGCCCTTCACGTTTCAGTG	(GA)25	9	55
47	RM206	FWD: TAGTTTAACCAAGACTCTC REV: GGTTGAACCCAAATCTGCA	(GA)19	2	55
48	RM207	FWD: CCATTCGTGAGAAGATCTGA REV: CACCTCATCCTCGTAACGCC	(CT)25	2	55
49	RM209	FWD: ATATGAGTTGCTGTCGTGCG REV: CAACTTGCATCCTCCCCTCC	(CT)18	11	55
50	RM211	FWD: CCGATCTCATCAACCAACTG REV: CTTCACGAGGATCTCAAAGG	(TC)3A(TC)18	2	55
51	RM212	FWD: CCACTTTCAGCTACTACCAG REV: CACCCATTTGTCTCTCATTATG	(GA)24	1	55
52	RM213	FWD: ATCTGTTTGCAGGGGACAAG REV: AGGTCTAGACGATGTCGTGA	(CT)17	2	55
53	RM215	FWD: CAAAATGGAGCAGCAAGAGC REV: TGAGCACCTCCTTCTCTGTAG	(GA)16	9	55
54	RM216	FWD: TTCCCCAATGGAACAGTGAC REV: AGGGTCTACCACCCGATCTC	(CT)18	10	55
55	RM219	FWD: CGTCGGATGATGTAAAGCCT REV: CATATCGGCATTCGCCTG	(GA)17	9	55
56	RM220	FWD: GGAAGGTAACTGTTTCCAAC REV: GAAATGCTTCCCACATGTCT	(CT)17	1	55
57	RM224	FWD: ATCGATCGATCTTCACGAGG REV: TGCTATAAAAGGCATTCGGG	(AAG)8(AG)13G(AG)	11	55
58	RM225	FWD: TGCCCATATGGTCTGGATG REV: GAAAGTGGATCAGGAAGGC	(CT)18	6	55
59	RM251	FWD: GAATGGCAATGGCGCTAG REV: ATGCGGTTCAAGATTCGATC	(CT)29	3	55
60	RM263	FWD: CCCAGGCTAGCTCATGAACC REV: GCTACGTTTGAGCTACCACG	(CT)34	2	55
61	RM307	FWD: GTACTACCGACCTACCGTTCAC REV: CTGCTATGCATGAACTGCTC	(AT)14(GT)21	4	55
62	RM322	FWD: CAAGCGAAAATCCCAGCAG REV: GATGAAACTGGCATTGCCTG	(CAT)7	2	55
63	RM333	FWD: GTACGACTACGAGTGTCACCAA REV: GTCTTCGCGATCACTCGC	(TAT)19(CTT)19	10	55
64	RM335	FWD: GTACACACCCACATCGAGAAG REV: GCTCTATGCGAGTATCCATGG	(CTT)25	4	55
65	RM337	FWD: GTAGGAAAGGAAGGGCAGAG REV: CGATAGATAGCTAGATGTGGCC	CTT64 8 (CTT)4-19-(CTT)8	8	55
66	RM475	FWD: CCTCACGATTTTCCTCCAAC REV: ACGGTGGGATTAGACTGTGC	(TATC)8	2	55
67	RM6344	FWD: ACACGCCATGGATGATGAC REV: TGGCATCATCACTTCCTCAC	(GAA)8	7	55

3.3.3 Polyacrylamide gel electrophoresis and SSR alleles scoring

The amplified PCR products were electrophoresed in an 8.0% polyacrylamide gel and detected through silver staining as described by Panaud *et al*., (1996). Fragment size was determined using a standardized DNA ladder. Based on the expected PCR product size given in the GRAMENE website (http://www.gramene.org), size of the most intensely amplified bands was identified as alleles for the SSR loci. Differently sized amplified bands were scored as different genotypes. The bands were recorded as (11, 22, 33. . .) to represent homozygous genotypes or (12, 13, 23. . .) to indicate the heterozygous genotypes and'?' was used to denote missing data.

3.3.4 Genotypic data analysis

The number of alleles per locus, expected and observed heterozygosity, Polymorphic Information Content (PIC) and *F* statistics such as Genetic differentiation (F_{st}), Wrights fixation index (F_{is}) and Total inbreeding coefficient (F_{it}) were calculated using GenAlEx 6.502 (Peakall and Smouse, 2012). The major allelic frequency and PIC of each marker were computed using Power Marker Version 3.25 (Liu and Muse, 2005). Nei genetic distance and identity along with Analysis of Molecular Variance (AMOVA) for the 12 *O. longistaminata* populations of Ethiopia and principal components analysis (PCA) for the 26 rice groups were also conducted by GenAlEx6.502.

Based on their computed pair wise genetic distance value, a dendrogram was drawn for the 430 *Oryza* accessions by using DARwin V6 software (Perrier and Jacquemoud-Collet, 2006). Unweighted Pair Group Method with Arithmetic Mean (UPGMA) for the 26 rice groups was drawn by the Poptree2 (http://www.ualberta.ca/~fyeh/fyeh) program.

Population structure was inferred via the STRUCTURE v. 2.3 model (Pritchard *et al*., 2000). Three replications were run with a burn-in period of 150,000 and 150,000 repeats of the MCMC method for K value 1 to 26. The optimal number of subpopulations (K) was determined using the ΔK approach as described by Evanno *et al*. (2005) using STRUCTUREHARVESTER v0.6.8. (Earl and von Holdt, 2011).

3.4 RAD-Seq based study

3.4.1 RAD library preparation and sequencing

Total genomic DNA extracted from young leaves using a standard CTAB procedure (Doyle, 1987) was concentrated, quantified by PicoGreen and normalized to a 50 ng/ul concentration (Poland *et al*., 2012). The RAD tag library was prepared as described by Clark *et al*., 2014. Briefly, genomic DNA of 87 *O. longistaminata* accessions from six populations (Gambella1, Gambella2, Gambella3, Gambella4, Gambella5 and Amhara3) were fragmented by *Pst*I and *Msp*I enzymes in a 96-well format. The digestion was held in a 15 µl reaction volume containing 8µl nuclease free water, 0.25µl of 20,000 U/ml of *MspI*, 0.25 µl of *PstI*-HF (High-Fidelity), 1.5µl 1X NEBuffer and 5µl DNA for 3 hrs at 37°C followed by enzymes inactivation at 80 °C for 20 min.

In the same tube/plate, sequencing adapters with sample specific indices (barcodes) were added and ligated to the *Pst*I end. An 8.5ul of ligation master mix consisting of 1ul 10X ligase buffer with ATP, 0.5ul 10uM *MspI* adapter, 1.5ul 10mM ATP, 0.1uL 2M U/ml T4 ligase and 5.4ul nuclease free water was used for the ligation reaction at 27 °C for 2 hrs and ligase inactivation at 65 °C for 20 min.

The ligated products were pooled in equal volumes and run for 20 minutes at 100V in a 2% agarose gel. A smear between 200 and 500bp were excised and purified with the QIAquick PCR Purification Kit (Qiagen). As described by Clark *et al*., (2014), polymerase chain reaction (PCR) was carried out by paired primers that can anneal to the *PstI* and *MspI* overhangs. The gel extracted PCR products with a 200-500 bp size were purified again and a 150-bp paired-end reads were generated by Illumina sequencing.

3.4.2 RAD-seq data processing and analysis

In accordance to the barcodes used for each *O. longistaminata* accession, the process_radtags program from the Stacks v1.48 package was used to demultiplex and sort the FASTQ sequence (Catchen *et al*., 2013a). During this process, reads having high sequencing quality with the right barcode and an unambiguous RAD site were retained. The demultiplexed reads were aligned against the Nipponbare reference genome sequence (version BROADs1, Ensembl release 64) via the Bawtie2 Alignment Tool version 0.11.3 (Langmead *et al*., 2009). The pstacks program extracted stacks that have been aligned to the reference genome. Loci data of the whole *O. longistaminata* accessions were merged into a catalog using cstacks and allele status in each accession was determined by the sstacks (Catchen *et al*., 2013a).

Using the adegenet package of R, genetic relationships among the Ethiopian *O. longistaminata* accessions were analyzed by the Discriminant Analysis of Principal Components (DAPC) (Jombart *et al*., 2010; Pan *et al*., 2010). The UNEAK pipeline in TASSEL 3.0.162 was also used for the base recalibration and phylogenetic tree construction at a minimum call rate of 0.5 and a minimum minor allele frequency of 0.01 (Lu *et al*., 2013).

Based on the complete data produced by Stacks, Nei's genetic distance, pairwise population differentiation (*Fst*) and fixation index (F_{is}) between the Amhara and Gambella *O. longistaminata* populations was computed.

4. Results

4.1 Morphometric and photosynthetic study

4.1.1 Descriptive statistics of morphological traits

The grand mean and standard error of the traits; Plant Height, Panicle Length, First Leaf Length, second Leaf Length, third Leaf Length, first Leaf Width, second Leaf Width and third Leaf Width, for all *O. longistaminata* individuals included in this study, were 176.4cm ±14.7; 26.33cm ±3.14; 36.67cm ±5.48; 52.47cm ±6.92; 57.1cm ±5.72; 1.63cm ±0.2; 1.58cm ±0.17 and 1.6cm ±0.17 respectively (Table 10). All the eight quantitative traits showed a remarkably significant differences (p<0.001) among the 62 Ethiopian *O. longistaminata* accessions. However, estimation of four phenotypic traits, namely, Plant Height (6.14%), first Leaf Width (9.46%), second Leaf Width (9.27%) and third Leaf Width (9.14%) revealed lower coefficients of variation (CV).

Table 10. Range of the eight quantitative traits along with the ANOVA and CV values from the 62 Ethiopian *O. longistaminata* accessions

Trait	Minimum	Maximum	Mean	Std Dev	MSE (122)	CV	F test (61)
PH	145.67	206.03	176.4	14.7	117.25	6.14%	5.47***
PL	15.70	32.83	26.33	3.14	8.85	11.31%	3.32***
LL1	25.40	47.18	36.67	5.48	25.82	13.85%	3.52***
LL2	33.53	64.40	52.43	6.92	44.76	12.80%	3.06***
LL3	43.67	69.17	57.1	5.72	38.08	10.83%	2.50***
LW1	1.23	2.07	1.63	0.20	0.02	9.46%	5.01***
LW2	1.20	1.90	1.58	0.17	0.02	9.27%	3.94***
LW3	1.20	2.10	1.60	0.17	0.02	9.14%	3.74***

***** (highly significant at p<0.001)**

Key; PH (plant height), PL (panicle length), LL1 (first leaf length), LL2 (second leaf length), LL3 (third leaf length), LW1 (first leaf width), LW2 (second leaf width), LW3 (third leaf width) MSE (Mean square error), CV (Coefficient of Variation) and numbers in parenthesis represent degree of freedom

4.1.2 Correlation Coefficients of morphological traits

Correlations between the eight variable traits of the assessed 62 Ethiopian *O. longistaminata* accessions showed a perfect and highly significant association of Plant Height with all the others except the third Leaf width (Table 11). Likewise, most of the evaluated morphological traits showed significant pairwise correlations. However, the third leaf length was exceptionally and insignificantly correlated with the first leaf width, second leaf width and third leaf width.

Table 11. Pairwise correlation coefficients of the eight morphological traits of *O. longistaminata*

	PH	**PL**	**LL1**	**LL2**	**LL3**	**LW1**	**LW2**	**LW3**
PH	1.00	0.63***	0.40***	0.42***	0.28*	0.35**	0.39**	0.22
PL		1.00	0.64***	0.61***	0.34**	0.54***	0.51***	0.36**
LL1			1.00	0.76***	0.43***	0.66***	0.53***	0.46***
LL2				1.00	0.59***	0.63***	0.51***	0.35**
LL3					1.00	0.19	0.08	0.03
LW1						1.00	0.87***	0.73***
LW2							1.00	0.78***
LW3								1.00

* Significant at p<0.05; ** Highly significant at p<0.01; *** Highly significant at p<0.001

Key; PH (plant height), PL (panicle length), LL1 (first leaf length), LL2 (second leaf length), LL3 (third leaf length), LW1 (first leaf width), LW2 (second leaf width), LW3 (third leaf width)

4.1.3 Principal Component Analysis and traits contribution

The eight agro-morphological characters assessed in this study were subjected to principal component analysis. Hence, the first two principal axes with eigenvalues greater than unity were selected. These two principal components explained 74% of the total variation (Table 12). The first axis which accounted for 56% of the variation was contributed by Plant Height (0.622), Panicle length (0.710), first Leaf Length (0.704), second Leaf Length (0.811) and third Leaf Length (0.814). Whereas, high loadings for the second component which described 18 % of the variance was accounted by first Leaf Width (0.865), second Leaf Width (0.913) and third Leaf Width (0.896). In concordance to the communality value, first Leaf Width (0.879), second Leaf Width (0.889) and third Leaf Width (0.808) contributed strongly to the divergence of the 62 Ethiopian *O. longistaminata* accessions (Table 12).

Table 12. Eigen values of the first two principal components of the eight morphological variables and the loading factors of each variable

Variable	Factor 1	Factor 2	Communality
PH	0.234	**0.622**	0.441
PL	0.402	**0.710**	0.666
LL1	0.478	**0.704**	0.742
LL2	0.358	**0.811**	0.786
LL3	-0.170	**0.814**	0.691
LW1	**0.865**	0.361	0.879
LW2	**0.913**	0.235	0.889
LW3	**0.896**	0.068	0.808
Eigen value	4.44	1.44	
Proportion	0.56	0.18	
Cumulative	0.56	0.74	

** Coefficients with an absolute value of greater than or equal to 0.489 are shown in bold*

Key; PH (plant height), PL (panicle length), LL1 (first leaf length), LL2 (second leaf length), LL3 (third leaf length), LW1 (first leaf width), LW2 (second leaf width), LW3 (third leaf width)

4.1.4 Cluster Analysis

Based on a combined analysis of the eight quantitative traits, dendrogram from the Euclidean cluster analysis of the 62 Ethiopian *O. longistaminata* accessions showed six non-monogenic clusters (Figure 10). Cluster III contained the highest number of accessions at 29 followed by cluster II and cluster IV which comprised of 13 and 10 accessions, respectively (Table 13). Cluster I had 4 accessions and each of cluster V and VI were represented by only three accessions.

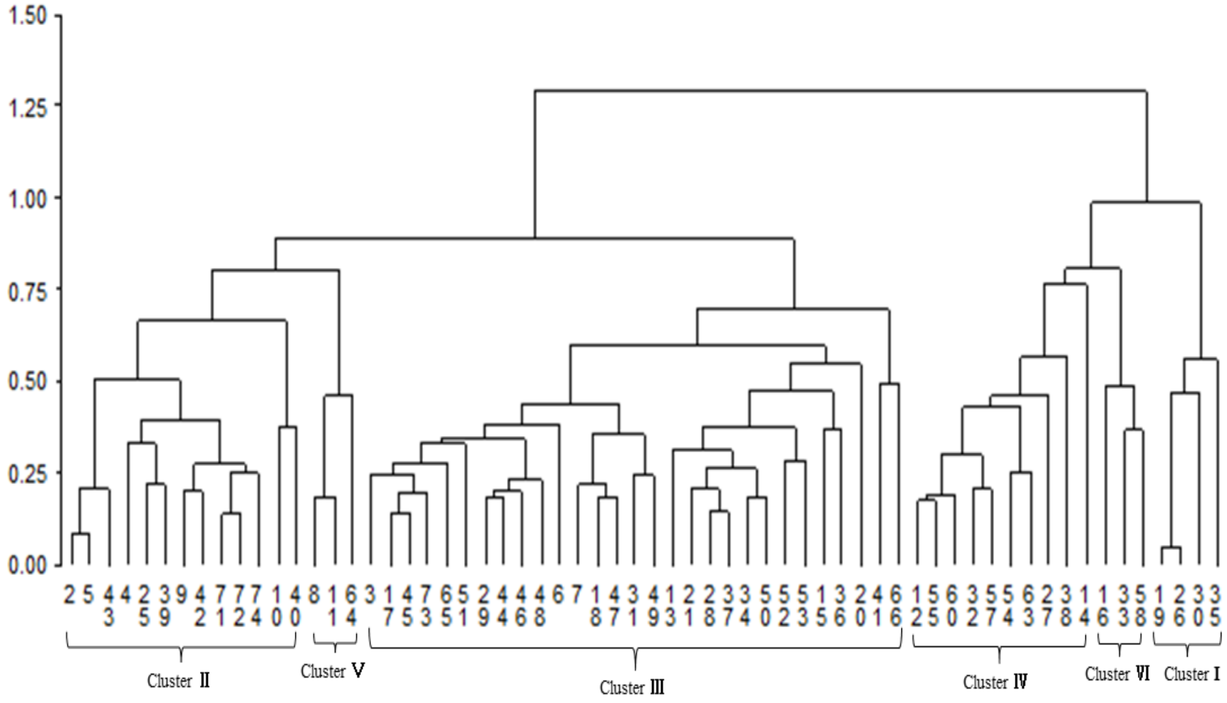

Figure 10. Dendrogram of 62 Ethiopian *O. longistaminata* accessions using Euclidian distance from eight quantitative traits (The six main clusters are labeled as I-VI with their own specific accession numbers as defined in Appendix I)

Table 13. Cluster membership of the 62 Ethiopian *O. longistaminata* accessions by eight morphological traits

Cluster	**Accessions**
I (4)	26, 19, 30, 35
II (13)	2, 5, 71, 72, 9, 42, 43, 25, 39, 74, 4, 10, 40
III (29)	17,45,28,37,47,18,29,44,50,34,73,46,23,7,48,3,49,31,65,52,53,13,51,15,36,6,41,66,20
IV (10)	12, 55, 60, 57, 32, 54, 63, 27, 38, 14
V (3)	8, 11, 64
VI (3)	58, 33, 16

Numbers in parenthesis indicate number of accessions grouped in to each cluster

Comparison of the mean values of the six independent clusters showed cluster VI to be with the least mean in all of the eight quantitative traits, cluster II with the highest values for Panicle Length (29.21), first Leaf Length (42.11), second Leaf Length (60.77), third Leaf Length (62.64), first Leaf Width (1.76) and second Leaf Width (1.68). While, cluster V showed best performance only for Plant Height (196.2) and third Leaf Width (1.64) (Table 14).

Table 14. Mean value of the eight quantitative traits of each cluster

Cluster/Traits	PH	PL	LL1	LL2	LL3	LW1	LW2	LW3
Cluster I	150.72	27.02	40.28	59.75	61.44	1.73	1.63	1.63
Cluster II	192.54	29.21	42.11	60.77	62.64	1.76	1.68	1.63
Cluster III	178.22	26.52	36.03	51.31	56.84	1.59	1.57	1.61
Cluster IV	162.41	23.23	33.92	47.22	52.34	1.59	1.52	1.53
Cluster V	196.20	28.31	34.29	48.94	52.28	1.71	1.66	1.64
Cluster VI	150.04	19.52	25.95	38.21	50.43	1.31	1.34	1.49
Grand mean	**171.69**	**25.63**	**35.43**	**51.03**	**56.00**	**1.62**	**1.57**	**1.59**

Accessions' clustering in relation to their populations of origin, showed inclusion of higher proportions of accessions from Gambella 1 (63.6%), Gambella2 (54.5%) and Gambella4 (50%) in cluster III (Table 15). Though Gambella3 accessions were distributed in five of the six clusters, relatively larger portion of them were equally residing under cluster III (35.7%) and cluster IV (28.6%). Moreover, majority of the accessions from Gambella 5 were also grouped in both cluster II (50%) and cluster III (37.5%). Amhara3 was highly distinguishable and 75% of the accessions were grouped in cluster II. Cluster I, V and VI had accessions from Gambella1, Gambella2, Gambella3 and Gambella4 populations with a smallest percentage ranging from 7.15 to 14.3 (Table 15).

