GENETIC DIVERSITY OF RICE (*Oryza sativa* L.) LANDRACES CONSERVED AT THE NATIONAL GENEBANK AS REVEALED BY SIMPLE SEQUENCE REPEAT (SSR) DNA MARKERS

EMMANUEL ABRAHAMU MAUSA

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN MOLECULAR BIOLOGY AND BIOTECHNOLOGY OF SOKOINE UNIVERSITY OF AGRICULTURE.

MOROGORO, TANZANIA.

2014

EXTENDED ABSTRACT

There is high degree of rice diversity in Tanzania. Increased human population pressure and activities has caused the population of wild rice to disappear at alarming rate. NPGRC collected over 125 accessions of rice landraces and 7 wild relatives for conservation and sustainable utilization in improvement programs, but no morphological or genetic diversity study that have been done to support any future collections. An investigation was conducted to determine the extent of genetic diversity and relationship among 79 rice (*Oryza sativa* L.) landraces and its wild relative conserved at the NPGRC in Tanzania. Fourteen quantitative morphological characters analysed indicated that rice germplasm conserved at the NPGRC has considerable diversity range. The yield related characters such as grain length (r = 0.360), flag leaf width (r = 0.511), and one hundred seed weight (r = 0.319) showed significant and positive association with grain yield per accession. Principal components analysis indicated that the first four components with Eigen values >1 accounted for 63.99% of the morphological variability among rice accessions studied. The cluster analysis grouped accessions into four groups, one group for wild rice accession and the rest three groups for mixed accessions. Five SSR primers were used to determine genetic diversity, 11 polymorphic alleles were revealed, alleles per locus ranged from 2 to 3. Primer RM 333 revealed large number of alleles. Jaccard's similarity coefficient revealed that most of the accessions evaluated were genetically similar, except few are distinct. Study also highlighted use of large number of SSR markers (>5) for efficient characterization of the rice accessions conserved at the genebank and those used in this study. This work is expected to be published in African Journal of Biotechnology (AJB); Academic Journal.

DECLARATION

I, Emmanuel A. Mausa, do hereby declare to the Senate of Sokoine University of Agriculture that this dissertation is my own original work done within the period of registration and that it has neither been submitted nor being concurrently submitted in any other institution.

_____ _____

Emmanuel A. Mausa Date

(MSc. Candidate)

The above declaration is confirmed:

_____ _____

Dr. Kusolwa P.M Date

(Supervisor)

COPYRIGHT

ACKNOWLEDGMENTS

I would like to express my sincere thanks to the Director General of the Tropical Pesticides Research Institute for allowing me to attend this course. Dr. Margaret M. Mollel, the Ag.Curator of the National Plant Genetic Resources Centre (NPGRC) for providing me field samples and William Hamisy the officer in-charge Biotechnology Laboratory for availing the Laboratory facilities that made this research work possible. Moreover, invaluable thanks should be extended to my supervisor Dr. P. M. Kusolwa from Sokoine University of Agriculture for his criticism and direction that made me completes this work.

I also express my heartfelt thanks to my beloved wife Damary, my daughters Glory and Gift for their support and encouragement. Last but not least, I am greatly indebted to every one who supports me morally and, materially. God bless you all. AMEN!

DEDICATION

This work is dedicated to my beloved wife late Damary Yesaya Nnko and Almighty God

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

LIST OFABBREVIATION AND SYMBOLS

EDTA	Ethylenediaminetetra acetic acid
FAO	UN Food and Agriculture Organization
IBPGRI	International Board for Plant Genetic Resources Institute
IRRI	International Rice Research Institute
MAFS	Ministry of Agriculture and Food Security
NPGRC	National Plant Genetic Resources Centre
PCA	Principal Components Analysis
RAPD	Random Amplified Polymorphic DNA
RFLP	Random Fragment Length Polymorphism
RM	Random Marker
SNPs	Single Nuclear Polymorphism
SSR	Simple Sequence Repeats
Taq	Thermus aquaticus
TPRI	Tropical Pesticides Research Institute
Tris	tris (hydroxymethyl) methylamine
UPGMA	Unweighted Pair Group Method with Arithmetic
UVL	Ultraviolet light
V/V	Volume by Volume
W/V	Weight by Volume

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background Information

Rice (*Oryza sativa* L.) is one of the most important food crops in the world, consumed daily by more than half of the human population (Muhammed *et al*., 2009). Most of the world's rice is cultivated and consumed in Asia, which constitutes more than half of the global population (Chakravarthi and Naravaneni, 2006). In developing countries rice accounts for 715Kcal per capita/day and provides 27% of global human per capita energy, 20% of per capita protein, and 3% dietary fat (FAO, 2002). Rice can contribute nutritionally significant amounts of thiamin, riboflavin, niacin, and zinc to the diet, but lesser amounts of other micronutrients. Rice also provides minerals, vitamins, and fiber, although all constituents except carbohydrates are reduced by milling (Mario, 2007).

Rice is one of the widely grown crops in Tanzania and is the second most important food crop in terms of number of households, area planted, and production volume after maize. It is estimated that, about 60% of Tanzanian population consumes rice and its derivatives per day each year (Kanyeka *et al*., 1994). Rice also plays a great role in human being by which, stalks are utilized as animal feed and thatching materials (Kanyeka *et al*., 1994). Besides meeting local consumption demands, the rice sector is a major source of income and employment in rural areas.

There is a high degree of rice diversity in Tanzania, including rice landraces and its wild weed relatives. Genetic diversity is influenced by selection, mutation, population size, and genetic drift (Hedrick, 2005) and understanding how each of these factors influences the genetic diversity of a population is critical to the conservation of species. Landraces of

rice played a very important role in the local food security and sustainable development of agriculture, in addition to their significance as genetic resources for rice genetic improvement (Tang *et al*., 2002). Among wild rice relatives *O. barthii, O. longistaminata, O. punctata, O. eichingeni,* and *O. brachyantha* have been identified in different agro-ecological zones of Tanzania (Kiambi *et al*., 2005). The wide variation in temperature, rainfall, topography, and soil in the country has provided a wide diversity of ecosystems resulting in rich diversity of these plant species. The diversity of the wild species is gradually being eroded for multiplicity of reasons. These include destruction of natural wild rice habitats to pave the way for expanding agricultural activities resulting from increasing human population pressure, overgrazing, changing in land use and deforestation (Kiambi *et al*., 2005). Therefore, to maintain crop diversity, collection, characterization and conservation of traditional landraces and wild/weed relatives are vital.

The National Plant Genetic Resources Centre (genebank) of Tanzania through multi-crop collection missions collected over 125 accessions of rice landraces and 7 wild relatives (Akonaay and Milinga, 2002), for the purpose of conservation and sustainable utilization of these plant genetic resources (Ching'ang'a, 2002). The knowledge of the amount, the extent, and distribution of genetic variation of the collected accessions is vital to the improvement and development of effective conservation strategies. The evaluation of the genetic variability of these accessions can therefore provide the basic information necessary to help gene-bank manager to multiply and properly conserve the genetic resources; Also to breeders in their breeding programs to plan for crosses to incorporate this different variability into the genetic background of elite rice lines, which in turn will generate new rice varieties. The analysis of genetic variation both within and among elite breeding materials is of fundamental interest in plant breeders as it contributes to

monitoring germplasm and can be used to predict potential gain (Kalyan and Rambabu, 2006).

The study of morphological and agronomical traits is the classic way of assessing genetic diversity for the gene-bank managers and plant breeders as well. Although morphological traits coupled with statistical methods have been successfully used in characterization of germplasm for conservation and selection of superior individuals in plant improvement programs, it is often disguised by the factors in the environment. Moreover, an assessment of genetic diversity based only on morphological traits might be biased because distinct morphological types can result from a few mutations and share a common genetic background (Lanaud and Lebot, 1997).

The significant advancements in molecular biology have shifted the focus of assessment of genetic diversity from relying on morphological markers to using molecular markers. The molecular or DNA markers systems have many advantages over the traditional morphological and protein markers that are used in genetic and ecological analyses of plant populations; firstly, an unlimited number of DNA markers can be generated; secondly DNA marker profiles are not affected by the environment and thirdly DNA markers, unlike isozymes markers are not constrained by tissue or development stage specificity (Yong-Jin *et al*., 2009).

Of the available molecular marker systems, microsatellites loci, or simple sequence repeats (SSR) have become major molecular markers for a wide range of studies in plants and animals after its emergence as a Polymerase Chain Reaction (PCR)-based genetic marker (Chen *et al*., 2002). SSR markers have many advantages over others marker systems; it has higher reproducibility which is most important in genetic analysis, is co-

dominant in nature, bands produced from the same set of primers are intuitively orthologous, abundant in genome of all species, and well distributed in their genome (Wang *et al*., 1994).

Since the inception of conservation of germplasm strategies by the National Plant Genetic Resources Centre (NPGRC), a good number of landraces of rice germplasm have been collected and seeds conserved in the genebank. However, no morphological or genetic diversity study that have been done to support any future collection strategies. The combined use of morphological and molecular methods for analysis of genetic diversity of rice landraces and its wild relatives conserved at the National Plant Genetic Resources Centre would provide useful information to allow for exploitation of the potential genetic resources available. This study therefore, has focused on both morphological and Simple Sequence Repeat (SSR) as marker systems for the assessment of genetic diversity of 79 accessions of rice germplasm conserved at the Plant Genetic Resources Centre (National gene-bank) in Tanzania.

1.2 Objectives

1.2.1General objective

Genetic diversity of rice (*Oryza sativa*) landraces and its wild relatives conserved at the National Plant Genetic Resources Centre in Tanzania.

1.2.2 Specific objectives

i. To determine morphological diversity of rice accessions conserved at the National Plant Genetic Resources Centre in Tanzania.
ii. To estimate genetic relatedness of rice accessions conserved at the gene - bank and other collections from different rice growing areas in Tanzania using SSR markers.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Systematic and Distribution of Rice

Rice (*Oryza sativa* L.), belongs to the family *Poaceae* and the genus *Oryza* (Sasaki and Burr, 2000). It's widely recognized as an excellent model plant for the study of grass genetics and genome organization due to its diploid origin (2n = 24). Also it has relatively small genome size (430 Mb), availability of whole genome sequences, relative ease of transformation, development of several key genomic mapping resources. Moreover, it has a considerable level of genetic polymorphism, large amount of well conserved genetically diverse material and the availability of widely collected, compatible wild species and is grown under diverse cultural conditions and over wide geographical range (Malik *et al*., 2010). Rice occupies almost one-fifth of the total land area covered under cereals. It is cultivated over an area of 1.5 billion hectares of land having overall worldwide production of 645 million tons per annum, and most of the world's rice is cultivated and consumed in Asia, which constitutes more than half of the global population. Approximately 11% of the world's arable land is planted annually with rice, and it ranks next to wheat (Chakravarthi and Naravaneni, 2006).

Tanzania is the second largest producer of rice in Eastern and Southern Africa after Madagascar with production level of 818 000 tones (MAFS, 2009). The cultivated area is 681 000 ha; this represents 18 % of Tanzania's cultivated land. Major rice production systems are lowland rain fed (74%), upland rain fed (20%), and irrigated lowland (6%) (Kanyeka, 1994), and it's planted in both Mainland and Zanzibar Islands. The average yield is very low, 1-1.5 t / ha compared to the estimated potential yield of 4-5 t/ha (MAFS, 2009). Farmers grow a number of traditional varieties, these varieties have long

maturity, and yield is affected by both biotic and abiotic constraints which include diseases, insect pests, weeds, and drought among others. In this case the future paddy rice production in Tanzania will depend on improved varieties and the plant breeders have to utilize the genetic diversity of rice germplasms conserved at the National Plant Genetic Resources Centre (genebank) for improvement and breeding programs.

2.2 Morphological Characteristics

Morphological characters have been routinely used to analyze genetic diversity, but did not have the resolution power for revealing polymorphisms in genetic analyses and/or for differentiating between closely related genotypes. The major disadvantages associated with it are the limited number of morphological characters and affected by the environment (Thaura *et al*., 2008). As a consequence different plant genotypes cannot always be distinguished (Meglic and Staub, 1993). Morphological markers are influenced by the environmental conditions, are labor intensive, and time consuming.

The discriminators mostly used in morphological diversity study of rice genetic resources are those approved by International Board for Plant Genetic Resources (IBPGR) and IRRI (IBPGR, 1980) rice advisory committee on rice standard descriptors. These descriptors include both qualitative and quantitative; seedling height, leaf bladed length, leaf blade width, ligules length, Culm length, Culm number after full heading, Culm diameter at flowering period, panicle length, 100 seed weight at maturity, grain length and grain width measured as quantitative characters, and later at the vegetative stage leaf blade pubescence, leaf blade color, basal leaf sheath color, leaf angle, flag leaf angle, ligule color, ligule shape, collar color, auricle color, Culm angle after flowering, internodes color after flowering, panicle type at near maturity, secondary branching, panicle exertions, seed coat color at maturity (Samarajeewa *et al.*, 2004).

2.3 Chemical Markers

Biochemical markers such as Isozymes and protein patterns, though minimally influenced by the environment, offer limited polymorphism and often do not allow discrimination between closely related genotype (Ram *et al*., 2010).

2.4 Genome Composition of Rice

The total length of rice nuclear genome was calculated to be 388.8Mb of DNA distributed among the 12 chromosome pairs, and includes the genes that encode some 38 000 proteins (Heslop and Thomas, 2007). The pseudo molecules are expected to cover 95.5% of the entire genome and an estimated 98.9% of the euchromatin (Heslop and Thomas, 2007). The genome also consists of one pair of circular mitochondrial DNA, and a circular chloroplast DNA (Takuji *et al*., 2005).

2.5 Molecular Markers for Genetic Studies

Advances in plant genetics and molecular biology have led to the development of many types of molecular markers that can be used to characterize germplasm. Different types of DNA markers are available, each method differing in principle, application, type, and amount of polymorphism detected, cost, and requirement (Michael, 1997). These include the first generation DNA markers system which employed southern blot based markers such as RFLP (restriction fragment length polymorphism), which results from mutation in recognition sites. RFLP is non-PCR based marker. Botstein *et al.* (1980) used RFLP marker for construction of genetic map for the first time. It was the first technology that enabled the detection of polymorphism at DNA sequence level. In this technique total genomic DNA is digested by restriction endonuclease followed by hybridization with a radioactively labeled probe or other tagging systems.

As a result, different sized hybridization fragments are revealed. Though RFLP marker have been developed in some crops like *Eucalyptus nitens* (Byrne *et al*., 1995), *Vigna subterranea* (Ntundu, 2006), it has been criticized primary related to the use of short-lived radioisotopes and the technical complexity involved in performing such analysis (Karp and Edwards, 1997), together with the high cost often limits their utilization in genetic analyses of large populations.

The second generation DNA markers for genetic analysis were those derived from (PCR) polymerase chain reaction (Mullis *et al.*, 1986). PCR revolutionized genetic and ecological analyses of populations in several ways because it had two major advantages over Southern blot based markers. First, it requires only small amounts of DNA to allow analysis at very early stages, thus reducing the need for plant nurseries. Second, it is inexpensive, and simple enough that large scale experiments can be carried out rapidly on a large scale (Ayard *et al*., 1997). RAPD, AFLP, and SSR are the major molecular marker systems with the other systems being modifications of these three.