Table 15. Clustering pattern of 62 Ethiopian *O. longistaminata* accessions from six geographic populations of origin over six clusters

Populations	No of accessions	Percentage of accessions in each cluster					
		Cluster **I**	Cluster **II**	Cluster **III**	Cluster **IV**	Cluster **V**	Cluster **VI**
Gambella 1	11		27.3	63.6		9.1	
Gambella 2	11		18.2	54.5	18.2	9.1	
Gambella 3	14	14.3	7.1	35.7	28.6		14.3
Gambella 4	14	14.3		50	21.4	7.15	7.15
Gambella 5	8		50	37.5	12.5		
Amhara 3	4		75	25			
Total No	**62**						

4.1.5 Pairwise genetic distance and dendrogram on the basis of populations of origin

Based on the Mahalanobis distance (D^2) value, all of the six populations of origin looked different from each other (Table 16). The inter-population distances among those six populations showed Amhara 3 and Gambella 3 (6.25) as the most distantly related and Amhara3 and Gambella2 (6.02) as the second distant populations (Table 16). On the basis of mean of the eight quantitative traits per population, the closely related populations were Gambella1 and Gambella5 (1.58) and Gambella 2 and Gambella 3 (2.22).

Table 16. Pairwise Generalized Square Distance (D^2) between populations from eight morphological traits

Populations/ CLS	Gambella 1	Gambella 2	Gambella 3	Gambella 4	Gambella 5	Amhara 3
Gambella 1	0					
Gambella 2	4.17	0				
Gambella 3	4.03	2.22	0			
Gambella 4	3.83	3.01	2.49	0		
Gambella 5	1.58	3.88	3.94	3.10	0	
Amhara 3	3.01	6.02	6.25	5.75	3.39	0

Means of the eight quantitative characters at populations level were also used to draw a dendrogram. The cladogram had two major clades where one of it clearly showed close relationship between Gambella1, Gambella2 and Gambella5 with a higher distinctness from Amhara3 and higher similarity between Gambella3 and Gambella4 in the second cluster (Figure 11).

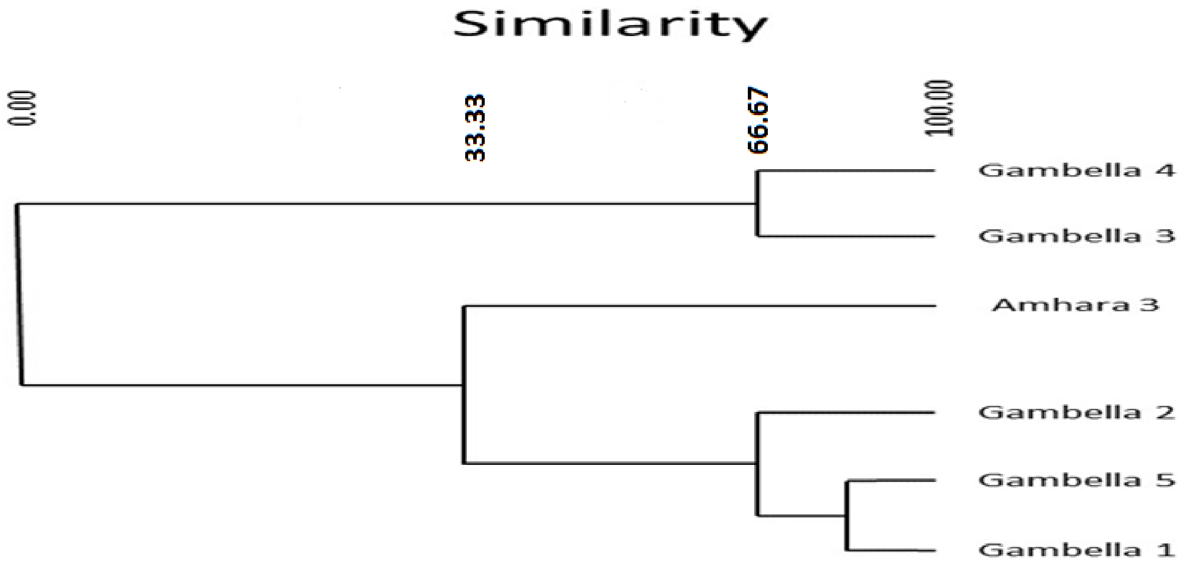

Figure 11. Dendrogram based on eight quantitative traits showing clustering of the six populations

In this study, further clustering analysis was also employed by using seven nuclear SSR markers that genotyped all of the investigated accessions. The UPGMA tree generated from the microsatellite markers showed close genetic similarity among Gambella1, Gambella2, Gambella3 and Gambella4 populations and distinctness of the Gambella5 and Amhara3 from each other and from the rest all groups (Figure 12).

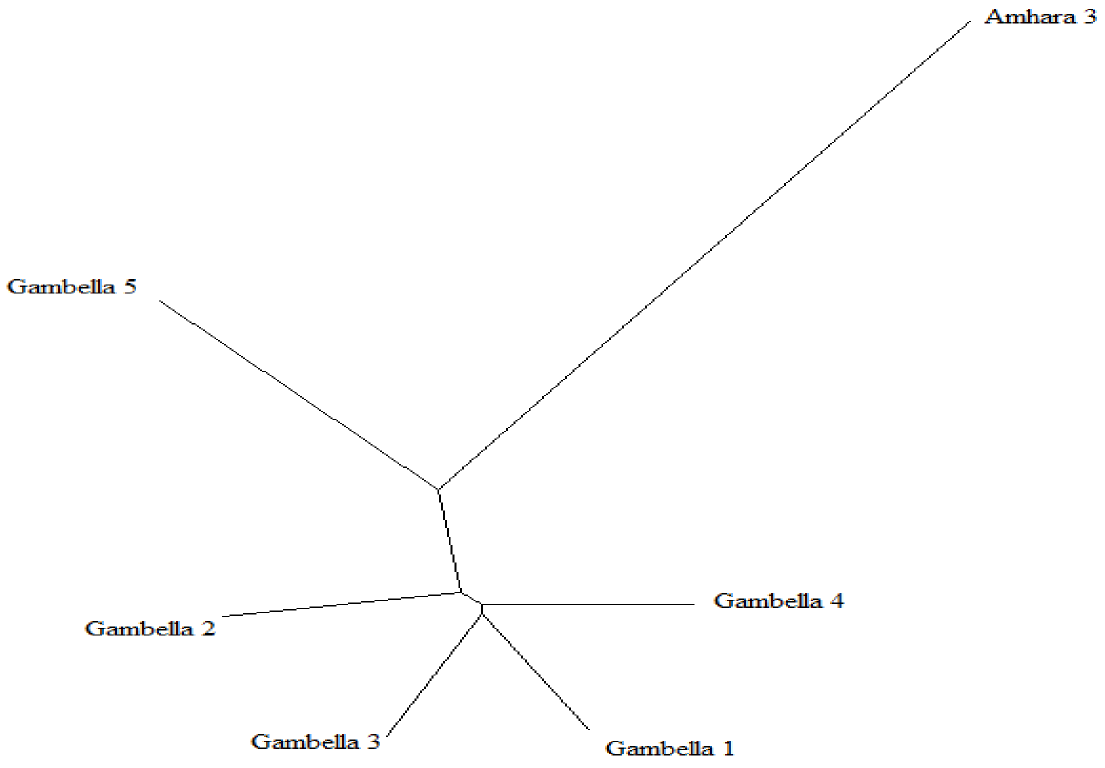

Figure 12. Cladogram of 6 *O. longistaminata* populations using 7 SSR primers over all individuals under study

4.1.6 Photosynthetic traits

Gas exchange measurements showed highly significant variation (P < 0.001) of Pn (Table 17). The Pn record across accessions ranged from 8.29 µmol CO_2 m^{-2} s^{-1} to 18.47 µmol CO_2 m^{-2} s^{-1} with a mean of 12.84 µmol CO_2 m^{-2} s^{-1}. Highly significant variation (P < 0.001) in *gs* was also observed. *g*s varied between a maximum of 0.1 mol H_2O m^{-2} s^{-1} and minimum of 0.05 mol H_2O m^{-2} s^{-1} with a 0.07 mol H_2O m^{-2} s^{-1} mean value. A wide and significant variation ($P < 0.001$) of *C*i which ranged from 12.14 µmol/mol to 134.65 µmol/mol with a mean value of 79.34 µmol/mol was also observed in this study. A highly significant difference ($P < 0.001$) with a mean of 3.14 mmol H_2O m^{-2} s^{-1} and range of 2.18 mmol H_2O m^{-2} s^{-1} to 4.05 mmol H_2O m^{-2} s^{-1} were noticed for the E. In this study, CE was the only leaf functional trait with a significant difference ($P < 0.05$) (Table 17). CE ranged from 0.07 and 0.57 with an average value of 0.15. Highly significant differences ($P < 0.001$) were also observed for both WUE and WUEi. WUE values ranged from 3.17 µmol CO_2/mmol H_2O to 5.17 µmol CO_2/mmol H_2O and WUEi was in a range of 134.92 µmol CO_2/mol H_2O and 221.45 µmol CO_2/mol H_2O. Both WUE and WUEi showed a mean of 4.08 µmolCO_2/mmol H_2O and 176.23 µmol CO_2/mol H_2O, respectively (Table 17).

Table 17. Range of the seven leaf gas exchange traits along with the ANOVA and CV values from the 62 Ethiopian *O. longistaminata* accessions

Trait	Minimum	Maximum	Mean	Std Dev	MSE (122)	CV	F test (61)
Pn	8.29	18.47	12.84	2.46	5.27	17.84%	3.48***
Gs	0.05	0.1	0.07	0.01	0.0001	13.94%	4.12***
Ci	12.14	134.65	79.34	26.8	252.07	17.61%	2.45***
E	2.18	4.05	3.14	0.47	0.13	11.38%	5.17***
CE	0.07	0.57	0.15	0.06	0.008	57.77%	1.53*
WUE	3.17	5.17	4.08	0.46	0.23	11.65%	2.75***
WUEi	134.92	221.45	176.23	18.7	444.74	11.96%	2.28***

* (significant at p<0.05), *** (highly significant at p<0.001),

Key; Pn (Photosynthetic rate), gs (stomatal conductance), Ci (intercellular CO_2 concentration), E (Transpiration rate), CE (Carboxylation efficiency), WUE (Water Use Efficiency), WUEi (Intrinsic Water Use Efficiency), MSE (Mean square error), CV (Coefficient of Variation) and numbers in parenthesis represent degree of freedom

4.1.7 Pairwise correlation analysis of photosynthetic traits

The Pearson pairwise correlation matrix for the seven leaf gas exchange traits revealed their inter-relationship. The rate of photosynthesis was positively and significantly ($P < 0.001$) correlated with all other traits except Ci which showed a highly significant ($P < 0.001$) and negative association (Table 18). Ci had also a highly significant ($P < 0.001$) and negative correlation with CE, WUE and WUEi. gs was negatively and highly significantly ($P < 0.001$) correlated with WUEi.

Furthermore, highly significant ($P < 0.001$) and positive associations was observed between gs and E, CE with both WUE and WUEi and among WUE and WUEi. Similarly, a positive but significant ($P < 0.05$) relationship was shown only between E and CE. gs was found to have a positive but non-significant correlation with Ci, CE and WUE. E and WUE were also associated positively but not in a statistically significant manner. Moreover, E had a non-significant and negative correlation with Ci and WUEi (Table 18).

Table 18. Pairwise Pearson's Correlation Coefficients for 62 Ethiopian *O. longistaminata* accessions

Traits	**Pn**	**Gs**	**Ci**	**Tr**	**CE**	**WUE**	**WUEi**
Pn	1.00	0.83^{***}	-0.44^{***}	0.80^{***}	0.54^{***}	0.61^{***}	0.48^{***}
Gs		1.00	0.02	0.93^{***}	0.24	0.15	-0.09
Ci			1.00	-0.04	-0.65^{***}	-0.72^{***}	-0.84^{***}
Tr				1.00	0.29^{*}	0.02	-0.04
CE					1.00	0.54^{***}	0.60^{***}
WUE						1.00	0.87^{***}
WUEi							1.00

* Significant at p<0.05; ** Highly significant at p<0.01; *** Highly significant at p<0.001

Key; Pn (Photosynthetic rate), gs (stomatal conductance), Ci (intercellular CO_2 concentration), E (Transpiration rate), CE (Carboxylation efficiency), WUE (Water Use Efficiency) and WUEi (Intrinsic Water Use Efficiency)

4.2 Genetic diversity and relationships analysis

4.2.1 SSR genotyping and extent of genetic diversity

From the 430 wild and cultivated *Oryza* accessions (Table 7 and Table 8), a total of 440 alleles ranging from 2 (RM22 and RM171) to 19 (RM225) were genotyped with an average of 6.57 alleles per locus (Table 19). Based on their polymorphic information content (PIC), 26 (38.9%) of the 67 SSR markers used in this study were highly informative (PIC >0.5).

Table 19. Number of allels and genotypes, observed heterozygosity, gene diversity and polymorphism of the 67 SSR markers over 430 rice accessions

Marker	**Major Allele Frequency**	**Genotype No**	**Allele No**	**Gene Diversity**	**Heterozygosity**	**Polymorphism**
RM 1	0.336	20	7	0.76	0.407	0.721
RM5	0.74	14	9	0.438	0.23	0.422
RM6	0.37	11	8	0.678	0.225	0.612
RM 11	0.755	10	6	0.405	0.323	0.377
RM14	0.612	4	3	0.477	0.022	0.365
RM17	0.543	9	4	0.594	0.366	0.526
RM19	0.922	6	4	0.147	0.044	0.142
RM21	0.466	14	7	0.686	0.144	0.639
RM22	0.944	2	2	0.106	0	0.101
RM23	0.44	12	8	0.657	0.18	0.595
RM24	0.667	6	4	0.472	0.155	0.399
RM25	0.408	7	6	0.67	0.003	0.602
RM26	0.859	5	4	0.25	0.002	0.231
RM29	0.915	5	5	0.159	0	0.152
RM30	0.737	10	7	0.436	0.044	0.414
RM31	0.496	9	7	0.664	0.007	0.617
RM42	0.461	6	5	0.611	0.005	0.53
RM44	0.442	15	8	0.69	0.099	0.638
RM49	0.863	7	5	0.244	0.166	0.228
RM50	0.385	7	5	0.667	0.117	0.598
RM55	0.804	6	4	0.335	0.215	0.311
RM60	0.953	3	3	0.09	0	0.087
RM80	0.789	16	9	0.369	0.13	0.359

Table 19 continued

Marker	Major Allele Frequency	Genotype No	Allele No	Gene Diversity	Heterozygosity	Polymorphism
RM84	0.541	20	11	0.661	0.235	0.633
RM 105	0.373	8	5	0.688	0.051	0.625
RM 110	0.563	9	6	0.518	0.21	0.413
RM 118	0.656	10	7	0.543	0.052	0.519
RM125	0.674	11	5	0.506	0.358	0.47
RM126	0.735	5	4	0.421	0.196	0.379
RM129	0.579	7	6	0.509	0.163	0.405
RM133	0.485	8	7	0.536	0.031	0.428
RM134	0.824	6	4	0.295	0.227	0.261
RM136	0.814	7	7	0.319	0.276	0.295
RM144	0.81	7	5	0.325	0.071	0.299
RM152	0.499	13	8	0.651	0.515	0.597
RM153	0.901	11	6	0.185	0.084	0.179
RM154	0.396	11	5	0.684	0.276	0.621
RM159	0.894	10	7	0.197	0.039	0.19
RM161	0.922	5	4	0.147	0.003	0.142
RM162	0.902	4	3	0.179	0.002	0.167
RM164	0.505	10	7	0.64	0.087	0.581
RM167	0.813	8	8	0.313	0.005	0.28
RM171	0.91	2	2	0.164	0	0.15
RM180	0.561	11	5	0.597	0.157	0.539
RM184	0.839	18	13	0.29	0.249	0.281
RM205	0.746	15	7	0.427	0.198	0.409
RM206	0.605	20	13	0.597	0.194	0.569
RM207	0.731	17	14	0.455	0.236	0.443
RM209	0.356	18	12	0.716	0.106	0.664
RM211	0.514	8	5	0.611	0.253	0.54
RM212	0.709	14	10	0.457	0.145	0.416
RM213	0.884	7	6	0.215	0.095	0.21
RM215	0.848	6	5	0.265	0.005	0.243
RM216	0.788	11	7	0.358	0.147	0.332
RM219	0.912	3	3	0.161	0	0.15
RM220	0.47	13	8	0.658	0.229	0.601
RM224	0.397	11	6	0.666	0.248	0.599
RM225	0.37	26	19	0.757	0.082	0.723
RM251	0.86	9	8	0.253	0.096	0.242
RM263	0.448	14	5	0.722	0.396	0.687
RM307	0.863	8	7	0.249	0.002	0.239

Table 19 continued

Marker	Major Allele Frequency	Genotype No	Allele No	Gene Diversity	Heterozygosity	Polymorphism
RM322	0.872	9	6	0.232	0.099	0.221
RM333	0.517	14	9	0.574	0.177	0.488
RM335	0.532	8	7	0.53	0.018	0.423
RM337	0.53	7	4	0.552	0.232	0.458
RM475	0.364	6	5	0.7	0.073	0.646
RM6344	0.522	14	9	0.616	0.183	0.551
Mean	**0.66**	**9.9**	**6.57**	**0.46**	**0.14**	**0.42**

At group level, Na was lowest (0.72) in the *O. longistaminata* accession from Niger and highest (4.35) in all *O. longistaminata* populations of Ethiopia (Table 20). Similarly, Ne ranged from 0.72 in *O. longistaminata* from Niger to 1.9 of *O. glumaepatula*. Expected heterozygosity (He) was in the range of 0.015 for *O. officinalis* to 0.421 of *O. glumaepatula*. When all the assessed groups were pooled as one, 1.96 No, 1.47 Ne, 0.13 Ho and 0.24 He were obtained. With the exception of the out-groups, all populations showed lower observed than expected heterozygosity. The study also showed a high Fixation Index (F) in both domesticated and wild AA-genome *Oryza* species. Based on their PIC value, *O. officinalis* had the lowest (2.94) and *Japonica* showed the highest (91.18).

Table 20. Group based genetic diversity parameters of *Oryza longistaminata* in Ethiopia and other rice types

Populations	N	Na	Ne	Ho	He	F	PIC
Gambella 1	30	2.088	1.521	0.139	0.267	0.457	0.65
Gambella 2	30	2.309	1.575	0.160	0.279	0.447	0.71
Gambella 3	30	2.088	1.518	0.150	0.251	0.424	0.60
Gambella 4	30	2.162	1.621	0.119	0.271	0.545	0.71
Gambella 5	30	1.897	1.353	0.132	0.199	0.309	0.59
Gambella 6	30	2.059	1.505	0.161	0.255	0.357	0.60
Amhara1	30	2.206	1.525	0.150	0.274	0.381	0.72
Amhara 2	30	2.294	1.518	0.145	0.271	0.427	0.74

Table 20 continued

Populations	**N**	**Na**	**Ne**	**Ho**	**He**	**F**	**PIC**
Amhara 3	30	2.132	1.465	0.167	0.259	0.402	0.72
Amhara 4	30	2.221	1.438	0.171	0.236	0.371	0.71
Amhara 5	30	2.029	1.383	0.132	0.216	0.395	0.59
Amhara 6	30	1.574	1.282	0.135	0.150	0.185	0.37
Pooled Ethiopian *O. longistaminata*	***360***	***4.35***	***1.91***	***0.15***	***0.38***	***0.53***	***0.84***
O. barthi	3	1.809	1.659	0.108	0.305	0.601	0.63
O.glumaepatula	6	2.426	1.991	0.095	0.421	0.753	0.87
O. officinalis	1	0.750	0.750	0.029	0.015	-1.000	0.03
O. nivara	6	2.397	1.869	0.107	0.402	0.720	0.85
O. rufipogon	6	2.368	1.892	0.110	0.400	0.698	0.84
Open-polinated *O. rufipogon*	4	2.059	1.771	0.119	0.378	0.688	0.82
Japonica	17	2.485	1.659	0.045	0.329	0.873	0.91
Indica	18	2.485	1.801	0.054	0.385	0.845	0.87
Japonica weedy type	1	1.044	1.044	0.074	0.037	-1.000	0.74
O. rhyzomatis	2	1.426	1.374	0.103	0.217	0.526	0.43
O. eichengeri	2	1.147	1.135	0.162	0.107	-0.511	0.22
O. latifola	1	1.162	1.162	0.279	0.140	-1.000	0.28
O. australiansis	2	1.132	1.116	0.154	0.108	-0.431	0.22
O.longistaminata from Niger	1	0.721	0.721	0.088	0.044	-1.000	0.08
Mean		**1.96**	**1.47**	**0.13**	**0.24**	**0.22**	**0.58**

Key; N (Sample size), Na (Actual number of alleles), Ne (Effective number of Alleles), Ho (Observed Heterozygosity), He (Expected Heterozygosity), F (Fixation Index), PIC (Polymorphic Information Content)

4.2.2 Comparative polymorphism among *O. longistaminata* populations and/or with other rice groups

In this study, majority of Ethiopian *O. longistaminata* populations were genetically diverse and highly polymorphic. Although most of the Ethiopian *O. longistaminata* populations had equivalent Na, Ne, Ho, He, F and PIC values, Gambella5 as compared to the rest Gambella populations and Amhara6 in contrast to the remaining Amhara populations showed the least genetic diversity parameters (Table 20).