Random amplified Polymorphic DNA (RAPD): This maker system was developed by Welsh and McClelland in 1991, Ram *et al*. (2010). RAPD is performed in condition resembling those of PCR using genomic DNA and a single short oligonucleotide. The DNA amplification product is generated from region that is flanked by a part of 10 bp priming site in the opposite orientation. RAPD markers had been used by various workers in different species of plants to find out genetic diversity. Quintela-Sabaris *et al*. (2005) and Ram *et al*. (2010), used RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection.

Amplified Fragment Length Polymorphism (AFLP): It is a combination of RFLP and RAPD. This involves major steps as follows; cutting of DNA with restriction enzymes and double stranded oligonucleotides adapters are ligated to the ends of DNA fragments. Selective amplification of sets of restriction fragments is usually carried out with P^{32} labeled primers designed according to sequence of adapters plus 1-3 additional nucleotides, gel electrophoresis and analysis of amplified fragments. This is highly method for detecting polymorphism throughout the genome and it's popular for species whose primers have already been developed. Kiambi *et al*. (2005) clearly demonstrated the usefulness of AFLP in studying diversity in rice populations and its power to discriminating between populations and individuals within population.

The third generation of molecular markers is the system that utilizes SNPs (single nucleotide polymorphisms), ESTs (Expressed Sequence Tag) Markers, and microarrays. Single nucleotide polymorphism SNP's, represent sites in the genome where DNA sequence differs by a single base when two or more individuals are compared. They may be individually responsible for specific traits or phenotypes, or may represent neutral variation that is useful for evaluating diversity in the context of evolution.

The frequency of occurrence of SNP in a genome is generally one SNP in every 100-3000 bp (Yong-Jin *et al*., 2009). The SNPs can be detected by two ways one is gel based assays and other is non gel based assays. Kim *et al*. (2010) pointed out that SNPs are widely used in breeding programs or several applications such as in marker assisted and genomic selection, association and QTL mapping, positional cloning, haplotype and pedigree analysis, seed purity testing, variety identification and monitoring the combinations of alleles that perform well in target environments. Although SNP's seems to have several advantages over other technologies, the study by Schlötterer (2004) shows that SNPs

genetic markers systems suffer a few shortcomings such as; they are usually biallelic, so the information content of a single SNP is limited particularly if one of the two alleles occurs at a low frequency, the development of a set of SNP markers is time consuming and cost intensive; also SNPs might be located at the hyper mutable sites which violets the assumption that they are bi-allelic often made when analyzing SNPs for population genetics purposes. For the case of ESTs markers; such markers are obtained by partial sequencing of random c-DNA clones and are useful in cloning specific genes of interest and syntenic mapping (mapping of gene loci that lie in the same order on the same chromosome). This is also used in full genome sequencing and mapping programmes and isolation of genes.

2.6 Simple Sequence Repeats Markers

Simple sequence repeats (SSR) markers are di-, tri-, and tetra-nucleotide tandem repeats containing loci of eukaryotic genomes. These are actually non-coding regions which remained conserved during the course of evolution and are ideal for DNA fingerprinting and varietal identification. They are valuable as genetic markers as they detect high level of allelic diversity based on the variability in the tandem repeats in the core unit and are co- dominant. It is demonstrated that these loci are very polymorphic due to changes in the number of repeating units among the individuals of populations. Each SSR locus can easily be amplified by using PCR knowing the DNA sequence flanking the repeat region specifically.

The limiting feature of the application of these markers is the need for prior sequence information for developing primers for locus-specific PCR amplification. This limitation is alleviated for the economically important species and the ones closely related, since primer sequences of the SSR DNA markers and the amplification conditions are available

in the published reports. However, when the reported PCR amplification conditions were applied, not all primer pairs produce specific markers, which are specific for the locus. Generation of complex banding patterns for SSR loci could be due to various reasons such as type of repeat, non-optimization of PCR conditions and the nature of genome. Though SNPs and ESTs DNA markers are available, SSR were chosen for the analysis of genetic diversity of rice landraces and wild relatives in this study because several works have shown that these markers are very powerful for differentiating individual germplasm accessions, particularly when they are closely related (Bligh *et al.,* 1999; Xu *et al*., 2004; Jeung *et al.,* 2005). The experimental procedure of SSR analysis has been reported by Yong-Jin *et al*. (2009) to be simpler and require only small amount of DNA template.

Furthermore, SSR markers have higher reproducibility which is most important in genetic analysis, is co-dominant in nature. Therefore, the SSR bands produced from the same set of primers are intuitively orthologous, abundant in genome of all species, and well distributed in their genome (Wang *at el*., 1994). The hyper-variable nature of SSRs produces very high allelic variations even among very closely related varieties. Therefore, SSR markers have been found to be ideal DNA markers for genetic mapping and population studies; because of their abundance. Genetic diversity studies using SSR marker have been developed for a number of crops like maize (Chin *et al*., 1996), sorghum (Tarmino *et al*., 1997), wheat (Roder *et al*, 1998), soybean (Cregan *et al*., 1999), *rice* (Ramendra and Bahar, 2011); wild rice (Samarajeewa *et al*., 2004), and *jatropha curcas* (Umamaheswari *et al.*, 2010).

The SSR markers have been increasingly applied by many scientists in rice germplasm. Priti *et al*. (2011) studied genetic diversity of popular of 29 rice varieties in India using 12 SSR markers and identified genotype specific alleles in 14 popular rice varieties which

can be employed in true identification germplasm in their country. Herrera *et al.* (2008) assessed genetic diversity in Venezuelan rice cultivars using simple sequence repeat markers to broaden the genetic bases of rice germplasm in the country. The genetic diversity reported was very low, but this work proved SSR to be an efficient tool in assessing the genetic diversity of rice genotypes.

Furthermore, Claudio *et al*. (2006) characterized the allelic diversity of 192 traditional varieties of Brazilian rice using 12 simple sequence repeat markers. The study revealed identical accessions with the same name, few with different names and a mixture of pure lines, indicating that SSR markers are fundamental to determining the genetic relationship between landraces. They further concluded that the most variable set of genotypes analyzed can be used as progenitors to increase the genetic variability to rice breeding programs. Bakari (2010) used SSR markers to analyze 70 samples of Tanzanian rice landraces, his study showed no significant genetic variation among the landraces studied. Other scientists like Umamaheswari *et al.* (2010); Virk *et al*. (1995); Samarajeewa *et al*. (2004), Huang *et al*. (2010); Marie *et al*. (2010); Malik *et al*. (2010); Wong *et al*., 2009; Qian and Hong, (2001), Ram *et al.* (2010) conducted genetic diversity studies of both cultivated crop species very successfully with this DNA markers.

2.7 References

Akonaay, H. B. and Millinga, P. L. (2002). A catalogue of plant genetic resources under ex-situ conservation at the NPGRC-Arusha. *Proceedings of the Second National Workshop on Plant Genetic Resources and Biotechnology*, Arusha, Tanzania, 6 – 10 May 2002. pp. 522 – 536.

Ayad, W., Hodgkin, G., Jaradat, A. and Rao, V. R. (1995). Molecular genetic Techniques for plant Genetic resources. *Report of an International Board for Plant Genetic Resources Institute Workshop,* Rome, Italy, 9 – 11 October, 1995. 141pp.

Bakari, M. M. (2010). Genetic diversity of some rice (*Oryza sativa* L.) landraces grown in Tanzania using simple Sequence repeats (SSR). Dissertation for Award of MSc Degree at Sokoine University of Agriculture, Morogoro, Tanzania, 72pp.

Bligh, H. F. J., Blackhall, N. W., Edwards, K. J. and McClung, A. M. (1999). Using amplified length polymorphisms and simple sequence length polymorphisms to identify cultivars of brown and white milled rice. *Crop Science* 39(6): 1715 – 1721.

Botstein, D., White, R. L., Slolnick, M. and Davis, R. W. (1980). Construction of genetic linkage map using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32(3): 314 – 331.

Byrne, M., Murrell, J. C., Tibbits, W. N. Moran, G. L. (1994). Detection and inheritance of RFLPs, in *Eucalyptus nitens*. *Theory Applied Genetics* 89: 397 – 402.

Chakravarthi, B. K. and Naravaneni, R. (2006). SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L.). *African Journal of Biotechnology* 5: 684 – 688.

Chen, X., Cho Y. G. and McCouch, S. R. (2002). Sequence divergence of rice Microsatellite in *Oryza Sativa* L. and other plant species. *Molecular Genetics and Genome* 268: 331 – 343.

Ching'ang'a, H. M. (1990). Role of the National gene bank in the conservation of Plant Genetic Resources in Tanzania. *Proceedings of the First National Workshop,* Arusha, Tanzania, 16 - 20 January 1990. pp. 30 – 41.

Chin, E. C., Senior, M. L., Shu, H. and Smith, J. S. C. (1996). Maize simple repetitive DNA sequences: Abundance and allele variation. *Genome* 38: 866 – 873.

Claudio, B., Tereza, C. O. B. and Paulo, H. N. R. (2006). Determination of genetic diversity of traditional varieties of Brazilian rice using microsatellite markers. *Genetic and Molecular Biology* 29(4): 676 – 684.

FAO (2002). Nutritional Contribution of Rice: Impact of biotechnology and Biodiversity in rice consuming countries Thailand. *Journal of Sensory Studies* 23(1): 1 – 13.

Gregan, P. B., Jervik, T., Bush, A. L., Shoemarker, R. C., Lark, K. G., Kahlar, A. L., Kaya, N., VonToai, T. T., Lohnes, D. G., Chung, J. and Specht, J. E. (1999). An integrated molecular genetics linkage map of soybean genome. *Journal of Crop Science* 39: 146 – 149.

Hedrick, P. W. (2005). *Genetic Population.* (3rd Ed.), Jones and Bartlett Publisher Company, Sudburry, USA. 737pp.

Herrera, T. G., Duque, D. P., Almeida, I. P., Nunez, G. T., Pieters, A. J., Martinez, C. P., and Tohme, J. M. (2008). Assessment of genetic diversity in Venezuelan rice cultivars using simple sequence repeat markers. *Electronic Journal of Biotechnology* 11(5): 1 – 14.

Heslop, H. and Thomas, S. (2007). Plant nuclear genome composition. *Encyclopedia of Life Science* 3: 1 – 8.

Huang, M., Xie, F., Yun, C. L., Zhao, X., Jojee L. and Madonna, D. (2010). Comparative analysis of genetic diversity and structure of rice using ILP and SSR markers. *Rice Science* 17(4): 257 – 268.

IBPGR/IRRI Rice Advisory Committee (1980). *Descriptors for Rice* (*Oryza sativa* L.) International Rice Research Institute, Philippines. 40pp.

Jeung, J. U., Hwang, H. G., Moon, H. P. and Jena, K. K. (2005). Fingerprinting Temperate japonica and tropical indica rice genotypes by comparative analysis of DNA markers. *Euphytica* 146(3): 239 – 251.

Kalyan, B. C. and Rambabu, N. (2006). SSR based DNA finger printing and diversity Study in rice (*Oryza sativa* L.). *African Journal of Biotechnology* 5(9): 684 – 688.

Kanyeka, Z. L., Msomba, S. W., Kihupi, A. N. and Penza, M. F. (1994). Rice ecosystems in Tanzania: Characterization and classification. *Research and Training Newsletter* 9:13-15.

Karp, A. and Edwards, K. J. (1997). Techniques for the analysis, characterization, and Conservation of plant genetics. Molecular techniques in the analysis of the extent of genetic diversity. (Edited by Ayad, W. G., Hodging, T., Jaradat, A. and Rao, V. R.), *Reports of an International Plant Genetic Resource Institute Workshop* 9 – 11 October, 1995, Rome, Italy pp. 9 – 47.

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

Kim, J. S., Ahn, S. G., Kim, C. K. and Shim, C. K. (2010). Screening of rice blast resistance Genes from aromatic rice germplasms with SNP markers. *Plant Pathology Journal* 26: 70 – 79.

Lanaud, C. and Lebot, V. (1997). Increasing the use of plant genetic resources. In: Molecular genetic techniques for plant genetic resources. (Edited by Ayad, W. G. Hodgkin, T. Jaradat A. and Rao V. R.), *Report of an International Plant Genetic Resource Institute Workshop* 9 – 11 October 1995, Rome, Italy. pp. 92 – 97.

Malik, A. R., Muhammad, S. M., Zabta, K. S. and Kazuko Y. S. (2010). Genetic analysis of basmati and non-basmati Pakistan rice (Oryza sativa L.) cultivars using microsatellite markers. *Pakistan Journal of Botany* 42(4): 2551 – 2564.

Marie, N. N., Fousseyni, C., Gezahegn, G., Mounirou, S., Roland, B., Gustave, D. and Fatondji, B. (2010). Morpho-agronomic and molecular characterization of oryza glaberrima germplasm from Mali. *African Journal of Biotechnology* 9(44): 7409 – 7417.

Mario, R. H. (2007). Rice (*Oryza sativa* L.), as a source of microelements and toxic Contaminants. National Council for Science and the Environment. [http:/ /www.eoearth.org] sited visited on 31/5/2011.

Meglic, V. and Staub, J. E. (1993). Molecular genetics markers and their relevance for Cultivar discrimination: a case study in cucumber. *Horticulture Technology* 3: 291 – 300.

Michael, T. J. (1997). Conservation of rice genetic resources. The role of international rice genebank. *Plant Molecular Biology* 35: 61 – 67.

Ministry of Agriculture and Food Security (2009). *National Rice Development Strategies.* Ministry of Agriculture and Food Security, Dar es Salaam, Tanzania. 9pp.

Muhammad, S. R., Rezwan, M., Samsul, A. and Lulfur, R. (2009). DNA Fingerprinting of Rice cultivars using microsatellite markers. *Australian Journal of Crop Science* 3(3): 122 – 128.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986**).** Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction. *Cold Spring Herbarium Symposium Quantitative Biology* 51: 263 – 273.

Ntundu, W. H., Shillah, S. A., Marandu, W. Y. F. and Christiansen, J. L. (2006). Morphological diversity of bambara groundnuts (*vigna subterranean* L.) landraces in Tanzania. *Genetic Resources and Crop Evolution* 54(2): 367 – 378.

Priti, U., Vikas K, S. and Neeraja, C. N. (2011). Identification of genotype specific alleles and molecular diversity assessment of popular Rice (*oryza sativa* L.) varieties of India. *International Journal of Plant Breeding Genetics* 5(2): 130 – 140.

Qian, W. S. and Hong, D. Y. (2001). Genetic variation within and among populations of Wild Rice *Oryza granulata* from China detected by RAPD and ISSR markers. *Theory Applied Genetics* 102: 440 – 444.

Quintela-Sabaris, C., Kidd, P. S. and Fraga, M. I. (2005). Identification of Metallic Ferrous ecotypes of *Citrus ladanifer* L. using RAPD Markers. *Journal of Bioscience* 30: 229 – 235.

Ram, C. J., Kailash, C. S., Sumit, K. N. and Amrita, P. (2010). DNA finger printing ofPromising Rice (*Oryza sativa* L.) accessions from India using RAPD markers. *International Journal of Integrated Biology* 10(3): 142 – 146.

Ramedra, N. S. and Bahar, I. (2004). Genetic variation of Bora rice (glutinous rice) of Asam as revealed by RAPDs. *PGR Newsletter of Biodiversity International* 144: 34 – 38.

Roder, M. S., Korzun, V., Wendehake, K., Plashke, J., Tixier, M. H., Leroy, P. and Ganal, M. W. (1998). A micro satellite map of wheat. *Genetics*149: 2007 – 2032.

Samarajeewa, P. K., Nitanthi, R. M. and Lyanage, A. S. (2004). Morphological and Molecular characterization of locally collected wild rice germplasm. *Annals of the Sri Lanka Department of Agriculture* 6: 177 – 186.

Sasaki, T. and Burr, B. (2000). International Rice Genome Sequencing Project: Effort to completely sequence the rice genome. *Crop in plant Biology* 3(2): 138 – 141.

Schötterer, C. (2004). The evolution of molecular markers- just a matter of fashion. *Nature Review Genetics* 5: 63 – 69.

Takuji, S., Masahiro, Y., Kazuko, Y. and Rie, Y. (2005). The map-based sequence of the Rice genome. *Nature* 436: 793 – 800.

Tang, S. X., Jiang, Y. Z., Wei, X. H., Li, Z. C. and Yu, H. Y. (2002). Genetic diversity of Isozymes of cultivated rice in china. *Acta Agronomy* 28: 203 – 207.

Tarmino, G., Tarchini, R., Ferrario, S., Lee, M. and Pe, M. (1997). Characterization & Mapping of simple sequence repeats (SSRs) in Sorghum bicolor. *Theory of Applied Genetics* 95: 66 – 72.

Thaura, G. H., Duina, P. D. and Iris, P. A. (2008). Assessment of genetic diversity in Venezuelan rice cultivar using simple sequence repeats markers. *Molecular Biology and Genetics* 11(5): 1 – 12.

Umamaheswari, D., Paramathma, M. and Manivannan, N. (2010). Molecular genetic Diversity analysis in seed sources of Jatropha (*Jatropha curcas* L) using ISSR markers: *Electronic Journal of Plant Breeding* 1(3): 268 – 278.

Virk, P. S., Newbury, H. J., Jackson, M. T. and Ford, L. B. V. (1995). The identification of duplicate accessions within a rice germplasm collections using RAPD analysis. *Theory of Applied Genetics* 90: 1049 – 1055.

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

Wong, S. C., Yiu, P. H., Bong, S. T. W., Lee, H. H., Neoh, P. N. P. and Rajan, A. (2009). Analysis of Sarawak Bario Rice Diversity using microsatellite markers. *American Journal of Agriculture and Biological Sciences* 4(4): 298 – 304.

Xu, Y., Beachell H., and McCouch S. R. A. (2004). Marker-based approach to Broadening the genetic base of rice in USA. *Crop Science* 44: 1947 – 1959.

Yong-Jin, J. P., Ju, K. L. and Nam, S. K. (2009). Simple Sequence Repeat Polymorphisms for Evaluation of molecular Diversity and germplasm classification of minor crops. *Molecules* 14: 4546 – 4569.

CHAPTER THREE

3.0 Morphological diversity of rice (*Oryza sativa* L.) Landraces and its Wild Relatives (*Oryza punctata*) Conserved at the National Plant Genetic Resources Centre in Tanzania

3.1 Abstract

Morphological characteristics have been routinely used as preliminary and general approach for assessing genetic diversity among morphologically distinct crops. The knowledge of the amount, the extent, and distribution of genetic variation of the rice accessions conserved at the NPGRC is vital to the improvement and conservation strategies. NPGRC collected and conserved a good number of rice landraces, but no morphological data base that have been created. A field trial was conducted to assess the morphological diversity of seventy nine rice landraces conserved at the NPGRC. Morphological data were recorded and scored following the IBPGR and IRRI advisory committee on rice standard descriptors. The results showed that there is a significant variation among rice landraces studied. The first four principal components (PCs) with Eigen values > 1 accounted for 63.99% of the total variance among accessions studied. Cluster analysis grouped rice accessions into four distinct major groups including one group of wild rice and three groups of mixed clusters of rice accessions from different regions. The yield related traits showed a significant and positive association with the grain yield per accession. Therefore, these traits may be used for genetic potential in the improvement or breeding and selection programs of rice landraces.

3.2 Introduction

Rice (*Oryza sativa* L.) is the principal staple food for more than half of the world's population, growing over 1.5 billion hectares of the overall worldwide production. Landraces of rice played a very important role in the local food security and sustainable development of agriculture, in addition to their significance as genetic resource for rice genetic improvement (Tang *et al*., 2002). Rice is the second most important food and commercial crop in Tanzania after maize. It is among the major sources of employment, income, animal feeds, thatching materials and food security for Tanzania farming households. Tanzania is the second largest producer and consumer of rice in Eastern and Southern Africa after Madagascar with production level of 818 000 tones (USDA world rice statistics, 2007; Kafitiri *et al.,* 2003). It's estimated that about 71 % of the rice grown in Tanzania is produced under rain fed conditions. Irrigated land presents 29% of the total with most of it in small village level traditional irrigations.

Farmers grow a number of traditional varieties which have long maturity and in most cases production is affected by both biotic and abiotic stress (Kanyeka, 1994). In this case the future paddy rice production in Tanzania will depend on improved varieties and the plant breeders have to utilize the genetic diversity of rice germplasms conserved at the National Plant Genetic Resource Centre for improvement and breeding programs.

Rice landraces provide adaptability gene for specific environmental conditions, and if these genes are appropriately used in breeding programs it may ensure optimum grain yield for our poor resources farmers. National Plant Genetic Resources Centre (NPGRC), as collected from a wide range of agro ecological areas of Tanzania over 125 accessions of rice landraces for conservation and sustainable utilization. Part of these collections is included in this study. The efficient management and utilization of the genetic potential of

plant genetic resources like rice and other crops requires a detailed knowledge including characterization of genetic diversity, morphological evaluation, and classification (Karp *et al*., 1997). Nguyen *et al*. (2007) in their diversity study identified some limitations (including low polymorphism, low heritability, late expression, and vulnerability to environmental influence) in estimating total genetic variation using morphological traits, but despite of these limitations morphological traits were useful for preliminary evaluation of 22 Asian aromatic rice cultivars.

It is fast, simple and can be used as a general approach for assessing genetic diversity among morphologically distinguishable cultivars. Moreover, Muhammad *et al*. (2012) were able to identify a number of genotypes with good yield potential in Pakistani rice on the basis of plant height and grain morphological traits. The genotypes identified would be used for the improvement and development of new rice varieties through breeding program; because morphological variation alone does not reflect the total variation which is necessary for breeding. Further, comprehensive investigation including molecular markers and qualitative characters are vital as they provide a complete view about the genetic variation of the genotype in question. Many tools are available for assessing variability and relationship among accessions including seed protein, isozymes, and various types of molecular markers. However, morphological characterization is the first step in the description and classification of the germplasms (Rabbani *et al*., 1998; Smith *et al*., 1991).

Multivariate analyses are useful for characterization, evaluation, and classification of plant genetic resources when large numbers of accessions are to be assessed for several morphological characters of interest (Peeters and Martinelli, 1989). They have been used successfully to order variation observed in quantitative traits in collection of many crop

germplasms. This includes soya bean (Perry and Mackintosh, 1991), alfalfa (Smith *et al*., 1995) and pea (Amurrio *et al*., 1995) among others.

The main objective of the present study were therefore to assess morphological characteristics of rice landrace accessions in order to explore the degree of similarity and /or differences for the purpose of providing the information that could be used to develop effective conservation strategies at the National Plant Genetic Resources Centre and also utilization for improvement of crops by plant breeders.

3.3 Materials and Methods

3.3.1 Plant materials

Rice germplasm used in this study include 79 rice accessions as part of NPGRC collections from a wide range of agricultural climatic zones in Tanzania (Table1). Accessions included in this study were 41 from eastern zone, 15 from southern zone, 2 from central zone, and 10 from Lake Zone and 9 from southern highland agricultural zones in Tanzania. Seventy seven were landraces and two accessions of wild relatives that are *Oryza punctata* and *Oryza Longistigmata*. Duplicates of this seed samples are deposited at NPGRC in Arusha, Tanzania.

Table 1: Accessions and collection locations of the 79 rice landraces germplasm conserved at the National Plant Genetic Resources Centre

Accession no.	Status	Region	Altitude	Altitude status	Longitude	Latitude
TZA 1516	Landrace	Dodoma	1730m	High	3518 E	0556 S
TZA 1517	Wild	Singida	1740m	High	3511 E	553 S
TZA1519	Landrace	Dodoma	850m	Medium	3506 E	749 S
TZA 1525	Landrace	Mbeya	1900m	High	3417 E	840 S
TZA 1533	Wild	Mbeya	1910m	High	3417 E	841 S
TZA 1534	Landrace	Mbeya	1930m	High	3423 E	842 S
TZA 1535	Landrace	Mbeya	1930m	High	3423 E	842 S
TZA 1536	Landrace	Mbeya	1930m	High	3423 E	842 S
TZA1552	wild	Coast	70m	Low	3831 E	63953 S
TZA 1557	Wild	Morogoro	320m	Low	373204 E	62603 S
TZA 2256	Landrace	Lindi	-	-	-	-
TZA 2283	Landrace	Lindi	-	-	-	-
TZA 2286	Landrace	Lindi	-	-	-	-
TZA 2339	Landrace	Lindi	-	-	-	-
TZA 2341	Landrace	Lindi	-	-	-	-
TZA 2356	Landrace	Mtwara	-	-	-	-
TZA 2679	Landrace	Morogoro	373m	Low	374011 E	61000 S
TZA 2683	Landrace	Morogoro	395m	Low	373820 E	60650 S
TZA 2691	Landrace	Morogoro	512m	Medium	373438 E	60729 S
TZA 2696	Landrace	Morogoro	525M	Medium	370437 E	65657 S
TZA 2698	Landrace	Morogoro	525M	Medium	370437 E	65657 S
TZA 2699	Landrace	Morogoro	525M	Medium	370437E	65657S
TZA 2703	Landrace	Morogoro	600m	Medium	365521 E	65733 S
TZA 2723	Landrace	Morogoro	700m	Medium	365537 E	71159 S
TZA 2735	Landrace	Morogoro	350m	Low	365641 E	74912 S
TZA 2744	Landrace	Morogoro	400m	Low	305356 E	74954 S
TZA 2747	Landrace	Morogoro	412m	Low	365333 E	75013 S
TZA 2748	Landrace	Morogoro	412m	Low	365333 E	75013 S
TZA2750	Landrace	Morogoro	412m	Low	365333 E	75013 S
TZA 2751	Landrace	Morogoro	412m	Low	365333 E	75013 S
TZA 2763	Landrace	Morogoro	412m	Low	364952 E	80108 S
TZA 2764	Landrace	Morogoro	412m	Low	364952 E	80108 S
TZA 2765	Landrace	Morogoro	412m	Low	364952 E	80108 S
TZA 2766	Landrace	Morogoro	412m	Low	364952 E	80108 S
TZA 2771	Landrace	Morogoro	412m	Low	364952 E	80108 S
TZA 2772	Landrace	Morogoro	375m	Low	361448 E	81208 S
TZA 2776	Landrace	Morogoro	362m	Low	361448 E	81208 S
TZA 2779	Landrace	Morogoro	375m	Low	361448 E	81208 S
TZA 2780	Landrace	Morogoro	375m	Low	361440 E	81224S
TZA 2782	Landrace	Morogoro	387m	Low	363034 E	80555 S
TZA 2784	Landrace	Morogoro	387m	Low	363034 E	80555 S
TZA 2794	Landrace	Morogoro	487m	Low	364440 E	85426 S
TZA 2796	Landrace	Morogoro	487m	Low	364440 E	85426 S
TZA 2805	Landrace	Tanga	187m	Low	364333 E	85149 S
TZA 2810	Landrace	Tanga	187m	Low	364333 E	85149 S
TZA 2820	Landrace	Tanga	187m	Low	385908 E	52410 S
TZA 2821	Landrace	Tanga	187m	Low	385908 E	52410 S
TZA 2823	Landrace	Tanga	187m	Low	385908 E	52410 S
TZA 2826	Landrace	Tanga	187m	Low	385908 E	52408 S
TZA 2837	Landrace	Tanga	425m	Low	383708 E	51254 S
TZA 2844	Landrace	Tanga	425m	Low	383708 E	51254 S
TZA 2845	Landrace	Tanga	425m	Low	383708 E	51254 S
TZA 2846	Landrace	Tanga	425m	Low	383708 E	51254 S
TZA 2852	Landrace	Tanga	412m	Low	38275 E	51016 S
TZA 2866	Landrace	Ruvuma	969m	High	352723 E	103613 S
TZA 2906	Landrace	Ruvuma	566m	Medium	352004 E	112036 S
TZA 2943	Landrace	Ruvuma	585 m	Medium	345024E	111840 S
TZA 3028	Landrace	Iringa	-	-	344438 E	102327 S
TZA 3572	Landrace	Ruvuma	978 m	High	374204 E	70001 S
TZA 3642	Landrace	Mtwara	-	-	401013 E	103941 S
TZA 3691	Landrace	Mtwara	-	-	-	-
TZA 3726	Landrace	Mtwara	215m	Low	394288 E	105460 S
TZA 3822	Landrace	Mtwara	25m	Low	392190 E	110300 S
TZA 3831	Landrace	Mtwara	660m	Medium	391334 E	103104 S
TZA 3846	Landrace	Lindi	-	-	383835 E	100317 S
TZA 3850	Landrace	Lindi	-	-	383844 E	1003195 S
TZA 3895	Landrace	Lindi	230m	Low	385753 E	1003195 S
TZA 4007	Landrace	Mwanza	-	-	330143 E	020528 S
TZA 4015	Landrace	Mwanza	1120m	High	330158 E	020542 S
TZA 4034	Landrace	Mwanza	1100m	High	325240 E	015614 S
TZA 4O38	Landrace	Mwanza	1000m	High	325234 E	015605 S
TZA 4161	Landrace	Kagera	1100m	High	310813 E	030626 S
TZA 4180	Landrace	Kagera	1125m	High	314642 E	025923 S
TZA 4184	Landrace	Kagera	1110m	High	314647 E	025921 S
TZA 4210	Landrace	Mwanza	1100m	High	325229 E	023822 S
TZA 4431	Landrace	Lindi	-	-	393558 E	094726 S
TZA 4617	Landrace	Mwanza	1100m	High	325456 E	025647 S

3.3.2 Experimental site

The Laboratory work was done at National Plant Genetic Resources Centre (NPGRC) laboratory based at Tropical Pesticides Research Institute (TPRI), Arusha. TPRI is located about 12km north east of Arusha – township, at latitude 3 ^{o}19’52’’ south and longitude 36^{o}37’27’’east at 1470 meters above sea level. The field evaluation was carried out at Makumira Lutheran Evangelical Farm Usa-River located at latitude 3^{o}45’19’’south and longitude 36^{o}49’59’’east at 1189 m above sea level. There are two cropping seasons in this area that mainly depends on irrigation. The first season is from mid January to late June whereas second season is from mid August to December. This experiment was conducted between January and June, 2012 cropping season.

3.3.3 Field and crop management

The rice germplasm lines used in this study consisted of 79 rice accessions that were collected from National Plant Genetic Resources Centre (NPGRC) situated at TPRI, Arusha. The experimental evaluation was conducted under irrigation system during January to June, 2012 growing season at Makumira farm, Arusha. The seeds of all 79 rice accessions were pre-germinated in aseptic condition on the wet filter papers for five days in the laboratory and later transferred into pots in the greenhouse. Seedlings of each accession were transplanted at the age 21 - day old in a randomized block of 50cm, and each accession was assigned to a single row spaced at 20cm apart with 5cm between plants. Ten seedlings per accessions were transplanted and one seedling per hill, due to shortage of seedlings there were no replication in this experiment. The germination of genebank seeds were very low, so only few seeds sprout, the reason may be due to long keeping of these seeds as most them were collected before 2002. Fertilizer was applied three times; farm yard manure was applied as a basal fertilizer just before transplanting, 10kg urea was applied during tillers formation and another 10kg was applied during

panicle initiations stage. Weeding was carried out using hands, and an Endosulphan (35% EC-Emulsifiable Concentrates) insecticide at the rate of 70ml per 20L was sprayed to control insect pests.