When the 12 *O. longistaminata* populations from Ethiopia were pooled as one, the SSR Polymorphism showed; Na = 4.35, Ne = 1.91, Ho = 0.15, He = 0.38, F = 0.53 and PIC = 84

(Table 20). Despite the 12 *O. longistaminata* populations had less F and PIC values, they showed higher actual and effective number of alleles, observed heterozygosity and gene diversity than the *Indica* and *Japonica* subspecies of the cultivated rice (*O. sativa*).

For almost all of the genetic diversity estimates, the Ethiopian *O. longistaminata* populations were superior to all AA genome wild *Oryza* species. Likewise, the whole out groups were less diverse than the *O. longistaminata* populations of Ethiopia.

4.2.3 Genetic distance and identity among the twelve *O. longistaminata* populations of Ethiopia

For each pair of the twelve Ethiopian *O. longistaminata* populations genetic distance and genetic identity were determined using Nei (1972). The highest distance (0.419) and lowest genetic identity (0.658) were observed between populations of Gambella 4 and Amhara 6. In contrast, the smallest genetic distance (0.089) and highest identity (0.915) were found between Gambella 5 and Gambella 6 (Table 21).

Table 21. Nei's Unbiased measures of genetic distance (below diagonal) and genetic identity (above diagonal) among each pair of the twelve *O. longistaminata* populations

	Gambela 1	Gambela 2	Gambela 3	Gambela 4	Gambela 5	Gambela 6	Amhara 1	Amhara 2	Amhara 3	Amhara 4	Amhara 5	Amhara 6
Gambela 1	****	0.913	0.873	0.889	0.847	0.859	0.847	0.838	0.809	0.782	0.785	0.693
Gambela 2	0.091	****	0.904	0.869	0.810	0.821	0.817	0.799	0.794	0.782	0.786	0.676
Gambela 3	0.135	0.101	****	0.861	0.827	0.826	0.822	0.816	0.810	0.774	0.769	0.703
Gambela 4	0.117	0.141	0.149	****	0.868	0.852	0.826	0.830	0.798	0.743	0.746	0.658
Gambela 5	0.166	0.210	0.189	0.141	****	0.915	0.805	0.805	0.804	0.761	0.789	0.679
Gambela 6	0.152	0.197	0.192	0.160	0.089	****	0.805	0.792	0.786	0.772	0.773	0.688
Amhara 1	0.166	0.202	0.196	0.191	0.217	0.217	****	0.900	0.869	0.768	0.851	0.801
Amhara 2	0.176	0.224	0.203	0.187	0.216	0.233	0.106	****	0.869	0.777	0.852	0.739
Amhara 3	0.212	0.230	0.211	0.225	0.219	0.241	0.141	0.141	****	0.783	0.827	0.790
Amhara 4	0.245	0.246	0.256	0.297	0.274	0.259	0.264	0.252	0.244	****	0.823	0.777
Amhara 5	0.242	0.241	0.263	0.294	0.237	0.258	0.161	0.160	0.190	0.194	****	0.773
Amhara 6	0.367	0.392	0.353	0.419	0.387	0.374	0.222	0.302	0.236	0.252	0.258	****

4.2.4 Principal component analysis

A scatter plot of the first two principal components for the 430 accessions showed five clusters (Fig. 13). The first and second axes accounted 8.96% and 7.26% of the molecular variance, respectively. When only the 12 groups of *O. longistaminata* from Ethiopia were considered, a clear geographic pattern emerged. *Oryza longistaminata* accessions of the six Gambella groups were located in the right half of Quadrant I and left half of Quadrant II. Whereas, *O. longistaminata* accessions of the six Amhara groups resided in Quadrant III. Despite very few Amhara1 and Amhara2 accessions positioned at the very bottom of Quadrant II, all of the Amhara accessions were residing in Quadrant III. Specifically, the whole accessions from Amhara6 formed their own distinct cluster near the middle of Quadrant III. Coordinate 1 of the PCA separated the whole Amhara and majority of the Gambella accessions from the rest of all wild and cultivated *Oryza* members.

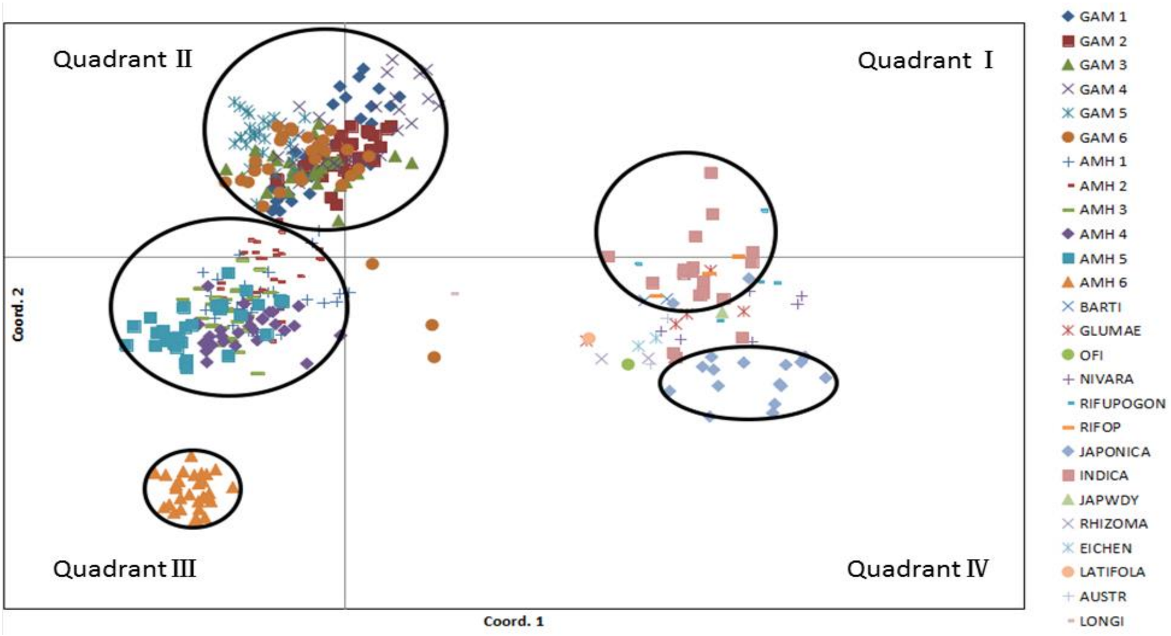

Figure 13. Two-dimensional scaling of principal component analysis for 26 *Oryza* groups from 67 SSR markers genotypic data

4.2.5 Cluster analysis

Clustering analysis based on UPGMA method at 1,000 boot strapping grouped the 26 rice groups into three main clusters comprising the genotypically unknown Japonica weedy type in Cluster I, AA genome *Oryza* species in Cluster II and the *O. officinalis* complex members in Cluster III (Figure 14). The AA genome *Oryza* species in Cluster II were further partitioned as sub-cluster A and sub-cluster B. In the *O. longistaminata* group (sub-cluster A) the accession from Niger was separated from the 12 Ethiopian *O. longistaminata* groups. Moreover, the Amhara groups differed from the Gambellas and undergone to continuous partitioning on the basis of their geographic provenance.

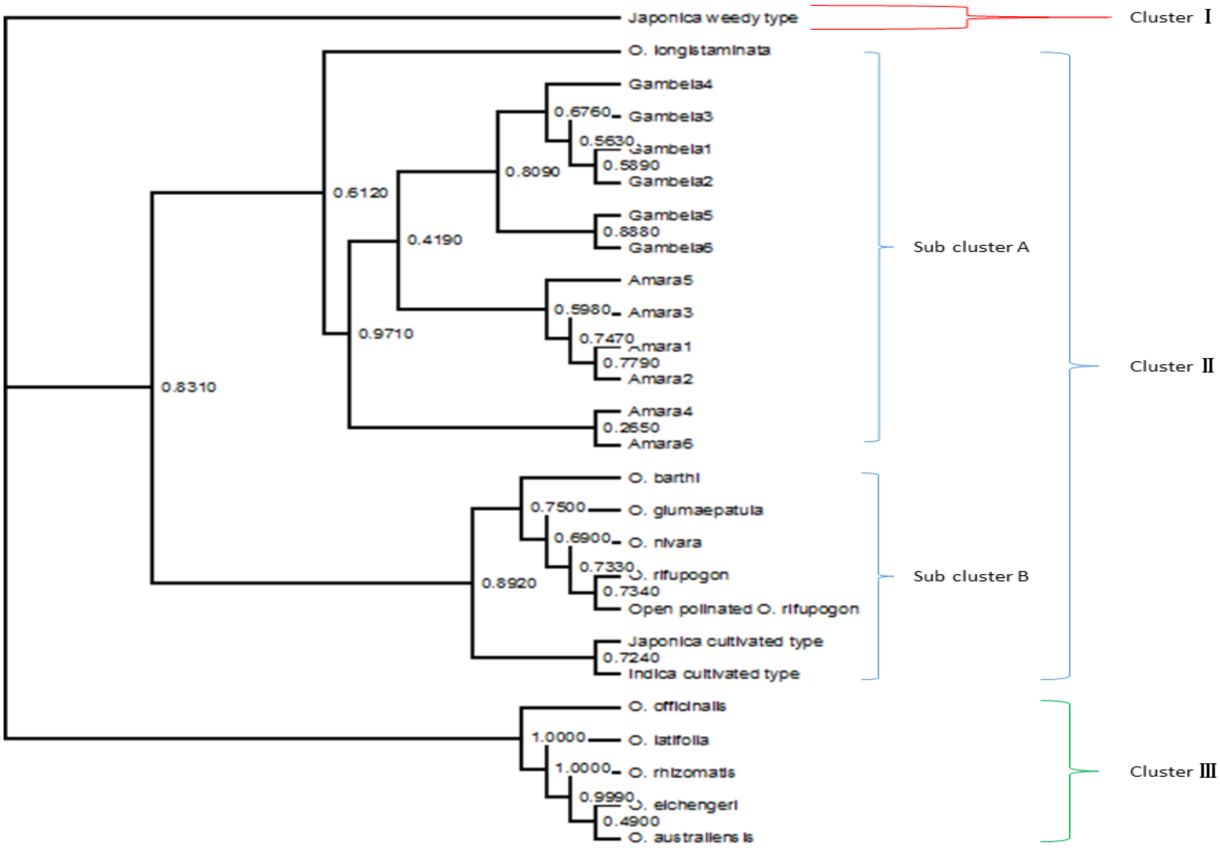

Figure 14. UPGMA tree of 26 rice groups based on corrected genetic distance (numbers indicated the bootstrapping values)

Unlike the dendrogram generated by UPGMA analysis, the hierarchical clustering of the 430 accessions showed 2 main groups categorizing all *O. longistaminata* accessions in one and the rest *Oryza* members in the other cluster (Figure 15). As sub-cluster A of Fig. 14, the main cluster of *O. longistaminata* accessions later separated the Amhara groups from their Gambella counterparts. Regarding population admixture, almost all *O. longistaminata* accessions were residing in their own populations. However, some accessions of Gambella6 with Gambella5, Gambella2 with Gambella4 and Amhara2 with Amhara1 were admixtured. Interestingly, no *O. longistaminata* accession was found to be admixed in any other cluster of either the AA or non-AA genome wild and cultivated rice.

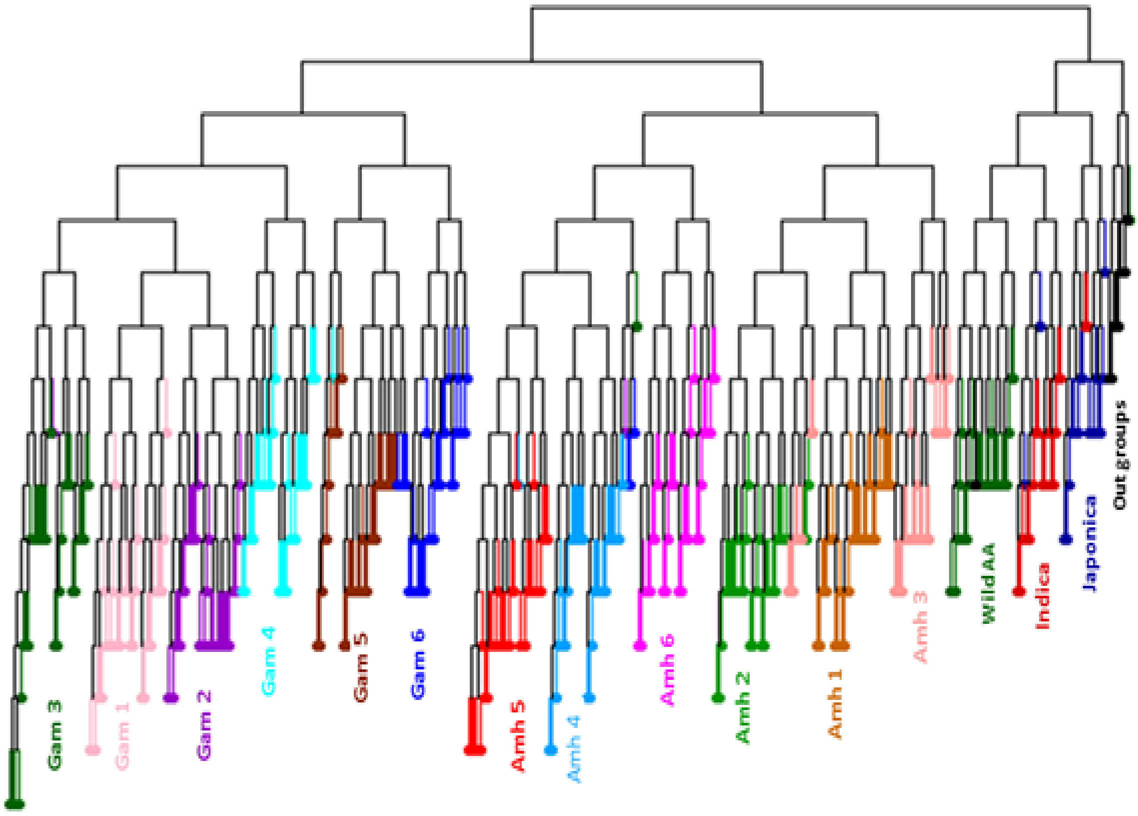

Figure 15. Dendrogram based on Ward’s minimum variance for 430 rice accessions

4.2.6 Population structure

STRUCTURE analysis suggested that the optimal number of subpopulations for all individuals in the study is three (Figure 16). At K=3, all Gambella populations of *O. longistaminata* comprised one cluster, all Amhara's *O. longistaminata* hold the second, and all the wild and cultivated AA genome *Oryza* species along with the out-groups occupied the third cluster (Figure 17).

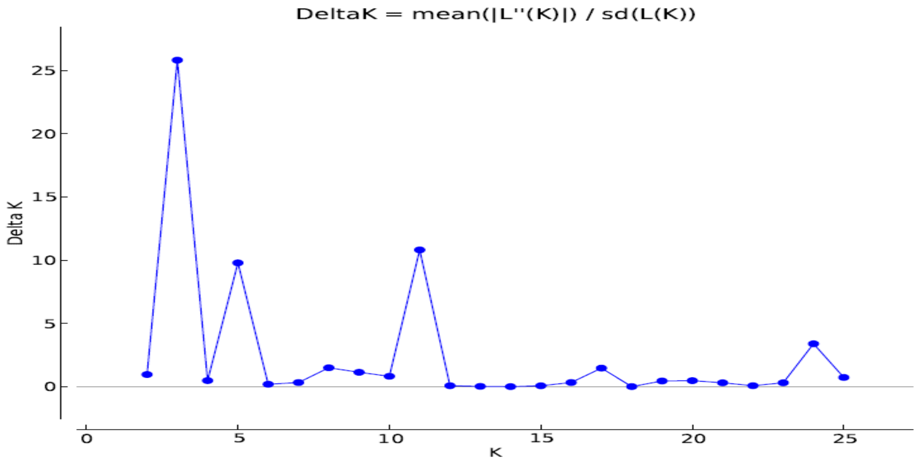

Figure 16. The modal value ΔK at k=3 (the highest peak)

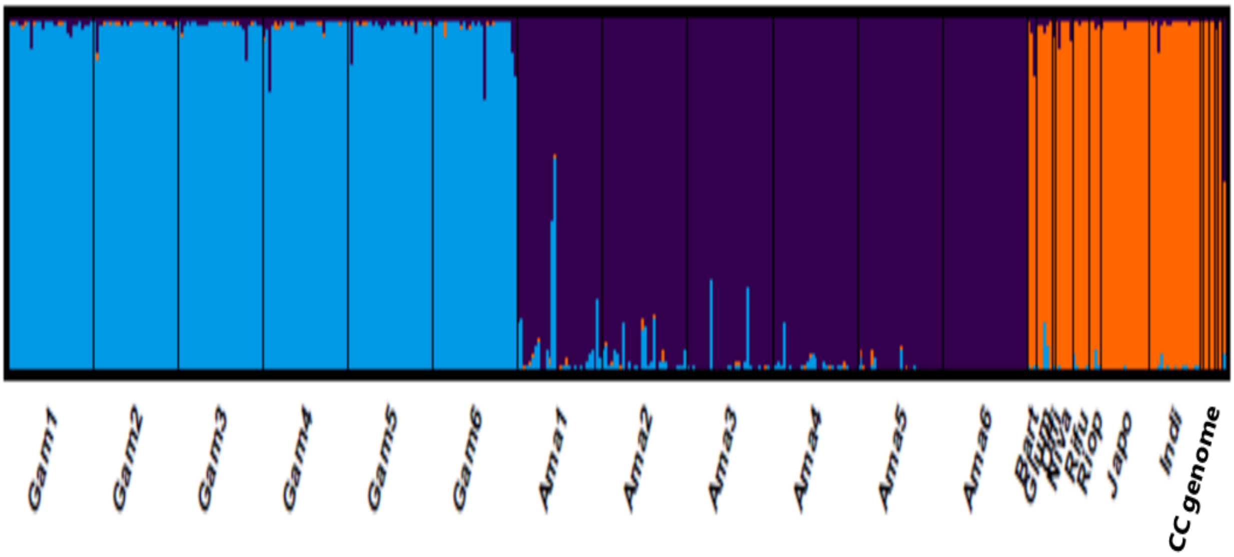

Figure 17. STRUCTURE bar graph of *Oryza* groups at k=3

4.2.7 Genetic differentiation and Variance Analysis of *O. longistaminata* in Ethiopia

From the 67 SSR loci, the 12 *O. longistaminata* populations from Ethiopia showed a highly significant (P> 0.001) genetic differentiation (F_{st}) = 0.299, Wrights fixation index (F_{is}) = 0.547 and Total inbreeding coefficient (F_{it}) = 0.683 (Table 22). In particular, the strongly significant F_{st} value showed a higher divergence or differentiation among the 12 *O. longistaminata* populations of Ethiopia. The indirect estimation of gene flow using F_{st} showed the average number of migrants per generation (N_m) = 0.586 (Table 22). Based on those 12 *O. longistaminata* populations, the Analysis of Molecular Variance (AMOVA) showed 30 % of the variance among populations, 38% variance among accessions and 32 % variance within accessions (Table 23).