3.4 Morphological Traits and Data Collection

Plant morphological data including vegetative, grains and yield characters were recorded following the International Board for Plant Genetic Resources (IBPGR) and IRRI (IBPGR, 1980) rice advisory committee on rice standard descriptors. Fourteen morphological characters were scored in this study. Days to heading was recorded at flowering stage when flowers were fully open on five randomly selected plants per each accession. Plant height (cm); number of productive tillers; culm length (cm); culm number; flag leaf length (cm); flag leaf width (cm) were recorded on five plant selected randomly within each accession just after full heading. Number of days to maturity was recorded at 80% maturity that was indicated by browning of spikelet; 100 seed weight (g); grain yield per accessions (g); grain length (mm); grain width (mm) and grains length-width ratio were recorded after harvesting. All data were recoded on a collection form.

3.5 Statistical Analysis

The quantitative morphological traits were subjected to Principal Component Analysis Procedure of GenStat computer programme version 4; using correlation matrix to define the pattern of traits variation between accessions and those principle components with Eigen values greater than 1.0 were selected. To group rice accessions in such a way that patterns of similarity and dissimilarity would be revealed, the trait data matrix was subjected to cluster analysis using simple average linkage method and the computations were performed using the statistical programme “R” (The Foundation for Statistical

Computing, 2004). On feeding in the observed morphological data, the program undergoes three major steps of a hierarchical analysis, which is, creating a distance or dissimilarity matrix, clustering or agglomeration, and plotting. The program produces a cluster dendogram which represents the relationships among the accessions under study in terms of approximate distances based on morphological traits.

Cluster analysis allows one to identify groups of variables that are similar among them, especially when they are closely related. The phenotypic distance matrix was created by calculating the distance between each pair of accession for each trait. The distance between two accessions for quantitative traits was scored as zero if their phenotypes matched and as one if they did not. The distance index for quantitative traits was determined by averaging all the distance in the phenotypic value for each quantitative trait divided by the respective range (Ortiz *et al*., 1998). The phenotypic distances for these traits were transformed/normalized into a 0-1 scale. Thus, the phenotypic distance between two accessions was calculated as the sum of by summing the individual traits distances between them, and dividing by the total number of traits recorded in both accessions (Ortiz *et al*., 1998).

3.6 Results

3.6.1: Descriptive statistics for 14 quantitative traits

Genetic variation was displayed for the traits evaluated as showing in table 2, by standard error of mean, percentage coefficient of variation, standard deviation and maximum and minimum values.

Table 2: Basic statistics for 14 quantitative traits of landraces and wild relatives

Traits	Mean ±SE	Maximum	Minimum	% CV	STDV
Culm length	88.98± 2.140	160	54	21.38	19.02
Culm number	4.443 ± 0.0715	6	4	14.30	0.635
Flag leaf length	68.27 ± 1.284	111	24	16.72	11.41
Flag leaf width	1.672 ± 0.0437	3	1	23.21	0.388
Grains L/W ration	4.476 ± 0.0459	6	3	9.120	0.408
Grains width	36.29 ± 2.302	2	1	56.38	20.46
Grain length	9.440 ± 0.0855	11	6	8.046	0.760
100 grain weighty	2.956 ± 0.0553	4	1	16.63	0.492
Days to heading	124.8 ± 1.876	159	74	13.36	16.68
Days to maturity	168.4 ± 1.840	183	120	9.712	16.35
Productive tillers	12.99 ±0.497	30	6	34.03	4.421
Panicle length	24.39 ± 0.239	28	17	8.724	2.128
Plant height	114 ± 2.130	180	73	16.54	18.93
Yield per accession	195 ± 15	601	19	18.63	13.6

SE - Standard errors of a mean, % CV – percentage coefficient of variation, STDV - Standard deviation

Grain width (8.046%), days to heading (9.12%), and panicle length (8.724%) were the traits that had the coefficient of variance less than 10%. The rest or majority of the traits had coefficient of variance above 10% and the highest 56.38% was observed for the grain width. Plant height had a wide range of 180cm – 73cm with mean height of 114cm and 16.54 coefficient of variance. Maximum plant height was observed in accession TZA 1517 a wild rice collected from Singida Region and the dwarf variety among the accessions studied was TZA 3028 having the plant height of 75cm collected from Iringa Region. Flag leaf length and its width is very important growth characters in which maximum photosynthesis is accrued. Maximum flag leaf of 96cm was recorded in accessions TZA 2744 and minimum leaf length of 24cm was observed in TZA 2339. Standard deviation was 11.41 for flag leaf length with 16.72% coefficient of variance and 68.27 mean values. Days to heading showed the highest range (74 – 159 day) and coefficient of variation 13.36%. TZA 2356 had minimum days of heading (74 days) among rice accessions studied. Days to maturity had a range of 120 – 183 days with a mean value of 168 days and coefficient of variance was 9.712%. Maximum days of maturity (183) was observed in several rice accessions studied, TZA 2784, TZA 3028,

TZA 2794, TZA 2691, TZA 2866, TZA 3822, TZA 3581, TZA 2852, TZA 1516, TZA 2744 TZA 4210, and TZA 2763 while the minimum (120days) was recorded in TZA 1517 and TZA 1557 wild rice varieties collected from Singida and Morogoro Regions respectively.

Maximum panicle length of 28cm was observed in TZA 1557, TZA 2794, TZA 2820 and TZA 2744 whereas minimum of 17.2 cm in TZA 2748. Maximum yield per accession of 601grams was observed in TZA 3846, while the minimum value (19 g) of TZA 1517. The coefficient of variability of grain length was 8.046% and standard deviation value of 0.760 was observed in this trait which revealed that low amount diversity in grain length. The maximum value was 11mm and minimum 5.9mm observed in TZA 4007 and TZA 1517 respectively. The grain width/ length ratio ranged from 6.091 to 3.281mm with mean value of 4.476, standard deviation of 0.408 and coefficient of variance 9.12%, which showed low amount of variability in this character. The maximum value of grain width / length ratio of 6.091mm was observed in TZA 3581, while minimum value of 3.28mm was observed in TZA 2796.

3.6.2 Correlation studies

Correlation matrix for the 14 quantitative rice traits studied is presented in Table 3 below. Some of the traits exhibited positive correlations, while other showed negative association with one another. The characters such as flag leaf width (0.511), grain length (0.360), and 100 seed weight (0.319) showed significant positive association with the grain yield per accessions, whereas Culm number (-0.227) and grain length/width ratio (-0.254) showed negative correlation with grain yield per accessions. Culm length showed highest correlation (r = 0.9311) with the plant height followed by the panicle length (r = 0.502) and Culm number (r = 0.312). Hundred seed weights was significantly positively

correlated with grain length (r = 0.804), and grain width (r = 0.553) indicating a tendency that long and wider grain were heavier. Number of days to maturity (r = 0.658) showed positive correlation with number of days to heading. Based on correlation analysis, a number of traits were directly associated with other traits regardless of plant type or architectural configuration of the whole plant. Correlation matrix helps to determine pairs of characters that vary in the same or opposite direction and its useful guide; especially for plant breeders who may want to associate a set of traits in their breeding programs. Correlation between pairs of characters can be considered reliably significant, when the absolute values of the coefficient are greater than 0.20 (Fowler *et al*., 1990).

Table 3: Correlation matrix between 14 quantitative traits measured in rice accessions

CUL	CUN	FLL	FLW	GLW	GRL	GRW	GWT	GYD	NDH	NDM	NPT	PAL	PHT
CUL	**1.000**												
CUN	**0.310**	**1.000**											
FLL	**0.024**	**0.109**	**1.000**										
FLW	**0.223**	**0.056**	**0.140**	**1.000**									
GLW	**-0.236**	**0.059**	**0.237**	**-0.004**	**1.000**								
GRL	**0.000**	**-0.258**	**-0.110**	**0.123**	**-0.570**	**1.000**							
GRW	**-0.267**	**-0.168**	**0.180**	**0.139**	**0.498**	**0.406**	**1.000**						
GWT	**0.014**	**-0.312**	**-0.080**	**0.139**	**-0.238**	**0.804**	**0.553**	**1.000**					
GYD	**0.185**	**-0.227**	**0.044**	**0.511**	**-0.254**	**0.360**	**0.081**	**0.319**	**1.000**				
NDH	**-0.523**	**0.149**	**0.272**	**0.004**	**0.443**	**-0.226**	**0.286**	**-0.174**	**-0.163**	**1.000**			
NDM	**-0.463**	**0.171**	**0.312**	**-0.131**	**0.334**	**-0.016**	**0.377**	**0.018**	**-0.310**	**0.658**	**1.000**		
NPT	**0.160**	**-0.035**	**0.083**	**-0.177**	**-0.279**	**0.078**	**-0.194**	**-0.029**	**0.031**	**-0.285**	**-0.262**	**1.000**	
PAL	**0.433**	**0.162**	**0.202**	**0.049**	**-0.129**	**-0.057**	**-0.179**	**-0.086**	**0.169**	**-0.136**	**-0.136**	**0.127**	**1.000**
PHT	**0.931**	**0.312**	**0.072**	**0.253**	**-0.165**	**-0.098**	**-0.278**	**-0.117**	**0.198**	**-0.406**	**-0.437**	**0.127**	**0.502**	**1.000**

CUL = Culm length, CUN =Culm number, FLL =Flag leaf Length, FLW= Flag leaf width, GLW=grain length/ width ratio, GRL=grain length, GRW = grain width, GWT =100 seed weight, GYD= grain yield per accessions, NDH= number of days to heading, NDM= number of days to maturity, NPT= number of productive tillers, PAL= panicle length, PHT= plant height

3.6.2 Principal component analysis

Principal Components Analysis (PCA) is designed to reduce the number of variables to a smaller number of indices that show linear combinations of the original variables called principal components (Manly, 1994). Principal component analysis showing the contribution (factor scores) of each character among 79 rice accessions, Eigen values and percentage total variance accounted for by four principal components are presented in Table 4 below.

Table 4: Principal component analysis

Traits	Prin.1	Prin.2	Prin.3	Prin.4
Culm length (cm)	-0.49348	0.19639	-0.09229	0.02018
Culm number	-0.09090	0.46221	0.14225	0.24736
Flag leaf length(cm)	0.06678	0.40373	-0.01246	-0.50432
Flag leaf width (cm)	-0.12028	0.15975	-0.67832	-0.02098
Grain length/width ratio	0.26464	0.31723	-0.05536	0.08564
Grain width (mm)	0.05216	-0.10460	0.12240	-0.5309
Grain length (mm)	-0.013	-0.454	-0.254	0.198
100 grain weight (g)	-0.01890	-0.33919	-0.4625	-0.39536
Grain yield/accession (gm)	-0.21094	-0.28128	0.26800	-0.25160
Days to heading (d)	0.40064	-0.28128	-0.01231	- 0.06379
Days to maturity (d)	0.39692	0.26466	-0.02033	-0.14030
Number of Productive Tillers	-0.19274	-0.13552	0.5052	-0.22672
Panicle Length (cm)	-0.27690	0.27270	0.13438	-0.40270
Plant Height (cm)	-0.47359	0.27757	-0.07128	0.02705
Eigen value	**3.230**	**2.008**	**1.280**	**1.160**
% variance	**26.92**	**16.73**	**10.67**	**9.67**
Cumulative % variance	**26.92**	**43.66**	**54.33**	**63.99**

The four PCs with Eigen values > 1 accounted for 63.99 % of the morphological variability for the 14 quantitative traits of the rice germplasm tested. PC1 accounted 26.92% of the total morphological variation for traits evaluated. Morphological traits that loaded highly for PC1 were days to heading, days to maturity and grain length/ width ratio (Table 4). PC2 accounted for 16.73% of the total morphological variation among the accessions studied. Culm number, flag leaf length, grain length width ratio, and days to

maturity, panicle length, and plant height loaded highly in PC2 for the accessions studied. Similarly, PC3 contributed to 10.67% of the total morphological variation and the characters that contributed to this variation for PC3 was the number of the productive tillers. In addition, PC4 accounted for 9.67% of the total variability among accessions for the characters accessed in this study. Culm number contributed highly to PC4 (Table 4). PCA is a technique which identifies plant traits that contributes to most of the observed variation within a group of genotypes. The tool has a practical application in the selection of parent lines for breeding purposes (Manly, 1994).

The overall observations for the PCs analysis for the 14 quantitative traits included in this study indicated that most of the morphological variations for the quantitative characters studied were contributed by PC1 and PC2. The observed high loading for both PC1 and PC2 were mainly reproductive characters (days to heading and days to maturity), grain characters (grain length width ratio), culm number, and flag leaf length which represent vegetative characters, implying that these traits may be important in distinguishing the materials under study. Principal components were then subjected to the scatter plots and most accessions concentrated at the centre except for the wild accessions TZA 1557 and TZA 1517(Fig. 1).

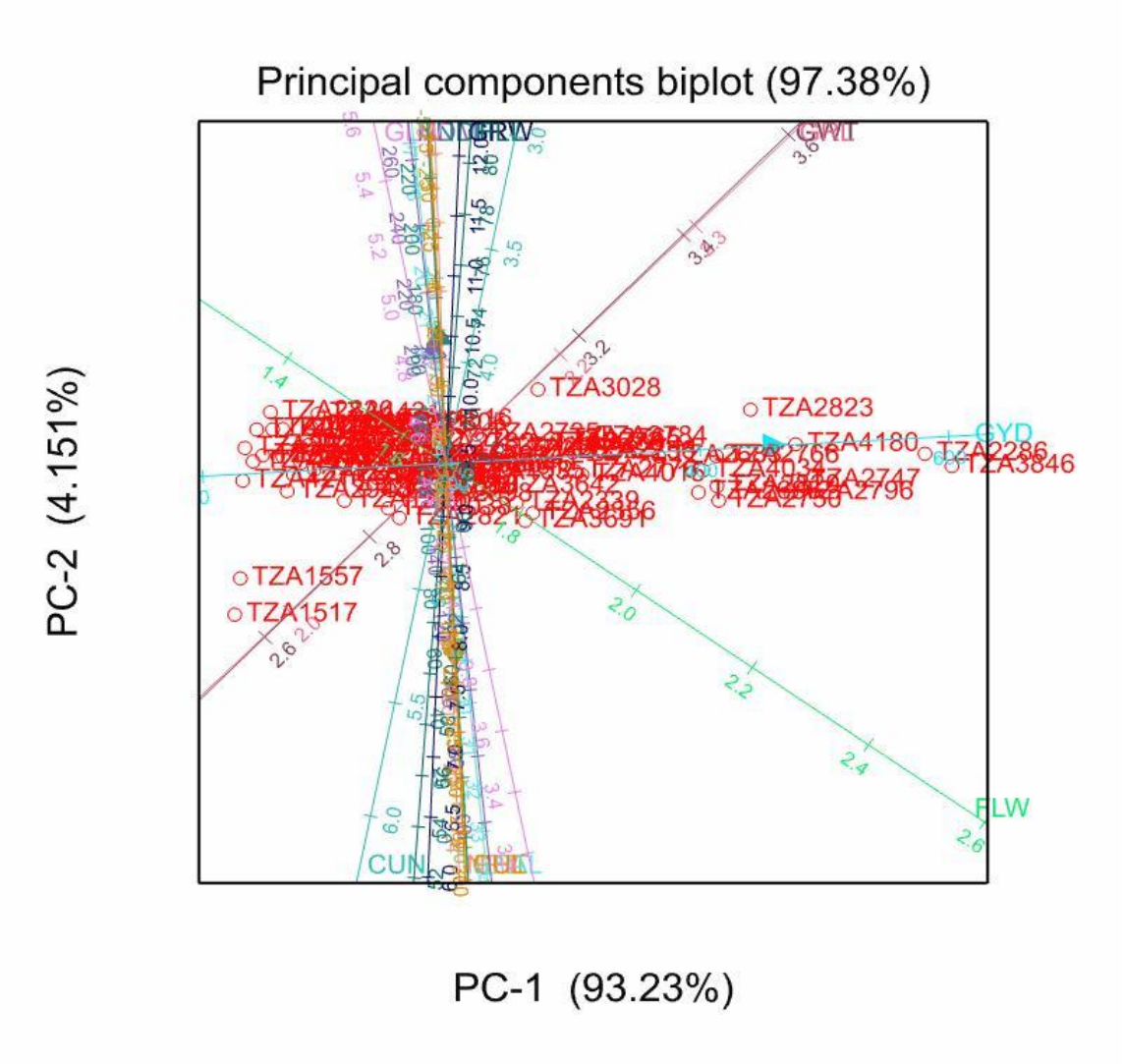

Figure 1: Scatter plot for two principal components based on 14 quantitative traits of rice accessions studied

3.6.3 Cluster analysis

Figure 2 below showed a dendrogram for average linkage cluster analysis for the 79 rice accessions evaluated based on the 14 quantitative morphological traits. The cluster analysis showed that 79 accessions of both rice landraces and wild relatives grouped in clusters with their respective distances.

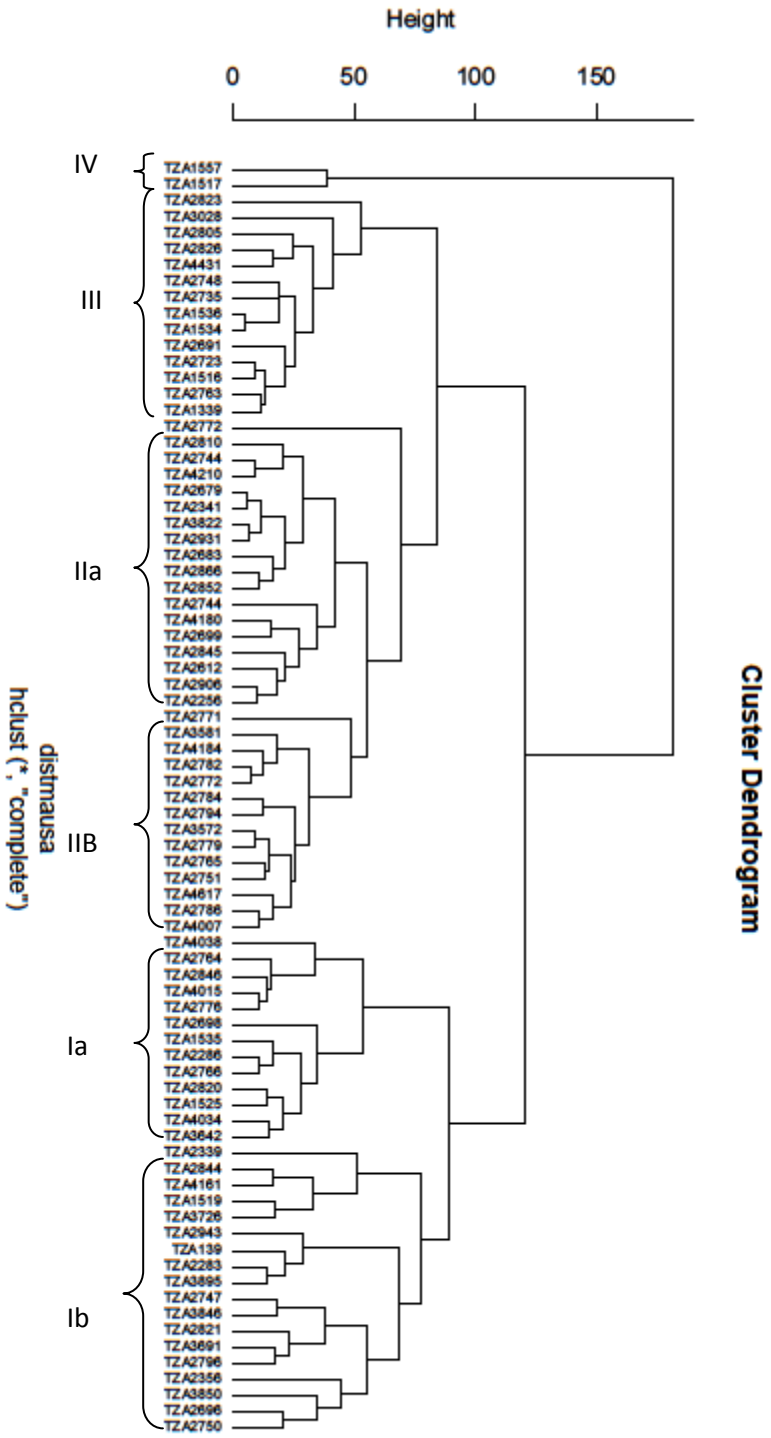

Figure 2: Dendrogram showing cluster analyses based on average linkage for the 14morphological relatedness of 79 rice accessions investigated

The phenotypic distance index based on the morphological traits ranged from 4 (between accessions TZA 1536 and TZA 1534 all from Mbeya which is in southern highlands agricultural ecological zone) to 60 for accession TZA 2772 from Morogoro in eastern agricultural ecological zone. Accessions of wild relatives formed a clear separate group with 2 accessions (TZA 1557 and TZA 1517) from Eastern Morogoro and Singida in central agricultural ecological zones. This grouping pattern indicates that these accessions are distinct from the rest of the materials studied. The cultivars in this group were characterized by their taller plants, longest panicle, longer culm, lower grain yield, shorter grain with low grain length/width ratio, early flowering and maturing. The first, second and third grouping included several clusters with 77 mixed accessions of rice landraces mainly collected from the Central, Lake Victoria, Eastern, Southern highlands, and Southern agricultural zones. The mixed accessions in these clusters may indicate that they constitute a heterogeneous group of accessions with different origins.

3.7 Discussion

A reasonable variation was displayed for the traits evaluated. Grain length/width ratio, grain width, and days to maturity, was the only traits with less coefficient of variance. The majority traits displayed higher coefficient of variance and the highest was recorded for the grain width. Accession TZA 4038 had the highest plant height, whereas TZA 3024 is a short variety having shorter plant height. Sadia *et al*. (2012) had also observed relatively greater range of plant height than the other characters in their studies. Plant height is of a paramount important as the reduction of it may develop their resistance to lodging and reduce substantial yield losses associated with this trait (Abbasi *et al*., 1995). Accession TZA 2356 collected from Mtwara was the one which had the minimum value for days to heading. The minimum days of heading were also reported by Sadia *et al*., (2012), when evaluated 68 primitive cultivars and commercial varieties under field

condition in Pakistan where they found a rice variety Ranbir-basmati with 51 days to heading. Maximum panicle length was observed in accessions TZA 1557, TZA 2794, TZA 2820, and TZA 2744 and minimum in TZA 2748. Although it contributes highly for higher yielding but it's not the only traits responsible, productive tillers per plant are another yield attributing traits (Abbasi *et al*., 1995). The highest coefficient of variability was observed in productive tillers per accessions in this study. Sadia *et al*. (2012) also observed and reported the higher coefficient of variability for number of productive tillers in Pakistan rice. Maximum grain length and width was observed in TZA 2845 and TZA 3895 respectively.

The correlation among the most reproductive characters demonstrates the association that exists between these characters. Rice grain yield improvement is the major character of economic interest. Therefore, its correlation with other character is most desirable. The yield related traits such as grain size as indicated by grain length and 100 seed weight demonstrated positive association with grain yield per accessions. Therefore, these traits may be useful to be included in the selection programs to improve rice yield. Days to heading were positively associated with days to maturity. Sadia *et al*. (2012) also reported positive correlation between days to heading and days to maturity among Pakistan rice. Plant height had positive association with panicle length showing the importance of plant height in improving panicle length in rice.

This result were in agreement with Zafar *et al*. (2004) and Sadia *et al*. (2012) who observed that plant height had highly significant and positive association with panicle length. Also panicle length and flag leaf length were correlated positively. These results are in agreement with the finding of Abdus *et al*. (2009) and Ramakrishna *et al*. (2006). Number of productive tillers per accession correlated positively with grain yield per

accessions. This implies that tillers produced the grains that are in good proportionate ratio and there was no high frequency of barren tillers.

Multivariate methods have been used successfully to classify and measure the patterns of morphological diversity in the relationship of species and germplasm collections of variety of crops (Perry and McIntosh, 1991; Smith *et al.,* 1995). In this study most of the morphological variation for rice accessions was accounted by the fist four principal components. The main quantitative characters which accounted for more variability in both PC 1 and PC 2 includes days to heading, days to maturity, grain length width ratio, culm number, flag leaf length and respectively. This suggests that accessions with high PC1 and PC2 value were yielding and vegetative traits, which are characterized by days to heading, days to maturity, grain size, culm number and flag leaf length . Hence these could be considered important traits in the rice accessions studied.

The cluster analysis for quantitative characters included in this study placed rice cultivars into four clusters with sub clusters for each, except cluster four which has only two accessions of wild rice. Clusters were grouped according to their morphological differences among them. This result is in agreement with Sadia *et al*. (2012) who also reported that in cluster analysis cultivars grouped together with greater morphological similarities, but cluster did not essentially include all the cultivars collected from the same origin. Cluster analysis also showed a distinct clear separate group of accessions of wild rice (*O. punctata*) from Morogoro and Tanga. On the other hand, the first, second and third cluster together with their sub clusters grouping included mixed accessions of 77 rice landraces collected from different regions. This indicates that they consisted of the heterogeneous group of accessions with different origin. The reason could be a frequent exchange of seed materials practiced by farmers over this region resulting in a mixture of

landraces cultivated by farmers. The exchange of seed materials is not unique to farmers as reported by other researchers like Ashimogo and Rukulantile (2000) sited by (Wazael *et al*., 2004) that 35.4% of farmers in Dodoma, Iringa, and Morogoro Regions obtain maize (*Zea mays* L.) seeds from their fellow farmers, while only 60.1% grow their own seeds. Also the wide variation in temperature, rainfall, topography, and soil in the country could be another reason that caused heterogeneous grouping of the accessions seen.

3.8 Conclusion and Recommendations

Morphological traits were useful for preliminary evaluation and can be used as a general approach for assessing genetic diversity among morphologically distinct rice accessions. Based on the present study, rice germplasm conserved at the National Plant Genetic Resources Centre displayed a considerable range of morphological diversity for the most of the morphological traits studied. Several potential economical important traits such as grain size indicated by 100 seed weight, grain length and grain width demonstrated a significant and positive association with grain yield per accessions, days to heading and days to maturity.

Despite the limitation of morphological traits (including low polymorphism, low heritability, late expression, and vulnerability to environmental influences) in estimating total genetic variation, the present study indicated that they can be used to assess genetic diversity of rice accessions. These traits of economic importance may be used for genetic potential in the improvement programs of rice landraces.

Cluster analysis has proved to be an effective method in grouping rice accessions from different parts of Tanzania that may facilitate the conservation, management, and utilization of plant genetic resources by selecting accessions with good economic traits.

3.9 References

Abbasi, F. M., Sagar, M. A., Akram, M. and Ashraf, M. (1995). Agronomic and quality traits of some elite rice genotypes. *Pakistan Journal of Science Indian Research* 38: 348 – 350.

Abdus, S. K., Muhammad, I. and Muhammad, A. (2009). Estimation of genetic Variability and correlation for grain yield components in rice (*oryza sativa* L.). *American Eurasian Journal of Agriculture and Environment Science* 6(5): 585 – 590.

Amurrio, J. O., de Ron, A. M. and Zeven, A. C. (1995). Numerical taxonomy of Liberian pea landraces based on Qualitative and quantitative characters. *Euphytica* 82: 1995 – 205.

Ashimogo, G. and Rukulantile, H. (2000). *Danish International Development Agency Agricultural Sector Programme Surport. Consultancy on Seed Market Assessment in Dodoma, Iringa and Morogoro Regions.* Faculty of Agriculture, Sokoine University of Agriculture, Morogoro, Tanzania, 37pp.

Fowler, J. and Cohen, L. (1990). *Practical Statistics for Field Biology*. John Wiley and Sons Ltd., West Sussex, England. 141pp.

Kafiriti, E. M., Dondeyne, S., Msomba, S., Deckers, J. A. and Raes, D. (2003). Variations in agronomic characteristics of irrigated rice varieties: Lessons from participatory trials in South Eastern Tanzania. *Food, Agriculture and Environment* 1: 273 – 277.

Kanyeka, Z. L., Msomba, S. W., Kihupi, A. N. and Penza, M. F. (1994). Rice Ecosystems in Tanzania: Characterization and classification. *Research and Training Newsletter* 9: 13 - 15.

Karp, A. and Edwards, K. J. (1997). Techniques for the analysis, characterization, and Evaluation of plant genetic resources. Molecular techniques in the analysis of extent of genetic diversity. (Edited by Ayad, W. G., Jaranda, A. and. Rao, V. R.), *Report of International Plant Genetic Resource Centre Workshop,* 9 - 11 October 1995, Rome, Italy. pp. 39 – 43.

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

Manyl, B. F. J. (1994). *Multivariate Statistic Methods a Primer*. (Second Edition) Champman and Hall, London. 108pp.

Muhammad, S. R., Rezwan, M., Samsul, A. and Lulfur, R. (2009). DNA Fingerprinting of Rice cultivars using microsatellite markers. *Australian Journal of Crop Science* 3(3): 122 – 128.

Nguyen, L. H., Wakil, A. S., Yosei, O. and Yutaka, H. (2007). Genetic diversity of Morphological and the relationships among Asia aromatic rice (*oryza sativa* L.) Cultivars. *Journal of* Crop *Science* 31: 36 – 47.

Ortiz, R., Madsen, S. and Vuylsteke, D. (1998). Classification of African Plantain Landraces and banana cultivars using a phenotypic distances index of quantitative descriptors. *Crop Science* 35: 1245 – 1248.

Peeters, J.P., and Martinelli, J. A. (1989). Hierarchical cluster analysis as a tool to Manage variation in germplasm collections. *Theoretical Applied Genetics* 78: 42 – 48.

Perry, M. C. and Maclntosh, M. S. (1991). Plant genetic resources. Geographical patterns of variation in the USDA soyabean Germplasm Collection: Morphological traits. *Crop Science* 31: 1350 – 1355.

Rabbani, M. A., Iwabuchi, A., Murakami, Y., Suzuki, T. and Takanayagi, K. (1998).Phenotypic variation and relationships among mustard (*Brassica juncea* L.). Germplasm from Pakistani. *Euphytica* 101: 357 – 366.