Table 22. *F* statistic and probability P (rand > = data) of the 12 Ethiopian *Oryza longistaminata* populations

F-Statistics	Value	P
F_{st}	0.299	0.001
F_{is}	0.547	0.001
F_{it}	0.683	0.001
Nm	0.586	

Key; *Fst* (Genetic differentiation), *Fis* (Wrights fixation index), *Fit* (Total inbreeding coefficient) and Nm (Gene flow)

Table 23. Analysis of molecular variance (AMOVA) of the 12 Ethiopian *Oryza longistaminata* populations at three levels

Source	Df	SS	MS	Est. Var.	Percentage
Among populations	11	3103.608	282.146	4.435	30%
Among accessions	348	5592.333	16.07	5.684	38%
Within accession	360	1693.00	4.703	4.703	32%
Total	**719**	**10388.942**		**14.821**	**100%**

Key; Df (Degrees of Freedom); SS (Sum of squares); MS (Mean squares); Est. Var. (estimated variance) and % (proportion of genetic variability)

Genetic diversity of the Ethiopian *O. longistaminata* groups was further assessed on the basis of their geographic region (Table 24). The mean allelic patterns were Na = 3.41, Ne = 1.78 and He = 0.34 for Gambella and Na = 3.44, Ne = 1.81 and He = 0.35 for Amhara regional states of Ethiopia. Pairwise genetic differentiation, Nei genetic distance and identity among the Amhara and Gambella regions were 0.092, 0.11 and 0.896, respectively.

Table 24. Allelic diversity, heterozygosity and pairwise differentiation among *O. longistaminata* accessions from Amhara and Gambella regions of Ethiopia

Region	**Gambella**	**Amhara**
Actual number of alleles (Na)	3.41	3.44
Effective number of alleles (Ne)	1.78	1.81
Expected heterozygosity (He)	0.34	0.35
Pairwise Nei genetic distance	0.11	
Pairwise Nei genetic identity	0.896	
Pairwise genetic differentiation (Fst)	0.092	

4.3 RADSeq analysis

4.3.1 RAD tag generation and de novo assembly

In this study, a multiplexed and reduced genome sequencing strategy was used to explore genetic diversity and regional differentiation in Ethiopian *O. longistaminata*. Whole genome RAD tag libraries at *PstI* cut sites were created from individually barcoded and sequenced DNA fragments of 87 *O. longistaminata* accessions of 6 different populations. The processed_radtags program retained 98.2% (652,611,547) reads of the Fastq data (Table 25). However, the remaining 1.1%, 0.1% and 0.5% represented the ambigious barcode drops, low quality read drops and ambiguous rad tags drops, respectively.

Table 25. Sequence types generated after running the process_radtags program of Stacks

Type of Sequences	**No of reads**	**Proportion of reads**
Retained reads	652,611,547	0.982
Ambiguous bar code drops	7454638	0.011
Low quality read drops	624658	0.001
Ambiguous rad tags drops	3639933	0.005
Total	664330776	100

4.3.2 Phylogey and Discriminant Analysis of Principal Components (DAPC)

Reduced genome multiple alignments by the UNEAK pipeline in TASSEL was used to identify single nucleotide polymorphism markers. From such single nucleotide polymorphisms (SNPs), phylogenetic relationship among 87 *O. longistaminata* accessions of Ethiopia was visualized by a neighbor-joining tree (Figure 18). The phylogenetic tree generated in this study was principally organized by the geographic origins of the accessions. For instance, the most distantly related Amhara accessions were in a clear and well-separated cluster. However, accessions from the five Gambella populations were found to be admixtured.

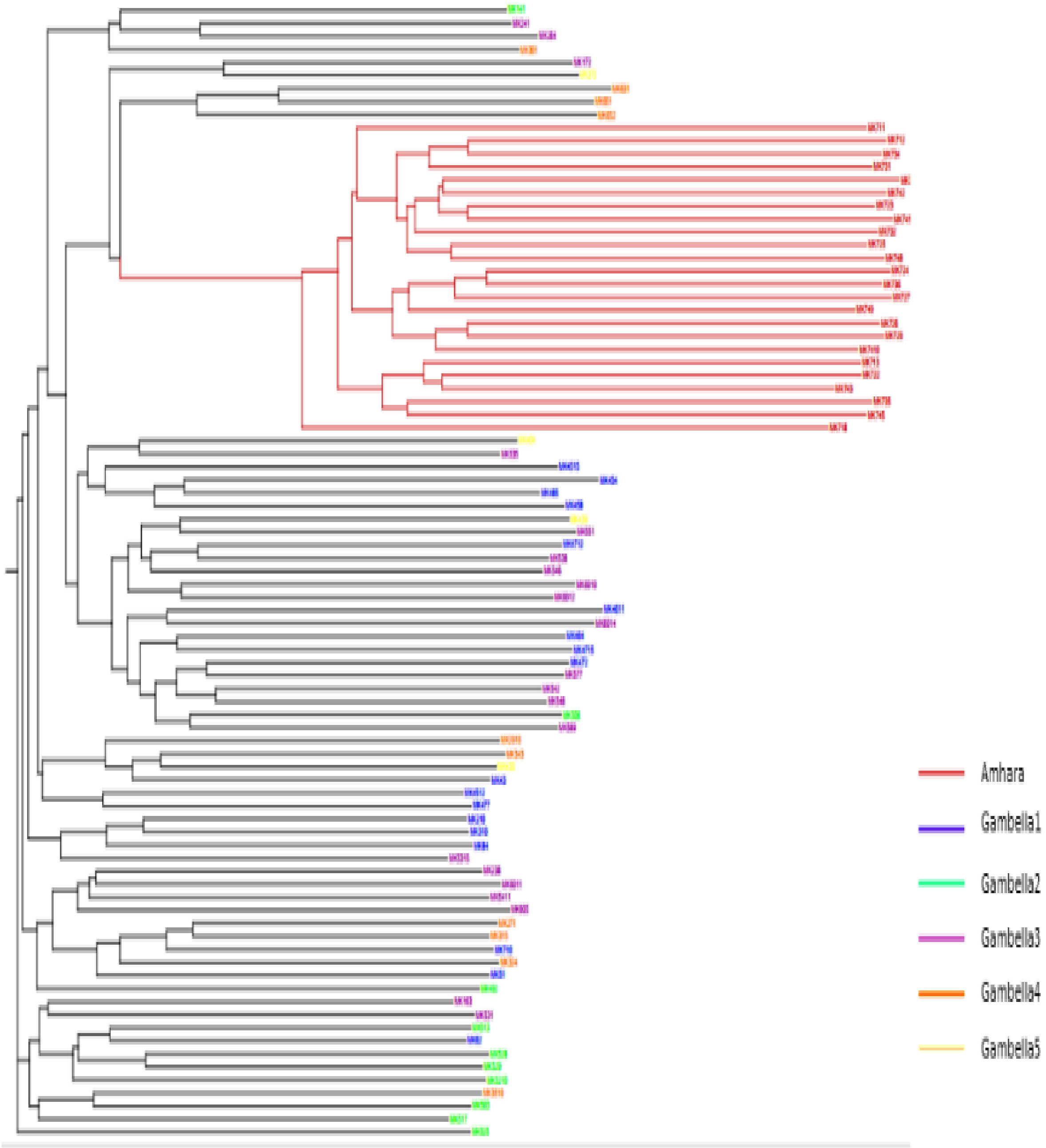

Figure 18. 30,256 RAD Tag SNPs based phylogenetic tree of 87 *O. longistaminata* accessions of Ethiopia. (Color of accessions' code illustrated the respective populations)

Like the Neighbor-joining tree, discriminant analysis of principal components (DAPC) from the adegenet R package of the 87 *O. longistaminata* accessions of Ethiopia showed lack of genetic differentiation among the Gambella accessions (Figure 19). However, accessions from Amhara were the most differentiated than those of Gambella's.

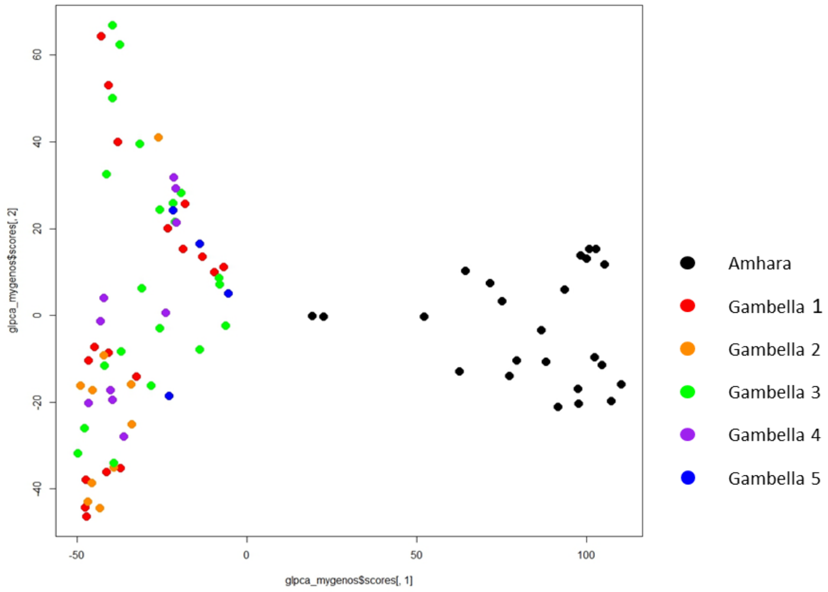

Figure 19. Accession based DAPC from 30,256 nuclear SNPs. Each dot represents a given *O. longistaminata* accession and the different colors stands for population of each accession

4.3.3 Region based differentiation

The Baysian Information Complex (BIC) versus number of clusters graph of this study showed a clear decrease of BIC until k = 2 clusters and then after BIC increased (Figure 20). In this case, the elbow in the curve also matched with the smallest BIC and clearly indicated retaining of the 2 clusters.

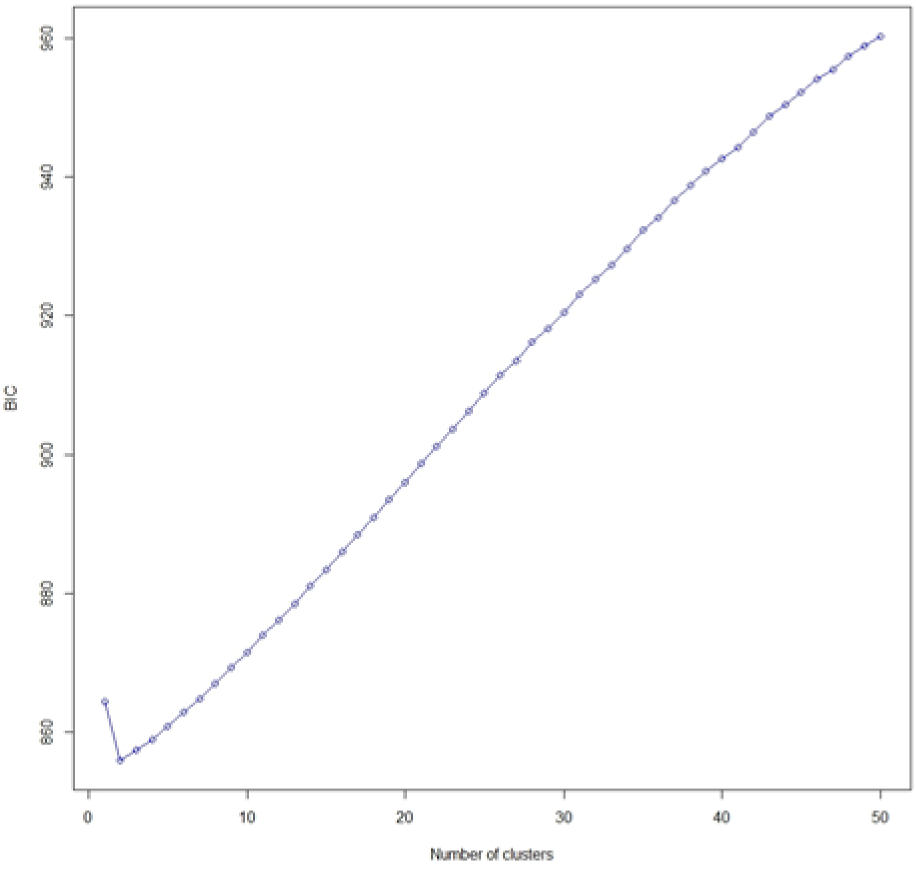

Figure 20. Graph of BIC versus clusters number of the Ethiopian *O. longistaminata* accessions from the RADSeq SNPs

From the analysis of the resulting multiple alignment, a clear geographical stratification of the Amhara population from the Gambellas along with evidence of admixture between the Gambella populations were observed. This pattern of clustering was totally alike with grouping of the whole accessions in to two clusters (Fig. 20).

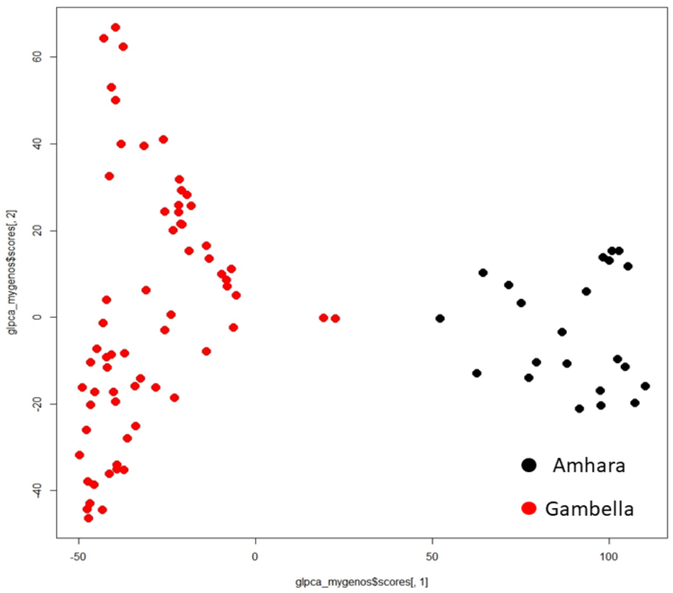

Figure 21. Region based DAPC plot generated by adegenet. Where, Gambella populations are designated in red and Amharas in black color

The RADSeq analysis in the present study showed the Nei's genetic distance for the Amhara and Gambella regions as 0.2436 and 0.2406, respectively. Pairwise comparison or genetic differentiation (*Fst*) between the Amhara and Gambella regions was estimated as 0.0607 (Table 25). Moreover, the genome wide pattern of Wright's inbreeding coefficient (*Fis*) value across the two regions was 0.3272.

Table 26. Region level genetic distance along with pairwise comparison of genetic differentiation (F_{st}) and genetic identity (F_{is}) values from the RADSeq analysis

	Amhara	Gambella
Nei's genetic distance	0.2436 ± 0.0011	0.2406 ± 0.0009
Fst with Nei's correction	0.0607 ± 0.0005	
Wright's inbreeding coefficient (*Fis*)	0.3272 ± 0.0018	

5. Discussion

5.1 Morphological and leaf gas exchange traits

5.1.1 Morphological traits of *O. longistaminata* and their implication for crossing

The eight quantitative traits evaluated in this study showed a very wide range of variation (Table 10). Such high level of diversity can be an indication for the great candidacy of the Ethiopian *O. longistaminata* accessions in rice improvement programs. In this study, analysis of variance in all of the eight quantitative characteristics also revealed a highly significant difference among the entire accessions. Such a rich genetic diversity status implicates presence of a substantial amount of variability in the assessed accessions and a need for an effective selection (Chouhan *et al*., 2014).

Phenotypic traits such as Plant Height, first Leaf Width, second Leaf Width and third Leaf Width were found to be traits with lower coefficients of variation (Table 10). Such a low variation in these traits suggests that they are more likely to be highly heritable traits (Molosiwa, 2012). It is well known that selection of varieties based on high heritability of traits make it easy for breeders to exploit their knowledge and skill in the transgressive breeding program (Chouhan *et al*., 2014). Therefore, utilization of such traits for hybridization of cultivated rice and *O. longistaminata* accessions and selection of desirable segregants could improve the cultivated rice gene pool.

5.1.2 Genetic variation for an ease of selection

The total variation from the sum of the two principal components of this study is relatively lower than 82.70% of total variation reported among 32 upland rice varieties (Lasalita-Zapico *et al*., 2010) and 82.1% of total variation from the first two principal components reported by Rajiv *et al*. (2010).

Clustering in this study partitioned the 62 Ethiopian *O. longistaminata* accessions in to six non-monogenic clusters having 3 to 29 accessions. The mean value for the clusters indicated existence of considerable differences in the mean values of different traits. The extreme mean values of the traits were also observed to fall in different clusters. This clustering pattern indicated a wide diversity of accessions between different groups and their close relationship within a cluster. It is, therefore, logical to expect a higher genetic diversity from accessions of different clusters than within a cluster.

According to Singh *et al*., (2011), accessions in clusters with highest mean values for Plant Height, Panicle Length and Leaf Width are quite promising for breeding. Thus, accessions in cluster II which were found to have highest values for six quantitative traits would be ideal for many of the traits under study. To further widen the gene pool of *O. longistaminata* or achieve a wide spectrum of variation, cluster II could be hybridized with clusters VI and V.

From the cluster analysis of the 62 *O. longistaminata* accessions, a tendency towards the effect of population of origin was noticed. Indeed, it is noteworthy that clustering pattern of accessions is consistent with their geographic proximity and ecological factors (Amsalu and Endeshaw, 1999). However, inclusion of fewer accessions with an extremely high or low mean values to Cluster I, V and VI limited manifestation of population effect. Besides, Gambella3 had accessions distributed patchily over five of the six clusters (Table 15). This scenario reflects a wide variation among accessions of a population or insignificant population differentiation from a high gene flow. Compared with accessions of the other Gambella populations, most of the accessions from Amhara3 were grouped into cluster II. This observation showed that accessions from Amhara population were relatively less variable than those from Gambella. Moreover, distinctness of the Amhara accessions could be associated with unique genetic, altitudinal and climatic factors (Melaku *et al*., 2013).

5.1.3 Relationships based on populations of origion

The Mahalanobis distance (D^2) value computed among the six populations revealed distinctness of Amhara3. For instance, the maximum inter-cluster distance was found between Amhara3 and Gambella3, and Amhara3 and Gambella2. These observations implicated a need to consider accessions from those distant populations for their favorable genetic resources and improvement of yield and yield related traits of the cultivated rice (Chouhan *et al*., 2014). The smallest genetic distance recorded among the Gambella populations further exemplified the impact of geographic proximity to the higher gene flow or the low genetic diversity among accessions of such Gambella populations.

The cladogram constructed on the basis of the eight quantitative traits also confirmed separation of Amhara3 from the other Gambella populations (Fig. 11). Though, the dendrogram showed a close relationship between Gambella populations, it did not distinguish them in terms of their geographic distance. Therefore, seven SSR markers were used to run a PCR on each replicates of the sixty two *O. longistaminata* accessions. The UPGMA generated from those microsatellite markers clearly implicated closer genetic similarity among Gambella1, Gambella2, Gambella3 and Gambella4 populations and distinctness of the geographically distant Gambella5 and Amhara3 populations both from each other as well as other populations (Fig. 12). In general, the dendrograms from both morphological and SSR data showed higher genetic diversity among accessions of the Ethiopian *O. longistaminata* populations and the different gene pools of the Amhara and Gambella regions of Ethiopia. Such distinct population structuring was entirely in harmony with the genetic diversity study of 320 Ethiopian *O. longistaminata* accessions from Gambella and Amhara populations (Melaku *et al*., 2013).

5.1.4 Leaf gas exchange traits of *O. longistaminata* and their implications for physiological trait improvement

Except Carboxylation Efficiency which had a significant difference ($P < 0.05$), all of the other photosynthesis related traits showed a highly significant variation ($P < 0.001$). Kiran *et al*., (2013) also reported similar pattern of high variations in the rate of photosynthesis, stomatal conductance, transpiration rate and internal CO2 concentration amongst different *Oryza* accessions. Such highly significant differences in turn indicated high genetic variability in the Ethiopian *O. longistaminata* accessions.