Ramakrishnan, S. H., Anandakumar, C. R., Sarvanan, S. and Malin, N. (2006). Association analysis of some yield traits in rice (*oryza sativa* L.). *Journal of Applied Science Research* 2(7): 402 – 404.

Sadia, T., Zahida, H. P., Yasin, M. M., Ashiq, M. R. and Shahid, M. M. (2012). Assessment of phenotypic variability in rice (*oryza sativa* L.) cultivars using multivariate analysis. *Pakistan Journal of Botany* 44(3): 999 – 1006.

Smith, S. A., Al-Dos, A. L. and Warburton, M. (1991). Morphological and agronomical variation in North African and Arabian Alfalfas. *Crop Science* 31: 1159 – 1163.

Smith, S. E., Guarino, L., Al-Dos, A. L. and Conta, D. M. (1995). Morphological and agronomic affinities among Middle Eastern alfalfa accessions from Oman and Yemen. *Crop Science* 35: 1188 – 1194.

Tang, S. X., Jiang, Y. Z., Wei, X. H., Li, Z. C., and Yu, H. Y. (2002). Genetic diversity of cultivated rice in china. *Acta Agronomy* 28: 203 – 207.

Wazael, H. N., Inga, C. B., Jorgen, L. C. and Sven, B. A. (2004). Analysis of genetic diversity in bambara groundnuts [*vigna subterranean* (L)] landraces using amplified fragment length polymorphism markers. *African Journal of Biotechnology* 3(4): 220 – 225.

Zafar, N., Aziz, S. and Masood, S. (2004). Phenotypic divergence for agro-Morphological traits among landraces genotypes of rice (O*ryza sativa* L.) from Pakistan. *International Journal of Agriculture Biology* 2: 335 – 339.

CHAPTER FOUR

4.0 Genetic diversity of rice landraces (*Oryza sativa* L.) and its Wild (*Oryza punctata*) Relative as Revealed by Simple Sequence Repeats DNA Markers

4.1 Abstract

Simple sequence repeat markers have been used in genetic diversity studies for different number of crops including rice landraces. SSR markers have higher reproducibility which is most important in genetic analysis, is co-dominant, bands produced from the same set of primers are intuitively orthologus, and are abundant in genome of all species. Genetic diversity of 79 rice landraces and its wild relatives conserved at the NPGRC in Tanzania were assessed using five SSR markers. A total of 11 polymorphic alleles were detected, number of alleles per locus ranged from 2 to 3 with an average of 2.2 per locus. The dendrogram grouped accessions into four clusters with sub clusters. A wild variety (*Oryza punctata)* collected from Singida Region which is in central agricultural zone formed a clear separate cluster. The Jaccard's similarity coefficient revealed that most rice accessions studied were genetically identical, except TZA 2696, TZA 2339, TZA 4210, TZA 3822, TZA 3850, TZA 2805, and TZA 1552 which were ascertain to be genetically different from other accessions. Therefore, advocated to be included in the rice improvement programs. This study also highlighted use of large number of molecular markers (> 5) for efficient characterization of the rice accessions conserved at the NPGRC than the ones that had been used in this present study which shows that most of the rice landraces studied were genetically identical at genomic level.

4.2 Introduction

Rice is the second most important food and commercial crop in Tanzania after maize in respect of local consumption as well as export. It's among the major sources of employment, income and food security for Tanzania farming households (Tusekelege *et al.*, 2011). Rice in Tanzania is grown by peasant farmers under varied ecological conditions in all regions of the country but at varied levels of importance. It is grown in swampy areas and river basins such as the Rufiji, Ruvu in Coast Region, Kilombero, and Wami, in Morogoro Region Pangani and Mombo in Tanga Region. It is also cultivated in Shinyanga, Mwanza, and Kigoma Region among other places (Monyo and Kanyeka, 1978). It is estimated that, about 60% of Tanzanian population consumes rice and its derivatives per day each year (Kanyeka *et al.*, 1994). Rice also plays a great role in human being by which, stalks are utilized as animal feed and thatching materials. Farmers through selection for many generations have developed multiplicity of rice varieties, each with particular traits valued by the communities that developed them, but genetic erosion is threatening plant genetic resources in Tanzania (Ching'ang'a, 1990) due to the rapid destruction of the earth's natural habitats which is now occurring at the alarming rate.

Therefore, there is a strong need to, collect, characterize, conserve not only the landrace genotype but also the wild relatives for future utilization in breeding of high yielding, superior quality, and better-adapted varieties in Tanzania. Germplasm characterization and evaluation will help to determine their trueness to type also their utilization potential in breeding and other development programmes (Moore, 1997). A part from using morphological and agronomical traits to characterize and evaluate rice germplasm, use of molecular or DNA marker techniques in determining diversity, will help to reveal the available genetic variability and identify population for conservation and sustainable utilization purposes. Molecular markers can reveal abundant difference among genotypes

at the DNA level providing a more direct, reliable, and efficient tool for germplasm characterization, conservation, management, and are not influenced by the environmental stress factors (Prabakaran *et al*., 2010). Molecular markers have been widely used for genetic diversity studies in a number of crop species, including the estimation of genetic diversity of rice landraces and its wild relative (Kiambi *et al*., 2005).

Among various PCR based markers, SSR markers are more popular in rice because they are highly informative, mostly monolocus, co - dominant, easily analyzed, cost effective (Gracia *et al*., 2004), require low amount of DNA, can easily be automated and allow high-throughput screening, can be exchanged between laboratories and are highly transferable between populations (Malik *et al*., 2010). SSR markers have been extensively used in the characterization of different crops including sorghum (Tulole *et al*., 2009); characterization and similarity relationship among apricot (Hormaza, 2002); evaluation of genetic diversity in wheat germplasm (Manifesto e*t al*., 2001); a simple sequence repeat based linkage map of Barley (Ramsay *et al*., 2000).

The success of these molecular markers in characterizing and analyzing genetic diversity in rice has been demonstrated by a number of studies. Priti *et al*. (2011) identified genotype specific alleles and molecular diversity assessment of popular rice (*Oryza sativa* L.) varieties of India using SSR marker. Ramendra and Bahar (2011), conducted genetic diversity studies to differentiate "Bora" rice (glutinous rice) using the same marker. Claudio *et al*. (2006) determined genetic variability of traditional varieties of Brazilian rice using SSR marker. Virk *et al*. (1995) assessed morphological and molecular characterization of locally collected wild rice germplasm using SSR marker. Samarajeewa *et al*. (2004) carry out experiment to identify duplicate accessions within rice germplasm collections using SSR molecular marker. Bakari (2010), conducted genetic diversity

study of some of rice landraces grown in Tanzania using simple sequences repeat markers. In his study, he reported a narrow genetic diversity of rice landraces and concluded that no significance genetic difference among rice landraces he studied. SSR markers are class of repetitive DNA sequence usually 2.6 bp that are distributed throughout whole genome and are flanked by highly conserved region (Chambers and Avoy, 2000). The current study aims to analyze genetic diversity and relationships between and within rice landraces and its wild relatives conserved at the National Plant Genetic Resources Centre using simple sequence repeat markers.

4.3 Materials and Methods

4.3.1 Plant materials

Young leaves from five plants of each rice accession were harvested from a screen house at the NPGRC-TPRI. Seventy percent ethanol was used as disinfectant for hand washing before the commencement of harvesting process. Plastic pestles were sterilized in an autoclave and then washed with 70% ethanol before a new set of leaves was extracted.

4.3.2 DNA extraction

The genomic DNA was extracted from the fresh young leaves following the protocol described by Ahmadikhah (2009) with a slight modification (modification by addition of 1 unit of 4M Ammonium Acetate before incubation of supernatant at -20 ^{o}C). Fresh leaf samples of 14 days-old seedling were used as the source of genomic DNA. Leaf tissue were placed in a 1.5 ml eppendorf tube and ground in 100µl DNA extraction buffer (1% CTAB, 700mM NaCl, 10mM Tris- HCl pH 8, and 50 mM EDTA pH 8) using hand operated homogenizer with a plastic pestle. After initial homogenization, 350µl of DNA extraction buffer (pre-warmed up to 65^{o}C and addition of 38mg/ml sodium bi-sulphite just before use) were added and vortexed for 60 seconds. After incubation for 45 minutes

at 65^{o}C with intermittent swirling for cell lyses, the mixture was emulsified with 0.7units of Chloroform: Isoamyl alcohol at 24:1 (v/v), vortexed for 5 minutes and then centrifuged at 14 000g for 5 minutes at 4^{o}C. About 0.7 units of cold isopropanal alcohol and 1 unit of 4M Ammonium Acetate were added to the supernatant and incubated in the freezer at -20 ^{o}C overnight. The supernatant was centrifuged at 14 000g for 5 minutes at 4^{o}C to pellet the DNA. The pellets were then washed with 70% ethanol, and allowed to air dry on the paper towel for one hour. To each tube 50µl of TE buffer (10mM Tris-HCl, 50mM EDTA, pH=8) were added and stored at 4^{o}C to dissolve the DNA pellets. To check the quality of DNA, 2µl of solution was loaded and run on a 3% agarose gel in a mini-gel at 4V/cm for 3 hours. The remaining DNA extracts solution was stored at -20 ^{o}C.

4.3.3 Microsatellite markers and PCR amplification

A set of 12 microsatellite primer pairs previously developed by Kalyan and Rambabu, (2006) distributed in the rice genome were selected. The identified primers were then screened against 79 rice accessions with a recommended PCR thermal profile. Based on better responsiveness in amplifying the target genomic region of DNA template, five primer pairs RM346, RM333, RM 519, RM 267, and RM 345 representing chromosome numbers 5, 6, 7, 10, and 12 of rice genome were selected and used for microsatellite analysis in this present study (Table 5). Polymerase Chain Reaction were carried out in a volume of 20µl containing 1µl of DNA templates from each accessions, 4µl of 1 X PCR buffer containing (50Mmol/l KCl, 10Mmol/l Tris-HCl, 1.5 Mmol/l $MgCl_2$), 1µl of each forward and reverse primer; 0.18 units of *Taq* DNA polymerase; 0.4µl of 200 µmol/l of each dNTPs (Qiagen GmbH, Germany) and required amount of sterile deionized water.

Amplification were carried out in oil-free thermal cycler (Whatman Biometra professional Basic Thermo cycler-GmbH, Germany) with the following program: initial

hot start and strand separation at 94^{o}C for 2 minute followed by 35 cycles of 30 second at 94^{o}C for template denaturation, 45 second primer annealing, the annealing temperature ranged from 55.3 ^{o}C to 60.3 ^{o}C depending on the primer used and 1minute at 72^{o}C for primer elongation, and finally 1 cycle of 7minute at 72^{o}C for final extension. Amplified products were stored at -20^{o}C until further use. The SSR primers used in this study and their corresponding annealing temperature is as shown in the table 5 below.

Table 5: List of SSR primer pairs used in the present study

locus	Forward primer	Tm (^{o}C)	Reverse primer	Tm (^{o}C)	Chromosome numbers	Base pair (cM)
RM346	cgagagagcccataactacg	59.8	acaagacgacgaggagggac	61.4	7	140-175
RM333	gtacgactacgagtgtcaccaa	60.3	gtcttcgcgatcactcgc	58.2	10	164-215
RM519	agagagcccctaaatttccg	57.3	aggtacgctcacctgtggac	61.4	12	119-138
RM267	tgcagacatagagaaggaagtg	58.4	agcaacagcacaacttgatg	55.3	5	137-160
RM345	attggtagctcaatgcaagc	55.3	gtgcaacaaccccacatg	56.0	6	152-167

4.4 Electrophoresis of Amplified Products

After amplification, a 10µl aliquot of the amplified SSR samples was mixed with 3µl of a loading buffer (0.4% (w/v) bromo-phenol blue, 0.4% (w/v) xylene cyanole and 5ml of glycerol) and electrophoresed at 4V/cm for 3 hours on 2.5% (w/v) agarose gel in 0.5xTBE buffer (10mM Tris-Borate, 1mM EDTA) containing 0.5 µg per ml of ethidium bromide. A 1000 bp DNA ladder (Qiagen GmbH, Germany) was used as a standard size marker to compare the molecular weights of amplified products. After electrophoresis, the gels were documented using an UVI Doc Gel Documentation System (UVITEC, Cambridge, UK).

4.5 Genetic Diversity Data Analysis

All genotypes were scored for the presence of the SSR bands. The data were entered into a binary matrix as discrete variable, 1 representing present band and 0 representing absence of a band. Thereafter, this data matrix was subjected to further analysis. The 0/1

matrix was used to calculate similarity matrix as Jaccard's coefficients. The resultant similarity matrix was used to construct dendrogram using UPGMA of R computer software to infer genetic relationships and phylogeny. Also Principal Components Analysis (PCA) based on clustering was done using GenStat version 4 Computer software.

4.6 Results

4.6.1 SSR analysis

Five informative primers were employed in this study which generated a total of 242 reproducible and scorable fragments across the rice accessions, out of which 11 were polymorphic. The number of amplified SSR bands per primer per sample ranged from 47 to 50 with an average of 48.4 bands per primer. The average number of polymorphic bands detected was 2.2 per primer. Plate 1 shows a DNA bands amplified from rice accessions included in this study using SSR marker RM 333.