In this study, the Pn record across the 62 Ethiopian *O. longistaminata* accessions showed an average of 12.84 µmol CO_2 m^{-2} s^{-1}. This mean value was slightly higher than the average Pn of the two *O. longistaminata* accessions evaluated by Kondamudi *et al*., (2016). Mean Pn in the 62 Ethiopian *O. longistaminata* accessions was also higher than mean of ten cultivated rice genotypes at three weeks after heading but lower than the mean values of the same genotypes at both panicle initiation and heading time (Kiran *et al*., 2013). Here a relatively closer Pn value of this study to the panicle initiation stage of the ten cultivated rice genotypes is because of conducting photosynthetic measures nearly at the same developmental stage. The average Photosynthetic rate recorded in this study was also superior to four cultivated rice accessions and a bit less than the mean photosynthetic rate from twenty accessions of seven wild rice species (Kondamudi *et al*., 2016). Comparison among seven wild *Oryza* species of the study from Kondamudi *et al*. (2016), indicated *O. longistaminata* to have the highest average Pn. The considerable genotypic differences and higher Photosynthetic rate observed in this study can thus be indicative for the use of *O. longistaminata* accessions as breeding resources for the improvement of the poor rate of photosynthesis displayed by the cultivated rice varieties.

In many studies, stomatal size and their frequency have been used as indicators of water loss (Venora and Calcagno 1991). And perhaps, wild *Oryza* species are following the route of environmental adaptation with large sized and fewer stomatas operating in patches and tandem (Kondamudi *et al*., 2016). The mean transpiration rate (E) of the 62 *O. longistaminata* accessions of this study, was almost less by half than the average E from 40 ILs (BC3F8) and their parents KMR3 variety of *O. sativa* and *O. rufipogon* (WR120) (Haritha *et al*., 2017). From this comparison, it is noticed that the Ethiopian *O. longistaminata* accessions were more efficient in minimizing water loss. The relatively lower E value recorded in this study could be associated with the relatively warmer and less humid field condition. However, leaf feature of the African wild rice (*O. longistaminata*) is capable of maintaining carbon gain via conserving water (Giuliani *et al*., 2013). A wide scale physiological assessment on *O. longistaminata* is, therefore, recommended for improving the low drought tolerance nature of most *O. sativa* varieties.

The leaf gas exchange trait called stomatal conductance is influenced by many other complex traits (Takai *et al*., 2006). Therefore, large genotypic difference to gs in this study can be a function of either stomatal density as concluded by Kawamitsu and Agata (1987) or stomatal aperture (Maruyama and Tajima 1990). The mean *g*s recorded in this study was significantly less than the average gs reported by Haritha *et al*., (2017); Kiran *et al*., (2013); Ohsumi et al., (2007a). This very little gs score from the Ethiopian *O. longistaminata* accessions can be explained by their thick and bushy leaves. Therefore, further information on the leaf structural and functional traits of *O. longistaminata* is needed for its utilization in the genetic improvement of rice.

According to Changhai *et al*., (2010), leaf transpiration efficiency is mediated by either stomata regulated transpiration rate or a stomata independent regulation of net CO_2

assimilation rate. Like the other leaf functional traits, transpiration efficiency or Water Use Efficiency (WUE) in this study showed a significant variation. The mean Water Use Efficiency computed from the 62 *O. longistaminata* accessions was more than twice that of both wild and cultivated rice collections (Kiran *et al.*, 2013), higher than the maximum record of IL106 line (Haritha *et al.*, 2017) and almost equivalent to accessions of different *Oryza* species (Giuliani *et al.*, 2013). The maximum transpiration efficiency observed in the assessed *O. longistaminata* accessions can be associated to the higher Mesophyll cell conductance to CO_2 diffusion (gm) per stomatal conductance to gas diffusion (gs) ratio (Flexas *et al.*, 2008, Flexas *et al.*, 2013; Barbour *et al.*, 2010). Similarly, the Intrinsic water Use Efficiency in this study was significantly higher than the Asian perennial wild rice *O. rufipogon* (Haritha *et al.*, 2017), eight wild and cultivated *Oryza* species (Giuliani *et al.*, 2013; Kiran *et al.*, 2013).

In the assessed *O. longistaminata* accessions, the mean intercellular CO_2 (*C*i) was 79.34 µmol/mol. This mean *C*i value was really incomparable with the higher mean values reported by Haritha *et al.*, (2017) and Kiran *et al.* (2013). The low amount of CO_2 in the intercellular air space can tightly be linked with the stomatal resistances to gas diffusion (Nobel 2009). The minimum intercellular CO_2 level recorded in this study can in turn limit the rise of photosynthetic rate.

Carboxylation efficiency (CE) is one of the limiting factors for CO_2 fixation (Niu *et al.*, 2004; Li *et al.*, 2009). In many crops, a strongly significant association between leaf photosynthetic rate and carboxylation efficiency is ensured by the amount and action of an enzyme ribulose-1,5-bisphosphate carboxylase (Jiang and Xu, 2001; Ding *et al.*, 2014). The mean carboxylation efficiency in this study (0.15) was about three fold of the CE reported by Kiran *et al.* (2013). Therefore, higher CE value in the evaluated *O. longistaminata* accessions reflected efficient carboxylation process by RUBISCO for the higher rate of photosynthesis.

Multiple correlations performed between different gas-exchange characteristics and other physiological traits revealed that the rate of photosynthesis is neither dependent on the leaf pigment content nor the leaf thickness (Kiran *et al*., 2013). In aggrement with the report from Haritha *et al*. (2017), associations among the photosynthetic parameters of this study showed positive and highly significant ($P < 0.001$) relationship of the Photosynthetic rate to five of the analyzed leaf gas exchange traits. Yeo *et al*., (1994) also reported positive association between photosynthetic rate and Carboxylation Efficiency. Moreover, Zhao *et al*., (2010) suggested selection of rice genotypes with higher CE value as basic breeding strategy than prioritizing accessions with high Chlorophyll content or leaf thickness. Therefore, the high CE values from *O. longistaminata* accessions raise their value in the development of highly photosynthetic rice breeds.

It is known that Pn is a function of *C*i and it in turn is a derivative of *g*s (Radin *et al*., 1988). A strong association between Pn and *g*s will maintain constant intercellular to atmospheric CO_2 ratio (Sharkey and Raschke 1981; Wong *et al*., 1985). Many other publications reported a strong correlation between *g*s and Pn as this study (Hetherington and Woodward, 2003). In rice, direct relationship of photosynthesis with stomatal aperture and *g*s was also implicated (Hirasawa *et al*., 1988; Kusumi *et al*., 2012).

In agreement with this study, a tight positive and highly significant ($P < 0.001$) correlation between Pn and E was reported by Giuliani *et al*. (2013). In both studies, gs was positively correlated with both Pn and E. Coupling between photosynthesis and transpiration processes evidenced the deterministic role of both gs and leaf structural traits (Barbour et al., 2010; Evans and von Caemmerer 1996; Evans 1999; Evans and Loreto 2000). For instance, trait-to-trait analysis suggested a negative relationship between thick leaf and E and/or a negative correlation between thick leaf and gs (Giuliani *et al*., 2013). These phenomena could be explained by the role of leaf stomatal arrangement in limiting E or leaf structural design to restrict water loss at the expense of photosynthesis.

5.2 Genetic diversity status of the Ethiopian *O. longistaminata* populations and their relatedness with other *Oryza* species

5.2.1 Patterns and promises of the Ethiopian *O. longistaminata* gene-pool

In this study, the Ethiopian *O. longistaminata* accessions showed higher actual and effective number of alleles than both *Japonica* and *Indica* subspecies of the cultivated rice. This observation supported the assumption where the course of domestication from wild to cultivated rice caused a significant loss of alleles (Brar and Khush 2003). The twelve *O. longistaminata* populations sampled from Ethiopia showed higher gene diversity than the rest rice groups of the study. Inclusion of higher number of *O. longistaminata* accessions from different agroecologies of Ethiopia could be a factor for such higher number of alleles and gene diversity.

The highest genetic diversity of those Ethiopian *O. longistaminata* accessions can further be explained by their high PIC value of 0.84. Genotyping of 192 cultivated rice germplasm lines using 61 SSR markers reported a very low PIC than this study (Nachimuthu *et al*., 2015). Actually, number of SSR markers used, repeats size of those SSR markers and their position across the genome will complicate the contrast between studies (Deu et al. 2008). However, the same set of SSR markers used over different AA and non-AA genome rice species of this study implicated the Ethiopian *O. longistaminata* populations as the highly diversified groups. Exceptionally, the whole genetic diversity measures of Amhara6 were inferior to both *Indica* and *Japonica* subspecies and even to other wild AA genome *Oryza* groups. As shown in Table 7, Amhara6 accessions are the only collections from a rice farm. Hence, the very low diversity status of Amhara6 could be associated with the strong weed out of *O. longistaminata* by farmers.

5.2.2 Genetic differentiation and features of natural *O. longistaminata* populations

Among the twelve Ethiopian *O. longistaminata* populations, highly significant F statistic values were reported (Table 21). The *Fst* value in this study showed a wider differentiation among the Ethiopian *O. longistaminata* groups. Such population differentiation is expected as genetic differentiation among populations is not evenly distributed across the sampled area (Cao *et al*., 2006). On the contrary, the AMOVA implicated the highest component of variance (38%) among accessions (Table 22). Thus, higher genetic variation was found within than among the Ethiopian *O. longistaminata* populations. The self-incompatible feature of *O. longistaminata* possibly raised the gene flow and variation among *O. longistaminata* accessions of a given population.

All of the Ethiopian *O. longistaminata* populations showed lower observed than expected heterozygosity (Table 21). This observation inferred a significant deviation of the Ethiopian *O. longistaminata* populations from the Hardy-Weinberg expectation. Fixation index of this study (F_{is}=0.547) was however lower than the F_{is}=0.75 of 320 *O. longistaminata* accessions (Melaku *et al*., 2013) and F_{is}=0.958 from 900 cultivated, wild and weedy rice individuals (Cao *et al*., 2006).

Pairwise geographical differentiation analysis among the Amhara and Gambella regions showed a higher differentiation (F_{st} = 0.092) than differentiation of the same regions (F_{st} = 0.064) (Melaku *et al*., 2013). Comparison between the two regions for all genetic variability parameters such as Na, Ne and Ho showed the Amhara region to be slightly higher than the Gambella (Table 23). This relatively wider genepool in the Amhara region can be endowed from its altitude (Kiambi *et al*., 2000), environmental heterogeneity (IUCN, 1990; Kiambi *et al*., 2001) and scarce water level which aggravate self-sterility (Kiambi *et al*., 2005).

5.2.3 Population structure, regional differentiation and geographic factors

The two axis of the PCA accounted for 16.22 % of the total genetic variation. To the right side of the Y-axis, a wide-open space existed between all *O. longistaminata* accessions and the rest rice groups (Figure 13). Furthermore, almost all of the Ethiopian *O. longistaminata* groups were clustered on the basis of their geographic locations. Near to the left corner of Quadrant IV, only 3 accessions from Gambella6 showed a relative proximity to the *O. longistaminata* accession from Niger. Thus, this pattern of distribution revealed the genetic distinction of Ethiopian *O. longistaminata* accessions.

The UPGMA clustering analysis merged *O. longistaminata* populations with their AA genome wild and cultivated relatives in one main cluster (Figure 14). To the contrary, hierarchical clustering of the 430 accessions assorted all *O. longistaminata* accessions in one group (Figure 15). Such a slight difference could be associated with feature of generating a more balanced and fewer clusters by the Ward's minimum variance than UPGMA (Odong *et al*., 2011). Here, both the *O. longistaminata* and other main clusters have undergone further division on the basis of their respective geographic location and genome type. However, very few *O. longistaminata* accessions looked admixed. Such intermingled accessions in a cluster could be attributed to the less genetic distance or high genetic identity (Table 24).

Population structure of this study sorted the Amhara and Gambella *O. longistaminata* populations of Ethiopia in two independent clusters (Figure 17). From this structure analysis, it is very clear that the Amhara and Gambella *O. longistaminata* groups of Ethiopia represent two different gene pools. Such distinct population structuring between the two regions can be a function of both geographical and climatic factors. The climatic difference among the two geographic regions was exemplified by the 25.7°C mean temperature and 896 mm annual rainfall for Gambella and 19.2°C mean temperature and 1468 mm annual rainfall for Amhara regions (Melaku, 2011). Difference among populations of the two eco-geographic regions can also be associated with diverse modes of adaptations to variable altitudinal levels, soil types and ecological conditions (Nachimuthu *et al*., 2015).

With respect to the pattern of regional comparison from the SSR markers of this study (Table 23), the genetic distance generated from the Restriction Associated DNA Sequencing (RADSeq) analysis was higher in the assessed accessions of Amhara than within Gambella region (Table 24). This result is expected for populations experiencing significant levels of gene flow (Catchen *et al*., 2013b).

The DAPC and phylogenetic tree generated from the genome wide RADSeq data of this study showed the strong impact of regional difference on the genetic relationships of the Ethiopian *O. longistaminata* accessions. The admixed Gambella accessions in both phylogenetic and DAPC analysis explicitly implicate the strong effect of edaphic, altitudinal and other physical factors. Furthermore, the admixtures can provide information about ancestry, relatedness between populations, history of the entire populations (Cromie *et al*., 2013). The two independent clusters inferred from the BIC versus number of clusters graph of the whole accessions was in harmony with the region based DAPC (Fig. 21).

In the present study, both SSR and RADSeq data were used to make pairwise comparisons of the two ecological regions of Ethiopia. Based on the RADSeq data, genetic differentiation between the Amhara and Gambella regions was estimated at 0.0607 (Table 24). This F_{st} value which is very close to zero couldn’t show a biologically significant differentiation among the regions (Pan *et al*., 2016). In contrast, the resolving power generated from the SSR markers of this study showed a better level of genetic differentiation (F_{st}) between the Amhara and

Gambella regions (Table 23). This variation among the two molecular tools could be associated with the higher variance in the allele frequency of SSR markers (Valdiser *et al*., 2015).

Wright's inbreeding coefficient (F_{is}), which measures reductions in observed heterozygosity with respect to the expectation under Hardy–Weinberg Equilibrium theory is used to detect a hidden population structure (Wright, 1931; Wright, 1978; Slatkin, 1991; Charlesworth, 1998). The positive F_{is} from the whole genome reduction technique of this study indicated the nonrandom mating or cryptic population structure (Nei, 1975; Nei, 1987; Nei and Kumar, 2000; Hartl and Clark, 2006; Holsinger and Weir, 2009). According to Catchen *et al*., (2013b), many genomic regions exhibited higher F_{is} values from the recurrent introgressive hybridization of the differentiated populations.

6. Conclusions and Recommendations

6.1 Conclusions and implications to improve cultivated rice

The *O. longistaminata* accessions in this study showed a highly significant variation when assessed using the 67 SSR markers, eight morphological and seven leaf gas exchange traits. Besides, cluster analysis from those morphological traits explained the high genetic diversity among the accessions. The low E and gs records with a maximum CE of the *O. longistaminata* accessions indicating their environmental adaptability and high Water Use Efficiency. The high Pn observed in the assessed *O. longistaminata* accessions further indicated the potential of *O. longistaminata* as a genetic resource for photosynthetic improvement of rice cultivars.

In this study, the microsatellite markers were used to show the differentiation level of *O. longistaminata* from other wild and cultivated AA genome rice species. Genotypic grouping through structure analysis, distance-based clustering and principal component analysis similarly confirmed distinctness of *O. longistaminata*. On the basis of their geographic regions, the assessed *O. longistaminata* populations of Ethiopia were genetically different and further subgrouping was in concurrence with geographical proximity. Population structuring by the SSR markers, cluster analysis from the eight quantitative traits and phylogeny from the RADSeq analysis indicated two gene pools comprising the Amhara populations in one group and the Gambella populations in the other.

Generally, the high level of genetic variation from the natural *O. longistaminata* populations along with the diversified morphological features and promising photosynthetic potential showed an ideal resource for rice improvement through modern breeding tools.

6.2 Recommendations and future research directions

In this study, morphological and molecular characterization of the Ethiopian *O. longistaminata* accessions showed the existence of high genetic diversity. However, further investigation on their eco-geographic pattern of distribution could be used as a benchmark for the selection of *in situ* conservation sites.

The present study also revealed a considerable significance of the Ethiopian *O. longistaminata* accessions as sources of genotypes with desirable attributes for high photosynthetic rate, less transpiration rate and great carboxylation efficiency. Such better performances should be supplemented with further efforts targeting genetic enhancement or favoring frequency of desirable genes of the accessions via selection.

Despite the important roles they could have played in rice breeding programs, potential of the Ethiopian *O. longistaminata* accessions has been under exploited by the worldwide breeding efforts. Hence, the accessions used in this study can serve as indicators for the utilization of *O. longistaminata* for the purpose of breeding.

References

Aggarwal, R.K., Brar, D.S., Nandi, S., Huang, N. and Khush, G.S. (1999). Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theor. Appl. Genet.* **98**:1320–1328

Akimoto, M., Shimamoto, Y. and Morishima, H. (1997). Genetic differentiation in *O. glumaepatula* and its phylogenetic relationships with other AA genome species. *Rice Genet. Newsl.* **14**: 37–39.

Ali, M.L., Sanchez, P.L., Yu, S., Lorieux, M. and Eizenga, G.C. (2010). Chromosome segment substitution lines: a powerful tool for the introgression of valuable genes from *Oryza* wild species into cultivated rice (*O . sativa* L.). *Rice* **3**: 218–234.

Aliyu, R., Adamu, A.K., Gregorio, G.B. (2013). Allelic data revealing interrelatedness in rice species (*Oryza sativa*, *Oryza glaberrima*, *Oryza barthii*) and the interspecific hybrids (NERICA). *J. of. P. Breed & Crop Sci*. **5**: 12-19

Amsalu, A. and Endashaw, B. (1999). Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution* **46**: 273–284.

Andrews, R.K., Good, M.J., Miller, R.M., Luikart, G. and Hohenlohe, A.P. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet*. **17**: 81–92.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* **3**: 1-7.

Barbour, M.M., Warren, C.R., Farquhar, G.D., Forrester, G.Y. and Brown, H. (2010). Variability in mesophyll conductance between barley genotypes, and effects on transpiration efficiency and carbon isotope discrimination. *Plant Cell Environ* **33**: 1176–1185.

Bezancon, G., Bozza, J.L. and Second G. (1977). Viability of *Oryza longistaminata* and the *Sativa* complex of *Oryza* in Africa. Ecological and evolutive aspects. **In**: IRAT-ORSTOM meeting on African rice species. Instituto de Recherche Agronomiques Tropicales-Office de la Recherche Scientifique et Technique d'Outre-Mer, Paris, pp 47-55.

Bezanc¸on, G. (2008). Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theor*. *Appl*. *Genet*. **116**: 903–913.

Boissot, N., Valdez, M. and Guiderdoni, E. (1990). Plant regeneration from leaf and seed-derived calli and suspension cultures of the African perennial wild rice, *Oryza longistaminata*. *Plant Cell Reports* **9**: 447-450.

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 gene introgression in rice. *P. Mol. Biol*. **35**: 35–47.

Brar, D.S. and Khush, G.S. (2003). Utilization of wild species of genus Oryza in rice improvement. **In**: Monograph on genus Oryza, Nanda JS, Sharma SD, editors. Enfield: Science Publ; p. 283–309.

Brar, D.S. and Khush, G.S. (1986). Wide hybridization and chromosome manipulation in cereals. **In**: Handbook of Plant Cell Culture, Vol. 4: Techniques and Applications, Evans, D.H., Sharp, W.R. and Ammirato, P.V. eds, MacMillan Publ Co., New York, USA pp 221–263.

Brar, D.S. and Khush, G.S. (2002). Transferring genes from wild species into rice. **In**: Quanttitative Genetics, Genomics and Plant Breeding. Kang MS, editor, CAB International, Wallingford, UK, pp 197–217.

Brar, D.S. and Singh, K. (2011). Oryza. **In**: C Kole, ed, Wild Crop Relatives: Genomics and Breeding Resources, Cereals. Springer-Verlag, Berlin, pp 321–365.

Brar, D.S., Elloran, R. and Khush, G.S. (1991). Interspecific hybrids produced through embryo rescue between cultivated and eight wild species of rice. *Rice Genet. News* **8**: 91–93.

Brondani, R.V., Zucchi, M.I., Brondani, C., Rangel, P.N., Borba, T.O., Rangel, P.N., Magalhães, M.R. and Vencovsky, R. (2005). Genetic structure of wild rice *Oryza glumaepatula* populations in three Brazilian biomes using microsatellite markers. *Genetica* **125**: 115-123.