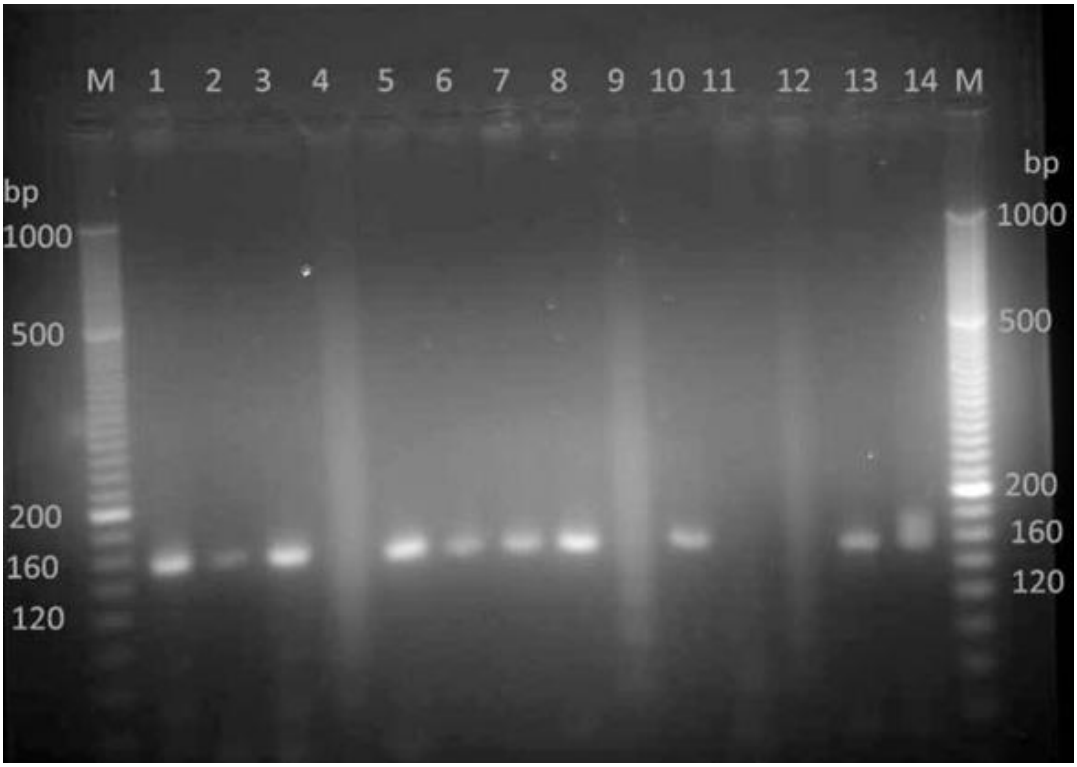

Plate 1: Amplification of DNA from accessions no. 1-14 using SSR marker RM 333

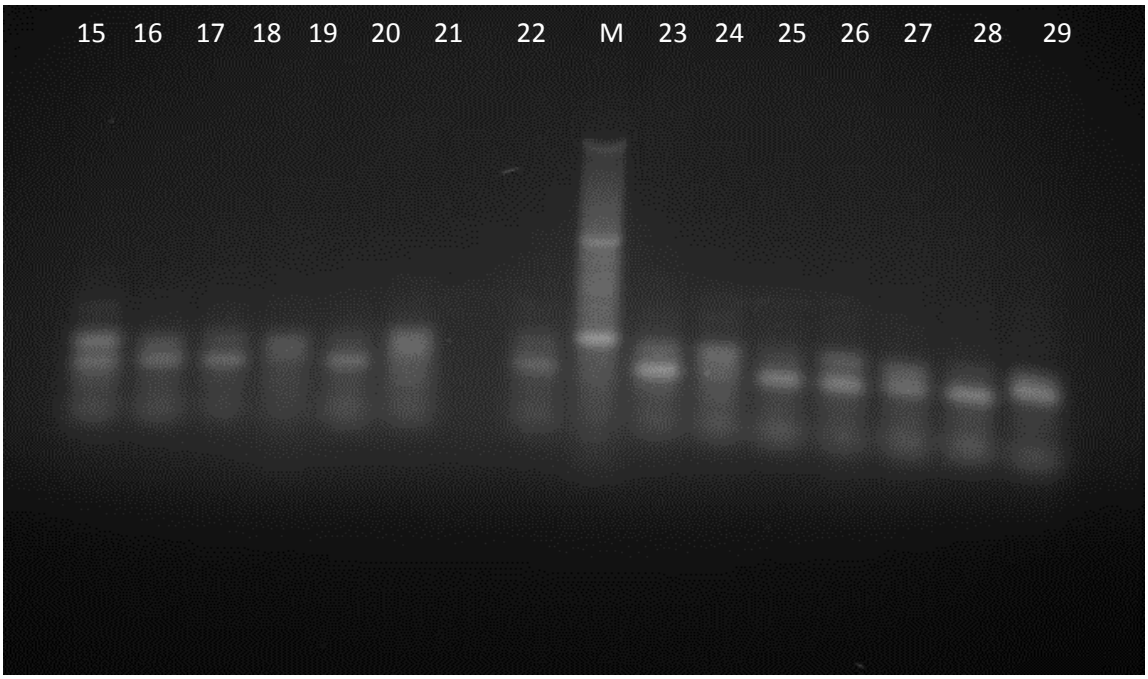

Plate 2: Amplification of DNA accession no.15 – 29 using SSR marker RM 333

4.6.2 Cluster analysis

The genetic distance for the 79 rice accessions based on the SSR markers varied from 0.0 to 1.5. The dendrogram based on the Jaccard's coefficient of genetic distance suggested the existence of four clusters with sub clusters. Cluster four has no sub cluster, all four clusters has 22 groups of accessions as summarized in Table 6 below.

Table 6: Distribution of genotype to different clusters based on the UPGM methods

Cluster	Number of genotype	Name of accessions
Ia	21	TZA1339, TZA2723, TZA2283, TZA1536, TZA1534, TZA2766, TZA2750, TZA2691, TZA2341, TZA1535, TZA1525, TZA4038, TZA2794, TZA2796, TZA2820, TZA2821, TZA2943, TZA3028, TZA3642, TZA4007, and TZA4034
Ib	4	TZA 2696, TZA 2782, TZA 2256, and TZA 2286
IIa	8	TZA2780, TZA2683, TZA2779, TZA 3846, TZA 2844, TZA 2810, TZA 2735, and TZA 2771.
IIb	12	TZA4180, TZA2784, TZA2852, TZA4015, TZA2748, TZA2837, TZA 2339, TZA4161, TZA3691, TZA 2845, TZA 2356, and TZA 2747
IIIa	5	TZA 4184, TZA 3850, TZA 3822, TZA 2751 and TZA 3895
IIIb	12	TZA 4617,TZA 3831,TZA 2846,TZA1516, TZA2699, TZA4431,TZA2776,TZA 2679, TZA 2763,TZA 4210, TZA 2744 and TZA 3572
IV	8	TZA2765, TZA 2866, TZA1533,TZA1519, TZA2823, TZA 2772, and TZA 1552

Cluster IV had 8 accessions including TZA 1552 a wild rice collected from Cost region with a genetic distance of 1.5, others were, TZA 1533, and TZA 2772. Cluster two had five accessions TZA 4184, TZA 3850, TZA 3822, TZA 2751, and TZA 3895. Moreover cluster three had twelve accessions, Cluster four had eight accessions, Cluster five had twelve accessions, Cluster six was the largest with 21accessions of rice landraces studied, and Cluster seven an out group had only four accessions collected from Morogoro and Lindi respectively. Microsatellite analysis indicated that, most of the rice landraces accessions studied were genetically related based on their similarity coefficient as showing in Fig. 4 below.

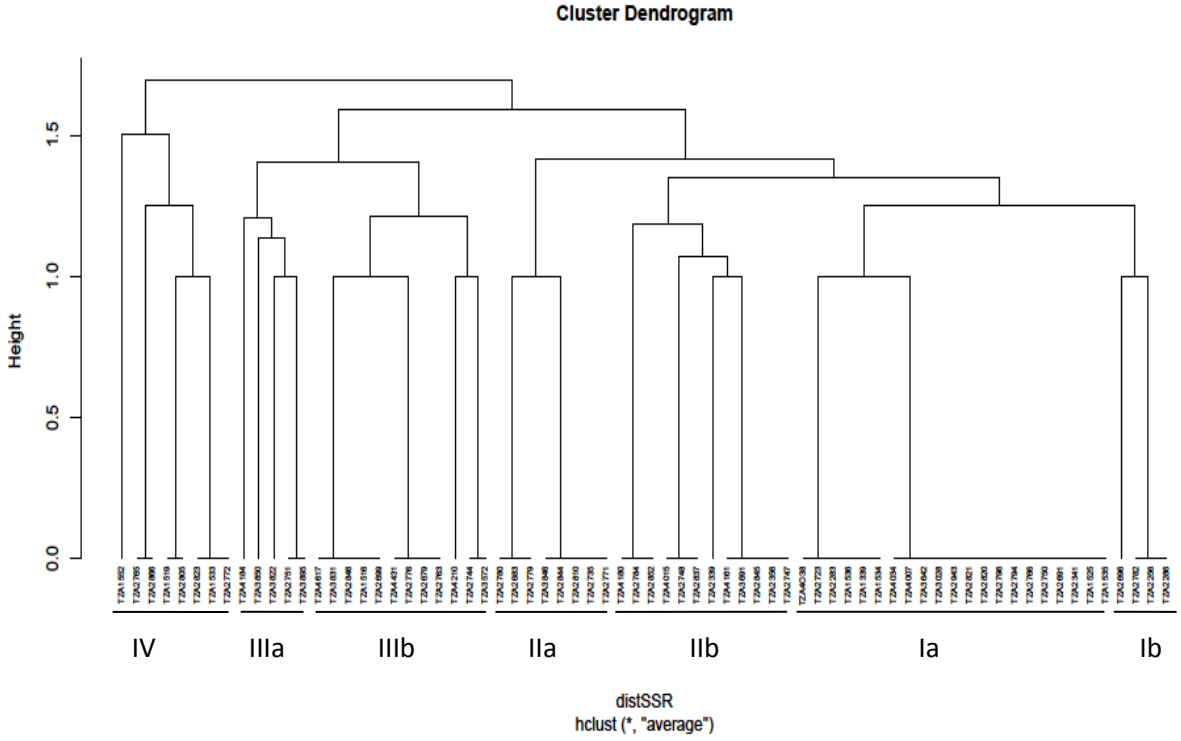

Figure 3: Dendrogram showing the clustering of 79 population of rice germplasm using 5 SSR primers

Key: The prefix TZA before the numbers is a designation for Tanzania accessions code.

4.7 Discussion

A total of 5 RM primers (Table 4) were utilized to provide genetic diversity among 79 accessions of rice landraces studied and a total of 242 bands were scored. All five SSR markers generated polymorphic patterns. A total of 11 alleles were detected among the genotypes. The number of allele per locus ranged from 2 (RM 346, RM 519, RM 267, RM 345) and 3 (RM 333) with an average of 2.2 per locus. Same results were also reported by Prabakaran *et al.* (2010) when investigated genetic variability of 12 rice landraces of India using five primers. However, this value is quite low compared with those reported in other places of the collections. In the report by Rahman *et al*. (2010), the number of alleles per locus varied from 3 to 8, with an average number of alleles per locus at 4.86 and Claudio *et al*. (2006) reported a range of 6-22 numbers of alleles per primer with an average of 14.7 alleles per locus. This indicates a less magnitude of diversity with reference to the fewer markers among the plant materials in this present study.

The genetic distance based on SSR data ranged from 0 to 1.5, with a total average distance of 0.75. This narrow range of genetic distance is supported by the results from the study done by Bakari (2010) on characterization of rice landraces collected from various places in Tanzania using SSR markers. The author reported some rice landraces with narrow genetic diversity that ranged from 0.76 to 0.98 distant coefficients. In Korea, Song *et al*. (2002) reported narrow genetic background of Korean rice germplasm as revealed by DNA fingerprinting with SSR markers and their pedigree information. In Japan a report by Hashimoto *et al*. (2004) showed that sake-brewing rice had low genetic diversity as compared to the cooking rice cultivars. The results from our study shows that no significant genetic differences among rice landraces conserved at the

National Plant Genetic Resources Centre as the coefficient of similarities was concentrated at a range of 1.0 to 1.5 genetic distances for the most of varieties studied.

Results from the dendrogram generated based on UPGMA from similarities or a genetic distance matrix has shown an overall pattern of variation as well as the degree of relatedness among accession of rice landraces. The dendrogram produced by cluster analysis in this study resulted into four clusters with sub clusters of 22 groups of rice accessions where all sample had a genetic distance ranged from 0.0 to 1.5 coefficient of similarity. Accessions from Morogoro, Tanga, Mbeya, Iringa, and Mwanza form a clear group (cluster 6) together with few accessions from Mtwara and Lindi. These accessions showed higher genetic similarity coefficient at level 1.0 indicating that they belong to similar genetic background. The reason, regarding this similarity cases were probably due to the selections made by farmers from a single landraces (Choudhury *et al*., 2001). According to Chakravarthi and Naravaneni (2006) high similarity coefficient indicates cultivars belonging to similar genetic background that may be caused by occasional out-crossing events that may occur spontaneously or selectively between cultivars. Also the cultural habits of exchanging seeds among farm families, villages, or regions sometime do contribute to the enhancing of genetic diversity within the regions by adding different genotypes to the local rice gene pool. This could be in part, due to good transportation systems in many places in Tanzania, which allow easy communication among farmers. The dendrogram reveals that the genotype that are genetically similar cluster more together. The independent grouping of accessions TZA 2696, TZA 2339, TZA 4210, TZA 3822, TZA 3850, TZA 2805, and TZA 1552 indicated that they are genetically different from other accessions. These divergent accessions may have many good agronomically important traits including pests and diseases resistance, tolerance to biotic and abiotic stress. Therefore, further evaluation of these accessions may be vital to

explore their unique characteristics from other materials included in this study, which can be utilized for direct selection and as parents of crosses with other accessions from different clusters. This results show that most of the rice landraces material kept at the National Plant Genetic Resources Centre comprises of duplicate accessions as out of 79 rice accessions studied only seven accessions were distinct. Principal Component Analysis (Fig.4) was also done to visualize genetic relationships among the genotypes studied; the results were similar to that of the cluster analysis having 22 distinct groups of rice accessions.

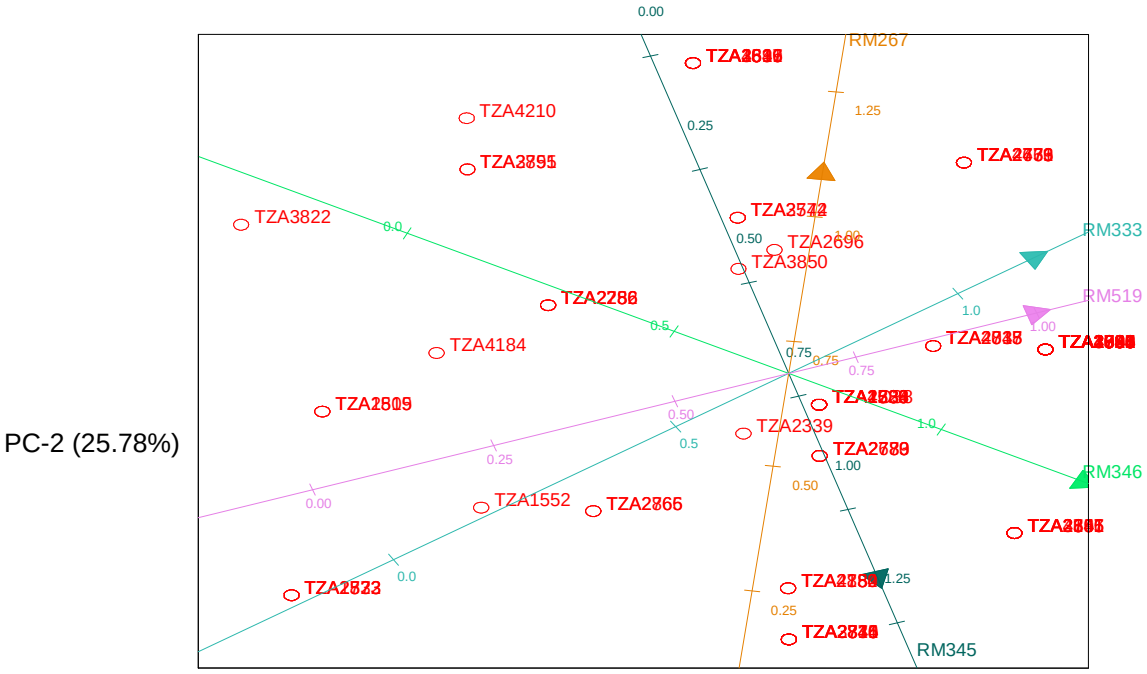

Figure 4: Principal Component Analysis scores grouping of 79 rice accessions based SSR markers (Principal components biplot 53.24%)

4.8 Conclusions

The overall results obtained by SSR analysis of rice accession in the present study shows that the rice accessions conserved at the National Plant Genetic Resource Centre are not a genetically diverse population as the genetic coefficient of most of the rice accessions studied concentrated at 0.0 to 1.0 genetic distances, and only six accessions been at 1.0 to 1.5 genetic distance levels. The accessions were grouped under six cluster similarity levels and one outer cluster of wild rice.