Brown, J.S. (1991). Principal component and cluster analyses of cotton cultivar variability across the U.S. Cotton Belt. *Crop Sci*. **31**: 915–922.

Buddenhagen, I. (1982). Bacterial blight of rice appears to be indigenous on wild rice species in West Africa. *International Rice Research Newsletter* **7**: pp 5.

Camussi, A., Ottaviano, E., Calinski, T. and Kaczmarek, Z. (1985). Genetic distances based on quantitative traits. *Genetics* **111**: 945–962.

Cao, Q., Lu, B., Xia, H., Rong, J., Sala, F., Spada, A. and Grassi, F. (2006). Genetic Diversity and Origin of Weedy Rice (*Oryza sativa* f. spontanea) Populations Found in North-eastern China Revealed by Simple Sequence Repeat (SSR) Markers. *Ann. of Bot*. **98**: 1241–1252.

Cao, S.Q., Zhai, H.Q., Yang, T.N., Zhang, R.X. and Kuang, T.Y. (2001). Studies on photosynthetic rate and function duration of rice germplasm. *Chin. J. Rice. Sci.* **15**: 29–34.

Cappers, R.J. (2007). Fields of Change. **In**: Progress in African Archaeobotany. Barkhuis, Groningen, Netherlands.

Catchen, J., Bassham, S., Wilson, T., Currey, M., Brien, O.C., Yeates Q. and Cresko, A.W. (2013a). The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Mol. Ecol*. **22**: 2864–2883.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A. and Cresko, W. (2013b). Stacks: an analysis tool set for population genomics. *Mol. Ecol*. **22**: 3124–3140.

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J.H. (2011). *Stacks*: Building and Genotyping Loci *De Novo* From Short-Read Sequences. *G3: Genes|Genomes|Genetics* **1**: 171-182.

Chang, T. T. (1985). Crop history and genetic conservation: Rice, a case study. *Iowa State J. Res.* **59**: 425–455.

Chang, T.T. (1976). The origin, evolution, cultivation, dissemination, and diversification of Asian and African rice. *Euphytica* **25**: 425–441.

Chang, T.T. and Bardenas, E.A. (1965). The morphology and varietal characteristics of the rice plant. *IRRI Tech. Bull*. **4**, pp 40.

Chang, T.T. and Vaughan D.A. (1989). Conservation and potentials of rice genetic resources. **In**: Biotechnology in Agriculture and Forestry. Y.P.S. Bajaj. ed. Springer Verlag (in press), Berlin.

Changhai, S., Baodi, D., Yunzhou, Q., Yuxin, L., Lei, S., Mengyu, L. and Haipei, L. (2010). Physiological regulation of high transpiration efficiency in winter wheat under drought conditions. *P. Soil Env.* **56**: 340–347.

Charlesworth, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Mol. Biol. Evol.* **15**: 538-43

Chen, X., Temnykh, S., Xu, Y., Cho, G.Y. and McCouch, R.S. (1997). Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa L.*). *Theor. Appl. Genet.* **95**: 553-567.

Chen, Z., Hu, F., Xu, P., Li, J., Deng, X., Zhou, J., Li, F., Chen, S. and Tao, D. (2009). QTL analysis for hybrid sterility and plant height in interspecific populations derived from a wild rice relative, *Oryza longistaminata*. *Breeding Science* **59**: 441–445.

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

Chevalier, A. (1932). Nouvelle contribution a'ltude systematique des *Oryza*. *Rev. Bot. Appl. Agric. Trop.* **12**: 1014-1032.

Chouhan, S.K., Singh, A.K., Singh, A., Singh, S.P., Singh, N.K. and Singh, P.K. (2014). Agro-morphological diversity in wild rice accessions of eastern Indo-Gangetic region of INDIA. *Bangladesh J. Bot.* **43**: 337-344.

Cissé, F. and Khouma, P.M. (2016). Improvement of wild rice *Oryza Longistaminata* through mutation induction. *J. of Food Sci. and Eng.* **6**: 82-89.

Clark, V.L., Brummer, E.J., Głowacka, K., Hall, M., Heo, K., Peng, J., Yamada, T., Yoo, H.J., Yu, Y.C., Zhao, H., Long, P.S. and Sacks, J.E. (2014). A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*. *Ann. Bot.* **114**: 97-107.

Claudia, K. and Knauer, K. (2013). Remote sensing of rice crop areas. *Internal Journal of Remote Sensing* **34**: 2101-2139.

Clayton, W.D. (1968). Studies in Gramineae: XVII, West African wild rice. *Kew Bull*. **21**: 487-488.

Cook, M.G., Evans, L.T. (1983). Some physiological aspects of the domestication and improvement of rice (Oryza spp.). *Field Crops Research* **6**: 205–218.

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

Cowen, N.M. and Frey, K.J. (1987). Relationships between three measures of genetic distance and breeding methods in oat (*Avena sativa* L.). *Genome* **29**: 97–106.

Cromie, G.A., Hyma, K.E., Ludlow, C.L., Garmendia-Torres, C., Gilbert, T.L., May, P., Huang, A.A., Dudley, A.M. and Fay, J.C. (2013). Genomic sequence diversity and population structure of Saccharomyces cerevisiae assessed by RAD-seq. *G3: Genes|Genomes|Genetics* **3**: 2163-2171.

CSA, 2007-2016. Central Statistical Agency of Ethiopia.

Dally, A.M. and Second, G. (1990). Chloroplast DNA diversity in wild and cultivated species in rice (genus *Oryza* section Oryza). Cladistic-mutation and genetic distances analysis. *Theor. Appl. Genet*. **80**: 209-222.

Dalmacio, R.D., Brar, D.S., Ishii, T., Sitch, T.A., Virmani, S.S. and Khush, G.S. (1995). Identification and transfer of a new cytoplasmic male sterility source from *Oryza perennis* into indica rice (*Oryza sativa*). *Euphytica* **82**: 221–225.

Das, B., Sengupta, S., Parida, S.K., Roy, B., Ghosh, M., Prasad, M. and Ghose, T.K. (2013). Genetic diversity and population structure of rice landraces from Eastern and North Eastern States of India. *BMC Genet*. 14:1-14.

Deshpande, V.G. and Ranjekar, P.K. (1980). Repetitive DNA in three Gramineae species with low DNA content. Hopper Seyler's Z. *Physiol*. *Chem*. **361**: 1223-1233.

Deu, M., Sagnard, F., Chantereau, J., Calatayud, C., He´rault, D., Mariac, C., Pham, J.L., Vigouroux, Y., Kapran, I., Traore´, P.S., Mamadou, A., Ge´rard, B., Ndjeunga, J.

and Bezancon, G. (2008). Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theor. Appl. Genet.* **116**: 903–913.

Dinesh, R.R., Kiran, A.G. and Thomas, G. (2010). Population genetic structure and conservation priorities of *Oryza rufipogon* Griff. populations in Kerala, India. *Curr. Sci.* **98**: 65–68.

Ding, Z.S., Li, T., Zhu, X.G., Sun, X.F., Huang, S.H., Zhou, B.Y. and Zhao M. (2014). Three photosynthetic patterns characterized by cluster analysis of gas exchange data in two rice populations. *Crop J.* **2**: 22–27.

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

Doyle, J.J., Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue, *Phytochem. Bull.* **19**: 11-15.

Doyle, J.J., Morgante, M., Tingey, S.V. and Powell, W. (1998). Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine*). *Mol. Biol. Evol.* **15**: 215–218.

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S. and Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. *Biochem. Genet.* **45**: 113-129.

Duistermaat, H. (1987). A revision of *Oryza* (Gramineae) in Malaysia and Australia. *Blumea* **32**: 157-193.

Duvall, M.R., Peterson, P.M., Terrell, E.E. and Christensen, A.H. (1993). Phylogeny of North American oryzoid grasses as construed from maps of plastid DNA restriction sites. *Am. J. Bot.* **80**: 83–88.

Earl, D.A. and von Holdt, B.M. (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. doi:10.100712686_011_9548_7.

Emerson, J.K., Merz, R.C., Catchen, M.J., Hohenlohe, A.P., Cresko, A.E., Bradshaw, E.W. and Holzapfe, M.C. (2010). Resolving postglacial phylogeography using high-throughput sequencing. *PNAS* **107**: 16196-16200.

Ethiopian Ministery of Agriculture, MoA. (2017). National rice research and development strategies of Ethiopia.

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

Evans, J.R. (1999). Leaf anatomy enables more equal access to light and CO2 between chloroplasts. *New Phytol* **143**: 93–104.

Evans, J.R. and Loreto, F. (2000). Acquisition and diffusion of CO2 in higher plant leaves. **In**: Photosynthesis: Physiology and Metabolism. RC Leegood, TD Sharkey, S von Caemmerer, eds, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 321–351

Evans, J.R. and von Caemmerer, S. (1996). Carbon dioxide diffusion inside leaves. *Plant Physiol* **110**: 339–346.

Food and Agricultural Organization, FAO (2010). FAOSTAT. Food and Agriculture Organization of the United Nations.

Food and Agricultural Organization, FAO (2011). FAOSTAT. Food and Agriculture Organization of the United Nations.

Farquhar, G.D. and Sharkey, T.D. (1982). Stomatal conductance and photosynthesis. *Ann. Rev. P. Physiol.* **33**: 317–345.

Feild, T.S., Arens, N.C. and Dawson, T.E. (2003). The ancestral ecology of angiosperms: emerging perspectives from extant basal lineages. *Int. J. Plant Sci.* **164**: 129–142.

Flexas, J., Niinemets, U., Gallé, A., Barbour, M.M., Centritto, M., Diaz-Espejo, A., Douthe, C., Galmés, J., Ribas-Carbo, M. and Rodriguez, P.L. *et al* (2013). Diffusional conductances to CO2 as a target for increasing photosynthesis and photosynthetic water-use efficiency. *Photosynth. Res.* **117**: 45-59.

Flexas, J., Ribas-Carbó, M., Diaz-Espejo, A., Galmés, J. and Medrano, H. (2008). Mesophyll conductance to CO2: current knowledge and future prospects. *P. Cell Environ.* **31**: 602–621.

Fu, X.L., Lu, Y.G., Liu, X.D. and Li, J.Q. (2008). Progress on transfering elite genes from non-AA genome wild rice into *Oryza sativa* through interspecific hybridization. *Rice Sci.* **15**:79–87.

Gao, L.Z. (2004). Population structure and conservation genetics of wild rice *Oryza rufipogon* (Poaceae): A region-wide perspective from microsatellite variation. *Mol. Ecol.* **13**: 1009–1024.

Gaut, B.S. (2002). Evolutionary dynamics of grass genomes. *New Phytol.* **154**: 15-28.

Ge, S., Guo, Y. and Zhu, Q. (2004). Molecular phylogeny and divergence of the rice tribe Oryzeae, with special reference to the origin of the genus *Oryza* Session **In**: The genus *Oryza*, its diversity, and its evolution pp 40-44.

Ge, S., Li, A., Lu, B.R., Zhang, S.Z. and Hong, D.Y. (2002). A phylogeny of the rice tribe Oryzeae (Poaceae) based on *mat*K sequence data. *Am. J. Bot.* **89**: 1967–1972.

Ge, S., Sang, T., Lu, B.R. and Hong, D.Y. (1999). Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proc. Natl. Acad. Sci. USA* **96**: 14,400-14,405.

Ghesquiere, A. (1987). Evolution of *Oryza longistaminata*. ORSTOM, Paris

Ghesquiere, A. (1985). Evolution of *Oryza longistaminata*. **In**: Rice Genetics. International Rice Research Institute, Malina, Philippines, pp 15-25.

Ghneim, H. T., Poso, D., Perez, I., Torrealba, G. N., Pieters, A., Martinez, C. and Tohme, J. (2008). Assessment of genetic diversity in Venezuelan rice cultivars using simple sequence repeats marker. *Electronic J. of Biotech*. **11**: 1-14.

Giuliani, R., Koteyeva, N., Voznesenskaya, E., Evans, M.A., Cousins, A.B. and Edwards, G.E. (2013). Coordination of leaf photosynthesis, transpiration and structural traits in rice and wild relatives (genus *Oryza*). *Plant Physiol.* **162**: 1632–1651.

Glaszmann, J.C. (1987). Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet*. **74**: 21–30.

Goodman, M.M. (1972). Distance analysis in Biology. *Syst. Zool*. **21**: 174-186.

GRiSP (Global Rice Science Partnership). 2013. Rice almanac, 4th edition. **In**: International Rice Research Institute, Los Baños, Philippines p. 283.

Guimara˜es, E.P. (2009). Food and Agriculture Organization of the United Nations (FAO) **In**: Cereals, The Banks and the Italian Economy, Viale delle Termi di Caracalla, Crop and Grassland Service (AGPC) – Room C-778, 00153 Rome, Italy, M.J. Carena (ed.) pp 99-126.

Guo, Y.L. and Ge, S. (2005). Molecular phylogeny of *Oryzeae* (*Poaceae*) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. *Am. J. Bot*. **92**: 1548–1558.

Gutierrez-Rodriguez, M., Reynolds, M.P. and Larque-Saavedra, A. (2000). Photosynthesis of wheat in a warm, irrigated environment: II. Traits associated with genetic gains in yield. *Field Crops Res*. **66** 51–62.

Haritha, G., Vishnukiran, T., Yugandhar, P., Sarla, N. and Subrahmanyam, D. (2017). Introgressions from *Oryza rufipogon* Increase Photosynthetic Efficiency of KMR3 Rice Lines. *Rice Science* **24**: 85−96.

Harlan, J. R., (1969). Ethiopia: A centre of diversity. *Econ. Bot*. **23**: 309–314.

Hartl, D.L. and Clark, A.G. (2006). Principles of Population Genetics, 4th edn. Sinauer Associates, Sunderland, Massachusetts

Harper, J.L. (1977). Population Biology of Plants. Academic Press, London.

Henry, R.J., Rice, N., Waters, D.E., Kasem, S., Ishikawa, R., Hao, Y., Dillon, S., Crayn, D., Wing, R. and Vaughan, D. (2010). Oryza: Utility and conservation. *Rice* **3**: 235–241.

Hetherington, A.M. and Woodward, F.I. (2003). The role of stomata in sensing and driving environmental change. *Nature*. **424**: 901–908.

Hirasawa, T., Iida, Y. and Ishihara, K. (1988). Effect of leaf water potential and air humidity on photosynthetic rate and diffusive conductance in rice plants. *Jpn. J. Crop Sci*. **57**: 112–118.

Hirasawa, T., Ozawa, S., Taylaran, R.D. and Ookawa, T. (2010). Varietal differences in photosynthetic rates in rice plants with special reference to the nitrogen content of leaves. *Plant Prod. Sci*. **13**: 53–57.

Holsinger, K.E. and Weir, B.S. (2009). Genetics in geographically structure populations: defining, estimating and interpreting FST. *Nat. Reviews Genetics* 10: 639–650

Horie, T., Lubis, I., Takai, T., Ohsumi, A., Kuwasaki, K., Katsura, K. and Nii, A. (2003). Physiological traits associated with high yield potential in rice. **In**: Rice Science: innovations and impact for livelihood, Mew TW, Brar DS, Peng S, Dawe D, Hardy B, eds. IRRI, Los Ban˜os, Philippines, pp 117–145.

Hovius, N., Stark, C.P., Tutton, M.A. and Abbott, L.D. (1998). Landslide-driven drainage network evolution in a pre-steady-state mountain belt; Finisterre Mountains, Papua New Guinea. *Geology* **26**: 1071–1074.

Hu, C.H. (1970). Cytogenetic studies of *Oryza officinalis* complex. III. The genomic constitution of *O. punctata* and *O. eichengeri*. *Cytologia* **35**: 304-318.

Hu, F.Y., Tao, D.Y., Sacks, E., Fu, B.Y., Xu, P., Li, J., Yang, Y., McNally, K., Khush, G.S., Paterson, A.H. and Li, Z.K. (2003). Convergent evolution of perenniality in rice and sorghum. *Proc. Natl. Acad. Sci. USA* **100**: 4050–4054.

Huang, H. and Kochert, G. (1994). Comparative RFLP mapping of an allotetraploid wild rice species (*Oryza latifolia*) and cultivated rice (*O. sativa*). *Plant Mol. Biol*. **25**: 633–648.

Hubbart, S., Peng, S., Horton, P., Chen, Y. and Murchie, E.H. (2007). Trends in leaf photosynthesis in historical rice varieties developed in the Philippines since 1966. *J. Exp. Bot*. **58**: 3429–3438.

IBM Corp (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

International Rice Research Institute, IRRI. (1964). Rice genetics and cytogenetics. Elsevier Publishing Company, Amsterdam, pp 274.

IRRI-IBPGR, 1980. Descriptors for rice (Oryza sativa L.). International Rice Research Institute, Manila, Philippines.

Ishii, T. and McCouch, S.R. (2000). Microsatellites and microsynteny in the chloroplast genomes of Oryza and eight other Gramineae species. *Theor. Appl. Genet*. **100**: 1,257-1,266.

Ishii, T., Nakano, T., Maeda, H. and Kamijima, O. (1996). Phylogenetic relationships in A-genome species of rice as revealed by RAPD analysis. *Genes Genet. Syst*. **71**: 195–201.

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

Jagadish, K.V., Cairns, J.E., Kumar, A., Somayanda, I.M. and Craufurd, P.Q. (2011). Does susceptibility to heat stress confound screening for drought tolerance in rice? *Funct Plant Biol*. **38**: 261–269.

Jena, K.K. (2010). The species of the genus *Oryza* and transfer of useful genes from wild species into cultivated rice*, O. sativa*. *Breeding Sci*. **60**: 518–523.

Jena, K.K. and Khush, G.S. (1984). Embryo rescue of interspecific hybrids and its scope in rice improvement. *Rice Genet. Newsl*. **1**: 133–134.

Jena, K.K. and Khush, G.S. (1990). Introgression of genes from *Oryza officinalis* Well exWatt to cultivated rice, *O. sativa* L. *Theor. Appl. Genet*. **80**: 737-745.

Jena, K.K. and Khush, G.S. (2000). Exploitation of species in rice improvement: opportunities, achievements and future challenges. **In**: Rice breeding and genetic: research priorities and challenges. Nanda JS, editor, Science Publ, Enfield, pp. 269–284.

Jena, K.K. and Kochert, G. (1991). Restriction fragment length polymorphism analysis of CCDD genome species of the genus *Oryza L*.. *P. Mol. Biol*. **16**: 831-839.

Jena, K.K., Khush, G.S. and Kochert, G. (1994). Comparative RFLP mapping of a wild rice, *Oryza officinalis*, and cultivated rice, *O. sativa*. *Genome* **37**: 382–389.

Jiang, H. and Xu, D.Q. (2001). The cause of the difference in leaf net photosynthetic rate between two soybean cultivars. *Photosynthetica* **39**: 453–459.

John, V.T., Thottapilly, G. and Awoderu, V.A. (1984). Occurrence of rice yellow mottle virus in some Sahelian countries in West Africa. *FAO Plant Protection Bulletin* **32**: 86-87.

Jombart, T., Devillard, S. and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94.

Jones, M.P., Dingkuhn, M., Johnson, D.E. and Fagade, S.O. (1996). Inter-specific hybridization: progress and prospects. Proceedings of workshop on: Africa/Asia joint research on inter-specific hybridization between the African and Asian rice species (*O. glaberrima* and *O. sativa*). West Africa Rice Development Association, Bouake, Cote D’Ivoire

Joshi, S.P., Gupta, V.S., Aggarwal, R.K., Ranjekar, P.K. and Brar, D.S. (2000). Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theor. Appl. Genet*. **100**: 1311–1320.

Judith, J. (2015). Morphological, Molecular and Quality Characterization of rice varieties and landraces from Eastern and Southern Africa. MSc. Sokoine University of Agriculture, Morogoro, Tanzania.

Kajala, K., Coushaff, S., Karki, S., Woodfield, H. and Tolley, B.J. (2011). Strategies for engineering a two celled C4 photosynthetic pathway into rice. *J. Exp. Bot.* **62**: 3001–3010.

Kanazawa, A., Akimoto, M., Morishima, H. and Shimamoto, Y. (2000). The distribution of Stowaway transposable elements in AA genome species of wild rice. *Theor. Appl. Genet*. **101**: 327-335.

Kanno, A. and Hirai, A. (1992). Comparative studies of the structure of chloroplast DNA from four species of Oryza: cloning and physical maps. *Theor. Appl. Genet*. **83**: 791-798.

Kanwal, K.S., Singh, R.M., Singh, J. and Singh, R.B. (1983). Divergent gene pools in rice improvement. *Theor. Appl. Genet*. **65**: 263– 267.

Katayama, T. and Onizuka, W. (1979). Intersectional F_1 plants from *Oryza sativa* X *O. ridleyi* and *O. sativa* X *O. meyeriana*. *Jpn. J. Genet*. **54**: 43-46.

Katayama, T., Onizuka, W. and Shin Y.B. (1981). Intersectional F1 hybrids obtained from the crosses *Oryza minuta* Presl. X *O. ridleyi* Hook. and *O. officinalis* Wall. X *O. ridleyi* Hook. *Jpn. J. Genet*. **56**: 67-71.

Katayama, T.C. (1961). Report on the collection trip to the Philippines and New Guinea. (mimeo.)

Katayama, T.C. (1963). Report on the collection trip to Borneo and Java. (mimeo.)

Kawamitsu. Y. and Agata, W. (1987). Varietal differences in photosynthetic rate, transpiration rate and leaf conductance for leaves of rice plants. *Japanese J. of Crop Sci.* **56**: 563–570.

Kellogg, E.A. (2001). Evolutionary history of grasses. *Pla. Physiol.* **125**: 1198-1205.

Khush, G.S. (1997). Origin, dispersal, cultivation and variation of rice. *P. 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., 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. Dissertation, University of Birmingham.

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. *Afri*. *J*. *of Biotech*. **4**: 308-317.

Kihara, H. (1963). Genome-analysis in *Oryza*. **In**: Symposium on rice problems. Tenth Pacific Science Congress of the Pacific Science Association, University of Hawaii, Honolulu, pp 57-66.

Kiran, T.V., Rao, Y.V., Subrahmanyam, D., Rani, N.S., Bhadana, V.P., Rao, P.R. and Voleti, S.R. (2013). Variation in leaf photosynthetic characterstics in wild rice species. *Photosynthetica* **51**: 350–358.

Kondamudi, R., Swamy, K.N., Rao, Y.V., Kiran, V.T., Suman, K., Rao, S.D., Rao, R.V., Subrahmanyam, D., Sarla, N., Kumari, R.B. and Voleti, R.S. (2016).Gas exchange, carbon balance and stomatal traits in wild and cultivated rice (Oryza sativa L.) genotypes. *Acta Physiol Plant* **38**: 1-9.

Krishnaswamy, N. and Chandrasekharan, P. (1957). Note on naturally occurring tetraploid species of *Oryza*. *Sci. Cult*. **23**: 308-310.

Kusumi, K., Hirotsuka, S., Kumamaru, T. and Iba, K. (2012). Increased leaf photosynthesis caused by elevated stomatal conductance in a rice mutant deficient in SLAC1, a guard cell anion channel protein. *J. Exp. Bot.* **63**: 5635–5644.

Langmead, B. (2010). Aligning short sequencing reads with Bowtie Curr Protoc Bioinformatics. doi: 10.1002/0471250953.bi1107s32.

Lasalita-Zapico, F.C., Namocatcat, J.A. and Carino-Turner, J.L. (2010). Genetic diversity analysis of traditional upland rice cultivars in Kihan, Malapatan, Sarangani Province, Philippines using morphometric markers. *Philip. J. Sci*. **139**: 177-180.

Launert, E. (1965). A survey of the genus *Leersia* in Africa. *Senckenb. Biol*. **46**: 129-153.

Li, C.B., Zhang, D.M., Ge, S., Lu, B.R. and Hong, D.Y. (2001). Differentiation and intergenomic relationships among C, E and D genomes in the Oryza officinaliscomplex (Poaceae) as revealed by multicolor genomic in situ hybridization. *Theor. Appl. Genet*. **103**: 197-203.

Li, H.W., Weng, T.S., Chen, C.C. and Wang, W.H. (1961). Cytogenetical studies of *Oryza sativa L*. and its related species. 1. Hybrids *O. paraguaiensis* wedd. X *O. brachyantha* A. Chev. et Roehr., *O. paraguaiensis* Wedd. X *O. australiensis* Domn and *O. alta* Swallen. *Bot. Bull. Acad. Sin* **2**: 79-86.

Li, Y., Gao, Y.X., Xu, X.M., Shen, Q.R. and Guo, S.W. (2009). Light-saturated photosynthetic rate in high-nitrogen rice (*Oryza sativa* L.) leaves is related to chloroplastic CO2 concentration. *J. Exp. Bot.* **60**: 2351–2360.

Lin, S.C. and Yuan, L.P. (1980). A mass screening method for testing grassy stunt disease of rice. In: Innovative approaches to rice improvement, hybrid rice breeding in China, International Rice Research Institute, Manila, Philippines pp 35–51.

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

Londo, J.P., Chiang, Y.C., Hung, K.H., Chiang, T.Y. and Schaal, B. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci.* **103**: 9578–9583.

Lu, B. R., Naredo, M. E., Juliano, A. B. and Jackson, M. T. (2000). Preliminary studies on taxonomy and biosystematics of the AA-genome *Oryza* species (Poaceae). **In**: Grasses: Systematics and Evolution. Jacobs, S.W. L., and Enerett, J. (eds.), CSIRO, Melbourne, Australia, pp. 51–58.

Lu, B.R. (1999). First record of the wild rice Oryza meridionalis in Indonesia. *IRRI Newslt.* **24**: 28.

Lu, B.R., Jackson, M. and Vaughn, D. (2010). Wild rice taxonomy. **In**: Rice knowledge bank, Int Rice Res Inst, Philippines.

Lu, B.R., Zheng, K.L., Qian, H.R. and Zhuang, J.Y. (2002). Genetic differentiation of wild relatives of rice as referred by the RFLP analysis. *Theor. Appl. Genet.* **106**: 101–106.

Lu, F., Ammiraju, J.S., Sanyal, A., Zhang, S., Song, R., Chen, J., Li, G., Sui, Y., Song, X., Cheng, Z., de Olivera, A.C., Bennetzen, J.L., Jackson, S.A., Wing, R.A. and Chen, M. (2009). Comparative sequence analysis of monoculm1-orthologous regions in 14 Oryza genomes. *Proc. Natl. Acad. Sci.* **106**: 2071–2076.

Lu, F., Lipka, A.E., Glaubitz, J., Elshire, R., *et al*. (2013). Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genetics* 9: e1003215

Lu, J.J. and Chang, T.T. (1980). Rice in its temporal and spatial perspective. **In:** Rice production and utilization. Edited by Luh, B.S. Westport, CT, USA, AVI Publishing Co. Pp. 1-74.

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

Mahalanobis, P.C. (1936). On the generalised distance in statistics. Proceedings of the National Institute of Sciences of India. **2**: 49–55.

Maheswaran, M. (2004). Molecular markers: History, features and application. *Advance Biotech*, Pp 17-24.

Makino, A., Mae, T. and Ohira, K. (1985). Photosynthesis and ribulose-1,5- bisphosphate carboxylase/oxygenase in rice leaves from emergence through senescence. Quantative analysis by carboxylation/oxygenation and regeneration of ribulose 1,5-bisphosphate. *Planta* **166**: 414–420.

Marathi, B., Ramos, J., Hechanova, S.L., Oane, R.H. and Jena, K.K. (2015). SNP genotyping and characterization of pistil traits revealing a distinct phylogenetic relationship among the species of Oryza. *Euphytica* **201**: 131–148.

Marlee, R.L. (2016). Population structuring of *O. longistaminata*. MSc thesis, University of Illinois, USA.

Maruyama, S. and Tajima, K. (1990). Leaf conductance in japonica and indica rice varieties: I. Size, frequency, and aperture of stomata. *Jpn. J. Crop Sci.* **59**: 801–808.

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

McCouch S.R., Chen, X., Panaud, O., Temnykh, S., Xu, Y., Cho, Y.G., Huang, N., Ishii, T. and Blair, M. (1997). Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol Biol.* **35**: 89-99.

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**: 199-207.

Melaku G. (2011). Genetic diversity of an African wild rice (*Oryza longistaminata* Chev. et Roehr) from Amhara and Gambella regions as revealed by microsatellite markers. MSc. Thesis, Addis Ababa University, Ethiopia

Melaku, G., Teklehaimanot, H., Tileye, F. and Samuel, K. (2013). Genetic diversity of wild rice *Oryza longistaminata* Chev. Et. Roher in Ethiopia as revealed by microsatellite markers. *Genet Resour*. *and Crop Evol*. **60**: 1047-1056.

Mengesha, M. H., (1975). Crop germplasm diversity resources of Ethiopia. **In**: Frankel, O.H. & J.G. Hawkes (Eds.), Crop Genetic Resources for Today and Tomorrow, pp. 449–453. Cambridge University Press, Cambridge.

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A. and Johnson, E.A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* **17**: 240-248.

Minitab 17 Statistical Software, [Computer software]. (2010). State College, PA: Minitab, Inc. www.minitab.com

Molosiwa, O.O. (2012). Genetic diversity and population structure analysis of bambara groundnuts (Vigna subterranea (L.) Verdc.) landraces using morphoagronomic characters and SSR markers. PhD thesis, University of Nottingham. Leicestershire, UK.

Morishima, H. (2000). Evolution and domestication of rice. **In**: Rice Genetics IV. Proc. of 4th Intern. Rice Genet. Symp. Science Publishers, India and IRRI, Philippines.

Morishima, H., and Oka, H. I. (1960). The pattern of interspecofic variation in the genus *Oryza*: Its quantitative representation by statistical methods. *Evolution* **14**:153–165.

Morishimia, H., Oka, H.I. and Chang, W.T. (1961). Directions of differentiation on populations of wild rice, *Oryza perennis* and *O. sativa f. spontanea*. *Evol*. **15**: 326-339.

Nachimuthu, V.V., Muthurajan, R., Duraialaguraja, S., Sivakami, R., Pandian, A.B., Ponniah, G., Gunasekaran, K., Swaminathan, M., Suji, K.K. and Sabariappan, R. (2015). Analysis of Population Structure and Genetic Diversity in Rice Germplasm Using SSR Markers: An Initiative Towards Association Mapping of Agronomic Traits in Oryza Sativa. *Rice* **8**: 1-24

Nayar N.M. (2014a). History and Early Spread of Rice **In:** Origins and Phylogeny of Rices. 1st Edition, Kindle Edition Pp 15-36.

Nayar, N.M. (2014b). Rice in the World **In:** Origins and Phylogeny of Rices. 1st Edition, Kindle Edition Pp 1-14.

Nayar, N.M. (1973). Origin and cytogenetics of rice. *Adv. Genet.* **17**: 153–292.

Nayar, N.M. (2010). Origin of African rice from Asian rice. **In**: Innovation and Partnerships to Realize Africa's Rice Potential, Second Africa Rice Congress, Bamako, Mali, Pp 1.18.1-1.18.13.

Nei. M. (1975). Molecular Population Genetics and Evolution. Elsevier, New York

Nei, M. (1987). Molecular Evolutionary Genetics. Columbia University Press, New York

Nei, M. and Kumar, S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York

Ng, N.Q., Chang, T.T., Williams, J.T. and Hawkes, J.G. (1981). Morphological studies of Asian rice and its related wild species and the recognition of a new Australian taxon. *Biol. Jour. Linnean. Society* **16**: 303–313.

Ngu, M.S., Sabu, K.K., Lim, L.S., Abdullah, M.Z. and Wickneswari, R. (2010). Genetic structure of *Oryza rufipogon* Griff. natural populations in Malaysia: implications for conservation and genetic introgression of cultivated rice. *Tropical P. Biol.* **3**: 227–239.

Niu, S.L., Li, L.H., Jiang, G.M., Gao, L.M., Li, Y.G., Peng, Y. and Liu, M.Z. (2004). Gas exchange and chlorophyll fluorescence response to simulated rainfall in *Hedysarum fruticosum* var. *mongolicum*. *Photosynthetica* **42**: 1–6.

Nobel, P.S. (2009). Physicochemical and Environmental Plant Physiology, Ed 4. Academic Press, San Diego

Normile, D. (2008). Reinventing rice to feed the world. *Sci*. **321**: 330–333.

Nyoka, G.C. (1983). Weed problems and control practices in deepwater and floating rice in Mali. In: Deat M, Marnotte P, (Eds.) Proceedings of the Second Biannual Conference. Abidjan, Côte d'Ivoire: West African Weed Science Society, pp 146-157.

Odong, T.L., Van Heerwaarden, J., Jansen, J., Van Hintum, T.L. and Van Eeuwijk, F.A. (2011). Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data? *Theor. Appl. Genet*. **123**: 195–205.

Ohno, Y. (1976). Varietal differences of photosynthetic efficiency and dry matter production in indica rice. *Trop Agric.* **53**: 115–123.

Ohsumi, T., Kanemura, K., Homma, T., Horie, T. and Shiraiwa, T. (2007a). Genotypic variation of stomatal conductance in relation to stomatal density and length in rice (Oryza sativa L.). *Plant Prod. Sci*. **10**: 322–328.

Ohsumi, A., Hamasaki, A., Nakagawa, H., Yoshida, H., Shiraiwa, T. and Horie, T. (2007b). A model explaining genotypic and ontogenetic variation of leaf photosynthetic rate in rice (*Oryza sativa*) based on leaf nitrogen content and stomatal conductance. *Ann. Bot.* **99**: 265–273.

Oka, H.I. and Chang, W.T. (1959). The impact of cultivation on populations of wild rice. *Oryza sativa f. spontanea*. *Phyton* **13**: 105–117.

Oka, H.I. (1956). Observations on natural populations of Formosan wild rice. *Ann. Rep. Natl. Genet*. Jpn. **6**: 46-47.

Oka, H.I. (1961). Report of trip for investigation of rice in Latin America. (mimeo).

Oka, H.I. (1977). Genetic variations of *Oryza glaberima*: their survey and evaluation. **In**: IRAT-ORSTOM meeting on African rice species. Instituto de Recherche Agronomiques Tropicales-Office de la Recherche Scientifique et Technique d’Outre-Mer, Paris, pp 77-86.

Oka, H.I. (1988). Origin of Cultivated Rice. Jap. Sci. Soc. Tokyo.

Oliveira G.X. (1994). Geographic distribution of wild *Oryza* species in Brazil. **In**: Morishima H and Martins PS (eds) Investigations of Plant Genetic Resources in the Amazon Basin with Emphasis on the Genus *Oryza*. Report of 1992/93 Amazon Amazon Project, The Monbusho International Scientific Research Program, Japan, and FAPESP, Brazil, pp 10-15.

Pan, Y., Wang, X., Sun, G., Li, F. and Gong, X. (2016). Application of RAD Sequencing for Evaluating the Genetic Diversity of Domesticated *Panax notoginseng* (Araliaceae). *PLoS ONE* 11: e0166419. doi:10.1371/journal.pone.0166419

Panaud, O., Chenayd, X. and McCouch, S. R. (1996). Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa L.*). *Mol. and General Genet*. **252**: 597-607.

Park, Y.J., Lee, J.K. and Kim, N.S. (2009). Simple sequence repeats polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. *Molecules* **14**: 4546-4569.

Parker, C. and Dean, M.L. (1976). Control of wild rice in rice. *Pesticide Science* **7**: 403-416.

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

Peeters, J.P. and Martinelli, J.A. (1989). Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theor. Appl. Genet*. **78**: 42–48.

Peng, Q., Guo, Q.H., Zhang, X.B., Cheng, D.G., Dai, S., Li, H.S., Zhao, S.J. and Song, J.M. (2012). Evolution in photosynthetic characteristics of wheat cultivars widely planted in Shandong Province since 1950s. *Sci. Agric. Sin*. **45**: 3883–3891.

Peng, S., Laza, R.C., Visperas, R.M., Sanico, A.L., Cassman, K.G. and Khush, G.S. (2000). Grain yield of rice cultivars and lines developed in the Philippines since 1966. *Crop Sci*. **40**: 307–314.

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

Perry, M.C. and McIntosh, M.S. (1991). Geographical patterns of variation in the USDA soybean germplasm collection. I. Morphological traits. *Crop Sci*. **31**: 1350–1355.

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

Poland, J.A., Brown, P.J., Sorrells, M.E. and Jannink, J.L. (2012). Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping by Sequencing Approach. *PLoS ONE* 7(2): e32253. doi:10.1371/journal.pone.0032253.

Prasad, V., Stromberg, C.E., Alimohammadian, H. and Sahni, A. (2005). Dinosaur coprolites and the early evolution of grasses and grazers. *Sci*. **310**: 1177–1180.

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Qian, W., Ge, S. and Hong, D. (2006). Genetic diversity in accessions of wild rice Oryza granulata from South and Southeast Asia. *Genetic Resources and Crop Evolution* **53**: 197–204.

Radin, J.W., Hartung, W., Kimball, B.A. and Mauney, J.R. (1988). Correlation of stomatal conductance with photosynthetic capacity of cotton only in a CO2 enriched atmosphere: Mediation by abscisic acid? *Plant Physiol.* **88**: 1058–1062.

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.

Rajiv, S., Thivendran, P. and Deivanai, S. (2010). Genetic divergence of rice on some morphological and physichemical responces to drought stress. *Pertinaka J*. **32**: 315-328.

Ram, T., Majumder, N.D., Krishnaveni, D. and Ansari, M.M. (2007a). Rice variety Dhanarasi, an example of improving yield potential and disease resistance by introgressing gene(s) from wild species (*O . rufipogon*). *Curr. Sci*. **92**: 987–992.

Ram, T., Majumder, N.D., Mishra, B., Ansari, M.M. and Padmavathi, G. (2007b). Introgression of broad spectrum blast resistance gene(s) into cultivated rice (*Oryza sativa* ssp indica) from wild rice, *O . rufi pogon* . *Curr*. *Sci*. **92**: 225–230.

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.

Sacks, E.J., Dhanapala, M.P., Tao, D.Y., Cruz, M.T. and Sallan, R. (2006). Breeding for perennial growth and fertility in an Oryza sativa/O. longistaminata population. *Field Crops Research* **95**: 39–48.

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. *Crop Sci*. **43**: 120–128.

Sanchez, P.L., Wing, R.A. and Brar, D.S. (2014). The Wild Relative of Rice. **In**: Genetics and Genomics of Rice, Zhang, Q. and Wing, R.A. (eds.), Springer Science+Business Media, New York, USA pp 9-25.

Sano, Y. (1980). Adaptive strategies compared between the diploid and tetraploid forms of *Oryza punctata***.** *Bot. Mag. Tokyo* **93**: 171-180.

SAS Institute (2009). Base SAS 9.1 procedures guide [Online]. Available at http://support.sas.com/documentation/onlinedoc

Second, G. (1985a). Evolutionary relationships in the *Sativa* group of *Oryza* based on isozyme data. *Genet. Sel. Evol.* **17**: 89-114.

Second, G. (1985b). Relations evolutives chez le genre *Oryza* et processus de domestication des riz. **In**: ORSTOM Etudes & Theses, Paris, pp 1- 189.

Second, G., Bezançon, G. and Bozza, J. (1977). Variabilite d'*O. longistaminata* et du complexe sativa des *Oryza* en Afrique. Aspects ecologiques et evolutifs. **In**: Reunion sur les especes africaines de riz. 25–26 Janvier, ORSTOM, Paris, France, pp. 47–55

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.

Shahi, B.B. (1985). Potential rice varieties for East Africa. **In**: International Rice Research Institute (eds.) Rice improvement in Eastern, Central and SNNPR Africa, Proceedings of the International Rice Workshop, April 9-19 1984, Lusaka, Zambia.

Sharkey, T.D. and Raschke, K. (1981). Effect of light quality on stomatal opening in leaves of *Xanthium strumarium* L. *Plant Physiol.* **68**: 1170–1174.

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

Sharma, S.D. and Sampath, S. (1985). Taxonomy and species relationship. **In**: Rice research in India. P.L. Jaiswal, ed. Indian Council of Agricultural Research, New Delhi, pp 21-43.

Sharma, S.D. and Shastry, S.S. (1965). Taxonomic studies in the genus *Oryza* L. III. *Oryza rufipogon* Griff. *Sensu stricto* and *O. nivara* Sharma *et* Shastry *nom. nov*. *Indian J. Genet. Plant Breed*. **25**: 157-167.

Shcherban, A.B., Vaughan, D.A., Tomooka, N. and Kaga, A. (2000). Diversity in the integrase coding domain of a gypsy-like retrotransposon among wild relatives in the *Oryza officinalis* complex. *Genetica* **110**: 43-53.

Shin, Y.B. and Katayama, T. (1979). Cytogenetical studies on the genus *Oryza*. XI Alien addition lines of *O. sativa* with a single chromosomes of *O. officinalis*. *Japan J. Genet*. **54**: 1–10.

Singh, A.K., Singh, P.K. and Rao, M. (2011). Diversity analysis in rice under rainfed condition. *Ann. Pl. Soil Res*. **13**: 58-61.

Sitch, L.A. (1990). Incompatibility barriers operating in crosses of *Oryza sativa* with related species and genera. **In**: Genetic Manipulation in Plant Improvement II, Gustafson, J.P. (ed). Plenum Press, New York pp. 77–94.

Snyderand C.S. and Slaton N.A. (2001). Rice Production in the United States –an Overview. *Better Crops* **85**: 3-7.

Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T.E., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., Fauquet, C. and Ronald, P.C. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science* **270**: 1804–1806.

Sotowa, M., Ootsuka, K., Kobayashi, Y., Hao Y., Tanaka, K., Ichitani, K., Flowers, J.M., Purugganan, M.D., Nakamura, I., Sato, Y., Sato, T., Crayn, D., Simon, B., Waters, D.L., Henry, R.J. and Ishikawa, R. (2013). Molecular relationships between Australian annual wild rice, *Oryza meridionalis*, and two related perennial forms. *Rice* **6**:26.

Souza, E. and Sorrells, M.E. (1991). Relationships among 70 American oat germplasm. I. Cluster analysis using quantitative characters. *Crop Sci*. **31**: 599–605.

Slatkin, M. (1991). Inbreeding coefficients and coalescence times. *Genet. Res*. **58**: 167-175.

Sun, C.Q., Wang, X.K., Li, Z.C., Yoshimura, A. and Iwata, N. (2001). Comparison of the genetic diversity of common wild rice (Oryza rufipogon Griff.) and cultivated rice (O. sativa L.) using RFLP markers. *Theor*. *Appl*. *Genet*. **102**: 157–62

Sun, X.P. and Yang, Q.W. (2009). Comparative study on genetic diversity of wild rice (*Oryza rufipogon* Griff.) in China and three countries in Southeast Asia. *Acta*. *Agro*. *Sinica*. **35**: 679–684.

Swamy, B.M. and Sarla, N. (2008). Yield-enhancing quantitative trait loci (QTLs) from wild species. *Biotechnol*. *Adv*. **26**: 106–120.

Taillebois, J. (1983). Une nouvelle perspective pour la production des semences hybrides F1: le transfert des caracters d'allogamie de l'espece *O. longistaminata* A. Chev. a l'espece *O. sativa* L. *Agronomie Tropicale* **38**: 303–307.

Takai, T., Matsuura, S., Nishio, T., Ohsumi, A., Shiraiwa, T. and Horie, T. (2006). Rice yield potential is closely related to crop growth rate during the late reproductive period. *Field Crops Research* **96**: 328–335.

Tanksley, S.D. and McCouch, S.R. (1997). Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science*, **277**: 1063–1066.

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, pp 22–27.

Tateoka, T. (1962a). Taxonomic studies of *Oryza*. I. *O*. *latifolia* complex. *Bot. Mag. Tokyo* **75**: 418-427.

Tateoka, T. (1962b). Taxonomic studies of *Oryza* II. Several species complexes. *Bot*. *Mag*. *Tokyo* **76**: 165–173.

Tateoka, T. (1964). Notes on some grasses. XVI. Embryo structure of the genus *Oryza* in relation to the systematics. *Am. J. Bot*. **51**: 539-543.

Tateoka, T. (1965). A taxonomic study of *O. eichengeri* and *O. punctata*. *Bot. Mag. Tokyo* **78**: 156-163.

Tateoka, T. and Pancho, J.V. (1963). A cytotaxonomic study of *Oryza minuta* and *O. officinalis*. *Bot. Mag. Tokyo* **76**: 366-373.

Temnykh S., Park, D.W., Ayres, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y.G., Ishii, T. and McCouch, S.R. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Thero. Appl. Genet*. **100**: 697-712.

Temnykh, S., Park, W. D., Ayes, N., Cartinhour, S., Hauck, N., Lipovich, L., Cho, Y. G., Ishii, T. and McCouch, S. R. (2000). Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa L.). Theor. App. Genet*. **100**: 697-712.

Teng, S., Qian, Q., Zeng, D.L., Kunihiro, Y., Fujimoto, K., Huang, D. and Zhu, L.H. (2004). QTL analysis of leaf photosynthetic rate and related physiological traits in rice (*Oryza sativa* L.). *Euphytica* **135**: 1–7.

Tesfaye, Z., Befekadu, A. and Aklilu, A. (2005). Rice production, consumption and marketing: The case of Fogera, Dera and Libokemkem districts of Amhara Region. Paper Presented at Rice Research and Promotion Workshop, June 3-4, 2005. Bahir Dar, Ethiopia.

Thottapilly, G. and Rossel, H.W. (1993). Evaluation of resistance to rice yellow mottle virus in Oryza species. *Indian J. of Virology* **9**: 65-73.

Tiffney, B.H. (2004). Vertebrate dispersal of seed plants through time. *Ann. Rev. Ecol. Evol. Syst.* **35**: 1–29.

Tzvelev, N.N. (1989). The system of grasses (Poaceae) and their evolution. *Bot. Rev*. **55**: 141-204.

Valdiser, P.A., Pappas, G.J., de Menezes, I.P., Müller, B.S., Pereira, W.J., Narciso, M.G., Brondani, C., Souza, T.L., Borba, T.C. and Vianello, R.P. (2016). SNP discovery in common bean by restriction‑associated DNA (RAD) sequencing for genetic diversity and population structure analysis. *Mol. Genet. Genomics* **291**: 1277-1291.

Vales, M. (1985). Study of complete resistance to Pyricularia oryzae of Oryza sativa Oryza longistaminata hybrids and their Oryza longistaminata parent. *Agronomie Tropicale Riz et Riziculture et Cultures Vivrieres Tropicales* **40**: 148-156

Vaughan, D.A. (1988). The genus *Oryza* L. current status of taxonomy. **In**: The International Rice Research Institute Research Paper Series Number 138.

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

Vaughan, D. A. (1994). The Wild Relatives of Rice **In**: A Genetic Resources. International Rice Research Institute, Los Banos, Philippines pp136.

Vaughan, D.A. (1989). The genus *Oryza* L.: current statist of taxonomy. **In**: IRRI Res. Paper Ser. P.138, Int. Rice Res. Inst., Manila, Philippines.

Vaughan, D.A. (1991). Biogeography of the genus *Oryza* across the Malay Archipelago. *Rice Genet*. *Newsl*. **8**: 73–75.

Vaughan, D.A. and Chang, T.T. (1980). Genetic diversity of rice germplasm of selected countries in tropical Asia; Bangladesh, Indonesia and Sri Lanka. *Plant Genet. Resour*. *Newsl*. **42**: 2-6.

Vaughan, D.A. and Morishima, H. (2003). Biosystematics of the genus *Oryza*. **In**: Rice: origin, history, technology, and production, Smith W, editor. John Wiley & Sons, Inc. p, New York, N.Y. (USA), pp 27-65.

Vaughan, D.A. and Muralidharan, V.K. (1989). Collection of wild relative of rice from Kerala State, India. *Plant Genet. Resour. Newsl.* **42**: 2-6.

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

Vaughan, D.A., Ge, S., Kaga, A. and Tomooka, N. (2008). Phylogeny and Biogeography of the Genus *Oryza*. **In**: Rice Biology in the Genomics Era. H.-Y. Hirano et al. (eds.), Springer-Verlag Berlin, Heidelberg, pp 219-234.

Vaughan, D.A., Kadowaki, K., Kaga, A. and Tomooka, N. (2005b). On the phylogeny and biogeography of the genus *Oryza*. *Breed Sci*. **55**: 113–122.

Vaughan, D.A., Morishima, H. and Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Curr. Opi. in Pla. Bio*. **6**:139–146.

Vavilov, N.I., (1951). The Origin, Variation, Immunity and Breeding of Cultivated Plants. *Chron. Bot*. **13**: 1–366.

Venora, G. and Calcagno, F. (1991). Study of stomatal parameters for selection of drought resistant Varieties in Triticum durum DESF. *Euphytica* **57**: 275–283.

Vitte, C. and Panaud, O. (2003). Formation of Solo-LTRs through unequal homologous recombination counterbalances amplifications of LTR retrotransposons in rice, *Oryza sativa* L. *Mol. Biol. Evol.* **20**: 528–540.

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

Wambugu, P.W., Furtado, A., Waters L.E, Nyamongo, D.O. and Henry, R.J. (2013). Conservation and utilization of African *Oryza* genetic resources. *Rice* **6**: 1-13.

Wang MX, Zhang HL, Zhang DL, Qi YW, Fan ZL, Li DY, *et al*. (2008). Genetic structure of Oryza rufipogon Griff. in China. *Heredity* **101**: 527–535.

WARDA (1997). Inter-specific hybridization: progress and prospects. Proceedings of Workshop on: Africa/Asia Joint Research on Inter-specific Hybridization between the African and Asian Rice Species (*O. glaberrima* and *O. sativa*). West Africa Rice Development Association, Bouake, Cote D'Ivoire

Waters, D., Nock, C.J., Ishikawa, R., Rice, N. and Henry, R.J. (2012). Chloroplast genome sequence confirms distinctness of Australian and Asian wild rice. *Ecol. and Evol*. **2**: 211–217.

Watson, L. and Dallwitz, M.J. (1992). The families of flowering plants. **In**: descriptions, illustrations, identification, and information retrieval. Version 2, June 2006, http://delta-intkey.com

Watt, G. (1891). Dictionary of economic products of India. Vol. 5. W. H. Allen, London.

Wenzi, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E. (2004). Diversity Arrays Technology (DArT) for whole-genome profiling of barley. *Proc. Nat. Acad. Sci. USA* **101**: 9915–9920.

Wong, S.C., Cowan, I.R. and Farquhar, G.D. (1985). Leaf conductance in relation to the rate of CO2 assimilation: I. Influence of nitrogen nutrition, phosphorous nutrition, photon flux density, and ambient partial pressure of CO2 during ontogeny. *Plant Physiol*. **78**: 821–825.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**: 97-159

Wright S (1978) Evolution and the Genetics of Population, Variability Within and Among Natural Populations. The University of Chicago Press, Chicago

Xiao, O.J., Grandillo, S., Ahn, S.N., McCouch, S.R. and Tanksley, S.D. (1996). Genes from wild rice improve yield. *Nature* **384**: 223–224

Yan, H.H., Xiong, Z.M., Min, S.K., Hu, Y., Zhang, Z.T., Tian, S.L. and Tang, S.X. (1997). The transfer of brown plant hopper resistance from *O. eichingeri* to *O. sativa*. *Chin. J. Genet*. **24**: 277-284.

Yang, H., Hu, L., Hurek, T. and Reinhold-Hurek, B. (2010). Global characterization of the root transcriptome of a wild species of rice, Oryza longistaminata, by deep sequencing. *BMC Genomics* **11**: 705–713

Yang, K.S., Ho, K.C., Li, H.W. (1965). Cytological studies of *O. sativa L.* and its related species. 8. Study on meiotic division of F1 hybrid of *O. sativa* (L.) X *O. brachyantha,* A. Chev. *et* Roehr. *Bot. Bull. Acad. Sin*. **6**: 32-38.

Yeo, M.E., Yeo, A.R., Flowers, T.J. (1994). Photosynthesis and photorespiration in the genus *Oryza*. *J. Exp. Bot*. **45**: 553-560.

Yue, B., Xue, W.Y., Luo, L.J. and Xing, Y.Z. (2006). QTL analysis for flag leaf characteristics and their relationships with yield and yield traits in rice. *Acta. Genet. Sin.* **33**: 824–832.

Zhang, Q. and Kochert, G. (1998). Independent amplification of two classes of Tourists in some *Oryza* species. *Genetica* **101**: 145-152.

Zhang, Q., Wang, C.L., Shi, A.N., Bai, J.F., Ling, S.C., Li, D.Y., Chen, C.B. and Pang, H.H. (1994). Genetic studies on resistance to bacterial blight (Xanthomonas oryzae pv. oryzae in nine rice cultivars. *Scientia Agricultura Sinica* **27**: 1-9.

Zhao, M., Ding, Z., Lafitte, R., Sacks, E., Dimaygua, G. and Holt, D. (2010). Photosynthetic characteristics in *Oryza* species. *Photosynthetica* **48**: 234–240.

Zhao, X.Q., Zhao, M., Xiao, J.T., Zhang, W.X., Guan, D.M., Wang, M.Y., Lu, J. and Zhang, N. (2003). Stomatal characteristics of leaves of high-photosynthetic efficiency progenies from a cross between *Oryza sativa* and *Oryza rufipogon* and their parents. *Acta*. *Agron*. *Sin*. **29**: 216–221.

Zheng, T.C., Zhang, X.K., Yin, G.H., Wang, L.N., Han, Y.L., Chen, L., Huang, F., Tang, J.W., Xia, X.C. and He, Z.H. (2011). Genetic gains in grain yield, net photosynthesis and stomatal conductance achieved in Henan Province of China between 1981 and 2008. *Field Crops Res*. **122**: 225–233.

Zhou, H.F., Xie, Z.W. and Ge, S. (2003). Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (Oryza rufipogon Griff.) in China. *Theor Appl*. *Genet*. **107**: 332–339.

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.

Zhu, Q., Zheng, X., Luo, J., Gaut, B.S. and Ge, S. (2007). Multilocus analysis of nucleotide variation of Oryza sativa and its wild relatives: severe bottleneck during domestication of rice. *Mol. Biol. Evol.* **24**: 875–888.

Zhu, X.G., Long, S.P. and Ort, D.R. (2008). What is the maximum efficiency with which photosynthesis can convert solar energy into biomass. *Curr Opin Biotechnol*. **19**: 153–159.

Appendix 1. Mean performances of 8 quantitative traits of 62 *O. longistaminata* accessions from three replications by accession

Accession No	PH	PL	LL1	LL2	LL3	LW1	LW2	LW3
2	181.94	27.84	42.88	63.43	65.39	1.74	1.68	1.71
3	186.33	27.17	37.40	46.93	57.63	1.60	1.58	1.55
4	195.37	26.42	47.18	63.23	57.90	1.73	1.73	1.9
5	183	27.07	41.47	64.27	65.07	2.00	1.83	1.63
6	187.33	26.00	33.87	49.47	51.50	1.50	1.70	1.67
7	183.78	24.77	38.44	56.69	58.76	1.84	1.69	1.62
8	196.67	27.30	32.13	46.87	53.67	1.67	1.80	1.63
9	192.17	29.27	42.57	61.57	63.97	1.77	1.60	1.57
10	206.03	32.83	40.33	64.40	67.27	1.77	1.67	1.47
11	197.73	27.82	35.22	47.30	57.20	1.70	1.67	1.52
12	158.83	21.82	32.00	46.37	56.53	1.45	1.55	1.50
13	178.83	25.52	33.39	45.27	55.13	1.45	1.54	1.57
14	160.87	29.17	46.47	48.40	43.67	1.80	1.63	1.63
15	170.73	22.33	30.67	46.77	62.37	1.30	1.20	1.23
16	156.43	15.70	25.50	40.50	51.63	1.23	1.3	1.5
17	185.67	27.73	37.47	51.57	54.33	1.77	1.73	1.73
18	181.23	26.50	37.50	60.71	58.74	2.03	1.87	2.1
19	146.50	28.60	42.20	62.00	64.40	2.00	1.80	1.90
20	180.83	26.68	26.37	41.50	52.38	1.35	1.38	1.32
23	177.95	28	30.48	51.42	53.57	1.45	1.50	1.53
25	191.17	30.97	45.47	57.97	53.87	1.80	1.67	1.6
26	145.67	27.97	42.20	61.97	63.80	1.67	1.57	1.63
27	165.6	22.67	35.47	45.80	44.47	1.57	1.73	1.47
28	174.58	25.67	31.58	50.93	55.68	1.43	1.47	1.53
29	181.88	28.08	38.63	50.02	58.13	1.55	1.48	1.55
30	151.1	27.57	39.03	57.10	54.87	1.90	1.90	1.77
31	173.5	26.25	37.95	58.8	58.1	1.8	1.75	1.75
32	165.15	22	32.85	47.15	54.4	1.6	1.5	1.48
33	146.57	21.70	25.40	33.53	52.87	1.30	1.37	1.53
34	172.09	25.73	34.37	51.49	58.61	1.40	1.34	1.38
35	159.6	23.93	37.70	57.93	62.70	1.33	1.23	1.2
36	167.73	26.07	34.50	54.07	62.90	1.37	1.3	1.4
37	175.2	24.02	28.18	50.78	56.02	1.48	1.62	1.57
38	162.63	23.43	32.00	39.93	62.57	1.50	1.40	1.67
39	191.97	29.63	43.13	58.81	58.72	1.67	1.57	1.47
40	201.58	31.02	40.22	56.02	67.28	1.56	1.61	1.69
41	167.65	26.80	42.88	46.67	62.92	1.73	1.67	1.73
42	195.63	25.53	41.83	62.00	63.03	2.07	1.80	1.67
43	184.5	27.8	45.83	63.10	62.03	1.47	1.50	1.53
44	180.43	25.77	39.37	53.10	55.63	1.57	1.57	1.67
45	184.98	26.22	37.33	52.40	57.59	1.70	1.64	1.70
46	178.43	25.87	35.56	50.67	55.86	1.62	1.54	1.66
47	178.68	27.51	38.77	57.42	60.25	1.61	1.56	1.49

Appendix 1 continued

Accession No	PH	PL	LL1	LL2	LL3	LW1	LW2	LW3
48	176.24	25.83	40.86	50.27	56.34	1.53	1.50	1.52
49	174.52	28.43	39.44	55.27	53.53	1.68	1.61	1.75
50	173.28	27.92	33.78	48.27	55.95	1.37	1.38	1.38
51	183.48	27.98	40.73	55.70	51.68	1.90	1.85	1.80
52	170.02	26.98	28.13	46.13	52.65	1.45	1.57	1.68
53	172.76	24.66	32.81	45.39	48.32	1.59	1.53	1.61
54	165.40	25.89	35.04	50.24	50.41	1.62	1.52	1.57
55	160.47	20.12	30.47	46.25	52.88	1.50	1.40	1.47
57	164.93	20.80	30.33	45.24	50.26	1.49	1.38	1.46
58	147.13	21.16	26.94	40.59	46.79	1.39	1.37	1.42
60	157.5	20.75	27.78	47.91	54.45	1.40	1.35	1.33
63	162.7	25.67	36.79	54.87	53.80	1.97	1.77	1.73
64	194.2	29.8	35.53	52.64	45.97	1.77	1.50	1.77
65	186.98	29.08	43.43	52.32	56.95	1.88	1.75	1.95
66	175.43	25.03	43.00	54.63	69.17	1.53	1.43	1.57
71	194.58	29.20	40.95	56.80	61.65	1.88	1.82	1.77
72	194.21	31.22	38.66	58.18	62.64	1.83	1.87	1.77
73	187.75	30.42	38.10	53.43	57.58	1.73	1.63	1.6
74	190.91	30.98	36.86	60.23	65.53	1.66	1.52	1.43
Grand mean	**176.40**	**26.33**	**36.67**	**52.43**	**57.10**	**1.63**	**1.58**	**1.60**

PH (plant height), PL (panicle length), LL1 (first leaf length), LL2 (second leaf length), LL3 (third leaf length), LW1 (first leaf width), LW2 (second leaf width), LW3 (third leaf width)