The first cluster had seven accessions with genetic distance of 1.5 coefficients of which TZA 2765, TZA 2866 were genetically identical with genetic distance of 1.25. The accessions TZA 1519, TZA 2805, TZA 2823, TZA 1533 and TZA 2772 showed high relationship with genetic distance of 1.0. Cluster two had five accessions of which two were genetically identical (TZA 2751 and TZA 3895) with genetic distance coefficient of 1.0, and three were distinct accessions. Cluster three had twelve genetically identical accessions with genetic coefficient of 1.0. Also cluster four had eight genetically similar accessions. Cluster five had twelve accessions, 3 with genetic distance of 1.25, two with genetic distance of 1.17 and six with genetic distance of 1.0. Cluster six had twenty five accessions with genetic distance of 1.0 coefficient. Based on this study, accessions TZA 2696, TZA 2339, TZA 4210, TZA 3850, TZA 2805 and TZA 1552 were genetically different from other accessions studied.

Therefore, may be selected as parents for hybridization with any of the landraces in other clusters. Most of the rice accessions studied was found to be genetically related with genetic distances of 1.0 coefficient thus indicating that genetically these accessions do not differ in genetic distance significantly. This result implies that rice landraces accessions conserved at the National Plant Genetic Resources Centre comprises of duplicates.

Therefore, they must be sorted out and the NPGRC Curator is necessitated to plan for core collections by identifying the hotspot areas where diversity of rice landraces can be collected, for the purpose of conservation and sustainably utilized by researcher and breeders for rice improvement programs.

4.9 Recommendations

SSR markers can effectively be used to assess the genetic diversity, identify duplicate accessions, and also to measure the extent of genetic relationship among accessions conserved at National Plant Genetic Resources Centre. But, the use of more number of DNA markers (>5) would be efficient to characterize the rice landraces conserved than the ones used in this study, which showed that most of the rice landraces were genetically identical accessions at genomic level among the genotypes studied. The knowledge of the degree of genetic relationship between rice accessions revealed by this work will be of importance to breeders in the crop improvement programs, and to genebank Curator to establish core collections as part of the germplasm collection management. Also may guide genebank manager to target germplasm collection areas, and this will be cost - effective where financial support is minimal.

4.10 References

Bakari, M. M. (2010). Genetic diversity of some rice (*Oryza sativa* L.) landraces grown In Tanzania using simple sequence markers (SSR). Dissertation for Award of MSc Degree at Degree at Sokoine University of Agriculture, Morogoro, Tanzania, 72pp.

Chakravarthi, B. K. and Naravaneni, R. (2006). SSR markers based DNA fingerprinting and diversity study in rice (*Oryza sativa* L.). *African Journal of Biotechnology* 5: 684 – 688.

Chambers, M and Avoy, M. (2000). Comparison of microsatellites and amplified Fragment length polymorphism markers for parentage analysis. *Molecular Ecology* 9: 1037 – 1038.

Ching'ang'a, H. M. (1990). The role of the national genebank in the conservation of Plant genetic resources in Tanzania. In: *Proceedings of the First Workshop on Plant Genetic Resources and Biotechnology.* (Edited by Shao, F. M. *et al.*), 16 – 20 January 1990, Arusha, Tanzania. pp. 30 – 41.

Choudhry, P. R., Kohli, S., Srinivasan, K., Mohapatra, T. and Sharma, R. P. (2001). Identification and classification of aromatic rice based on DNA fingerprinting. *Euphytica* 118: 241 – 251.

Claudio, B., Tereza, C. B., Paulo, H. R. and Rosana, P. B. (2006). Determination of Genetic Variability of traditional varieties of Brazilian rice using microsatellite markers. *Genetics and Molecular Biology* 29(4): 676 – 684.

Gracia, A. A. F., Benchimol, L. L., Antonica, M. M., Geraldi, I. O. and Deuza, A. P. (2004). Comparison of RAPD, RFLP, AFLP, and SSR markers for diversity studies in tropical maize inbred lines. *Euphytica*, 108: 53 – 63.

Hashimoto, Z., Mori, N., Kawamura, M., Ishii, T., Yoshinda, S., Ikegami, M., Takumi, S. and Nakamura, C. (2004). Genetic diversity and phylogeny of Japanese Sake-brewing rice as revealed by AFLP, Nuclear, and Chloroplast SSR markers. *Theoretical and applied Genetics* 109(8): 1586 – 1596.

Hormaza, J. I (2002). Molecular characterization and similarity relationship among Apricot (*Prunusarmeniaca*L. genotypes using simple sequence repeats. *Theoretical and Applied Genetics* 104: 321 – 328.

Kalyan, B. C. and Rambabu, N. (2006). SSR based DNA finger printing and diversity Study in rice (Oryza sativa L.). *African Journal of Biotechnology* 5(9): 684 – 688.

Kanyake, Z. L., Msomba, S. W., Kihupi, A. N. and Penza, M. F. (1994). Rice Ecosystems in Tanzania: Characterization and classification. *Research and Training Newsletter* 9: 13 – 15.

Malik, A. R., Muhammad, S. M., Zabta, K. S., and Kazuko, Y. S. (2010). Genetic Analysis of Basmati and non basmati Pakistani Rice (*Oryza sativa* L.) cultivars using microsatellite markers. *Pakistan Journal of Botany* 42(4): 2551 – 2564.

Manifesto, M. M., Schlaller, A. R., Hopp, H. E. and Dubcovsky, J. (2002). Quantitative Evaluation of genetic diversity in Wheat germplasm using molecular marker. *Crop science* 41(3): 682 – 690.

Monyo, J. H., and Kanyeka, Z. L. (1978). Rice Improvement in Tanzania. In: *Rice in Africa (Edited by Buddenhagen, I.W. and Persely, G. J.*), Academic Press Inc., London. 345 – 346.

Moore, J. M. (1987). Fruit germplasm preservation and management. *Chronica Horticulturae* 27(1): 243 – 270.

Prakaran, A., Paramasivam, K., Rajesh, T. and Dajarajan, D. (2010). Molecular Characterization of rice landraces using SSR markers. *Electronic Journal of Plant breeding*, 1(4): 512 – 516.

Priti, U., Vikas K, S. and Neeraja, C. N. (2011). Identification of genotype specific alleles and Molecular diversity assessment of popular Rice (*Oryza sativa* L.) varieties of India. *International Journal of plant breeding and Genetics* 5: 130 – 140.

Rahman, M. S., Sohag, M. K. H. and Rahman, L. (2010). Microsatellite based DNA Fingerprinting of 28 local rice (*Oryza Sativa* L.) varieties of Bangladesh. *Journal of Bangladesh Agriculture* 8(1): 7 – 17.

Ramedra, N. S. and Bahar, I. (2004). Genetic variation of Bora rice (glutinous rice) of Asamas revealed by RAPDs. *PGR Newsletter of Biodiversity international* 144: 34 – 38.

Ramsay, l., Macaulay, M., Fuller, J., Massari, A., Ganal, M., Powell, W. and Cadle, L. (2000). A Simple sequence Repeat – Based linkage map of Barley. *Genetics Society of America* 156(4): 1997 – 2005.

Samarajeewa, P. K; Nitanthi, R. M; and Lyanage, A. S. (2004). Morphological and Molecular characterization of locally collected wild rice germplasm. *Annals of the Sri Lanka Department of Agriculture* 6: 177 – 186.

Song, M. T., Lee, J. H., Cho, Y. S., Lee, S. B., Ku, J. H., Choi, S. H. and Hwang, H. G. (2002). Narrow genetic background of Korean rice germplasm as revealed by DNA fingerprinting with SSR markers, and their pedigree information. *Korean Journal of genetics* 24: 397 – 403.

Tulole, L. B., Gwanama, C., Mgonja, M., Folkertsona, R. and Mutegi, R. (2009). Genetic Variability Characterization of Tanzania Sorghum Landraces based on simple sequence repeats (SSRs) molecular and morphological markers. *African Crop Science Journal* 17 (2): 71 – 86.

Tusekelege, H. and Rweyemamu, C. (2011). Improving rice profitability through Increased productivity and better marketing focusing on Tanzania's central corridor. *Rural livelihood Development Company* 1 – 375.

Virk, P. S., Newbury, H. J., Jackson, M. T. and Ford, L. B. V. (1995). The identification of duplicate accessions within a rice germplasm collections using RAPD analysis. *Theory Applied Genetics* 90: 1049 – 1055.

CHAPTER FIVE

5.1 Extended Discussion

A reasonable variation was displayed for the morphological traits evaluated. Grain length/width ratio, grain width, and days to maturity, was the only traits with less coefficient of variance. The majority traits displayed higher coefficient of variance and the highest was recorded for the grain width. Accession TZA 4038 had the highest plant height, whereas TZA 3024 is a short variety having shorter plant height. Plant height is of a paramount important as the reduction of it may develop their resistance to lodging and reduce substantial yield losses associated with this trait. Accession TZA 2356 collected from Mtwara was the one which had the minimum value for days to heading. Maximum panicle length was observed in accessions TZA 1557, TZA 2794, TZA 2820, and TZA 2744 and minimum in TZA 2748. Although it contributes highly for higher yielding but it's not the only traits responsible, productive tillers per plant are another yield attributing traits. The highest coefficient of variability was observed in productive tillers per accessions in this study. Maximum grain length and width was observed in TZA 2845 and TZA 3895 respectively.

The correlation among the most reproductive characters demonstrates the association that exists between these characters. Rice grain yield improvement is the major character of economic interest. Therefore, its correlation with other character is most desirable. The yield related traits such as grain size as indicated by grain length and 100 seed weight demonstrated positive association with grain yield per accessions. Therefore, these traits may be useful to be included in the selection programs to improve rice yield. Days to heading were positively associated with days to maturity. Plant height had positive association with panicle length showing the importance of plant height in improving

panicle length in rice. Also panicle length and flag leaf length were correlated positively. Number of productive tillers per accession correlated positively with grain yield per accessions. This implies that tillers produced the grains that are in good proportionate ratio and there was no high frequency of barren tillers.

Multivariate methods have been used successfully to classify and measure the patterns of morphological diversity in the relationship of species and germplasm collections of variety of crops. In this study most of the morphological variation for rice accessions was accounted by the fist four principal components. The main quantitative characters which accounted for more variability in both PC 1 and PC 2 includes days to heading, days to maturity, grain length width ratio, culm number, flag leaf length and respectively. This suggests that accessions with high PC1 and PC2 value were yielding and vegetative traits, which are characterized by days to heading, days to maturity, grain size, culm number and flag leaf length . Hence these could be considered important traits in the rice accessions studied.

The cluster analysis for quantitative characters included in this study placed rice cultivars into four clusters with sub clusters for each, except cluster four which has only two accessions of wild rice. Clusters were grouped according to their morphological differences among them, but cluster did not essentially include all the cultivars collected from the same origin. Cluster analysis also showed a distinct clear separate group of accessions of wild rice (*O. punctata*) from Morogoro and Tanga. On the other hand, the first, second and third cluster together with their sub clusters grouping included mixed accessions of 77 rice landraces collected from different regions. This indicates that they consisted of the heterogeneous group of accessions with different origin. The reason could be a frequent exchange of seed materials practiced by farmers over this region

resulting in a mixture of landraces cultivated by farmers. Also the wide variation in temperature, rainfall, topography, and soil in the country could be another reason that caused heterogeneous grouping of the accessions seen.

Based on the number of SSR primer used in this study a total of 242 bands from 79 rice landraces were scored. All five SSR markers generated polymorphic patterns. A total of 11 alleles were detected among 79 rice genotypes. The number of allele per locus ranged from 2 (RM 346, RM 519, RM 267, RM 345) and 3 (RM 333) with an average of 2.2 per locus. However, this value is quite low compared with those reported in other places of the collections. This indicates a less magnitude of diversity with reference to the fewer number (5) of SSR markers among the plant materials in this present study.

The genetic distance based on SSR data ranged from 0 to 1.5, with a total average distance of 0.75. This narrow range of genetic distance is supported by the results from the study done in 2010, on characterization of rice landraces collected from various places in Tanzania using SSR markers. The author reported some rice landraces with narrow genetic diversity that ranged from 0.76 to 0.98 distant coefficients. The results from this study shows that no significant genetic differences among rice landraces conserved at the National Plant Genetic Resources Centre as the coefficient of similarities was concentrated at a range of 1.0 to 1.5 genetic distances for the most of varieties studied. Results from the dendrogram generated based on UPGMA from similarities or a genetic distance matrix has shown an overall pattern of variation as well as the degree of relatedness among accession of rice landraces.

The dendrogram produced by cluster analysis in this study resulted into four clusters with sub clusters of 22 groups of rice accessions where all sample had a genetic distance

ranged from 0.0 to 1.5 coefficient of similarity. Accessions from Morogoro, Tanga, Mbeya, Iringa, and Mwanza form a clear group together with few accessions from Mtwara and Lindi. These accessions showed higher genetic similarity coefficient at level 1.0 indicating that they belong to similar genetic background. The reason, regarding this similarity cases were probably due to the selections made by farmers from a single landraces and also the cultural habits of exchanging seeds among farm families, villages, or regions sometime do contribute to the enhancing of genetic diversity within the regions by adding different genotypes to the local rice gene pool. This could be in part, due to good transportation systems in many places in Tanzania, which allow easy communication among farmers.

The dendrogram reveals that the genotype that are genetically similar cluster more together. The independent grouping of accessions TZA 2696, TZA 2339, TZA 4210, TZA 3822, TZA 3850, TZA 2805, and TZA 1552 indicated that they are genetically different from other accessions. These divergent accessions may have many good agronomically important traits including pests and diseases resistance, tolerance to biotic and abiotic stress. Therefore, further evaluation of these accessions may be vital to explore their unique characteristics from other materials included in this study, which can be utilized for direct selection and as parents of crosses with other accessions from different clusters. This results show that most of the rice landraces material kept at the National Plant Genetic Resources Centre comprises of duplicate accessions as out of 79 rice accessions studied only seven accessions were distinct.

5.2 Extended Conclusions

Based on the present study, rice germplasm conserved at the National Plant Genetic Resources Centre displayed a considerable range of morphological diversity for the most

of the morphological traits studied. Despite the limitation of morphological traits (including low polymorphism, low heritability, late expression, and vulnerability to environmental influences) in estimating total genetic variation, the present study indicated that they can be used to assess genetic diversity of rice accessions. Cluster analysis has proved to be an effective method in grouping rice accessions from different parts of Tanzania that may facilitate the conservation, management, and utilization of plant genetic resources by selecting accessions with good economic traits. The overall results obtained by SSR analysis of rice accession in the present study shows that the rice accessions conserved at the National Plant Genetic Resource Centre are not a genetically diverse population as the genetic coefficient of most of the rice accessions studied concentrated at 0.0 to 1.0 genetic distances, and only six accessions been at 1.0 to 1.5 genetic distance levels. This result implies that rice landraces conserved at the National Plant Genetic Resources Centre comprises of duplicates accessions.

5.3 Extended Recommendations

Several potential economical important traits such as grain size indicated by 100 seed weight, grain length and grain width which demonstrated a significant and positive association with grain yield per accessions may be used for genetic potential in the improvement programs of rice landraces.

The use of more number of DNA markers (>5) would be efficient to characterize the rice landraces conserved than the ones used in this study, which showed that most of the rice landraces were genetically identical accessions at genomic level among the genotypes studied.

APPENDICES

Appendix 1: Quantitative traits recorded in Rice diversity study

Abbreviation	characters	Phenotypic scale
NDH	Number of days to heading	(d)
PHT	Plant height	(cm)
NPT	Number of productive tiller	(-)
CUN	Culm number	(-)
CUL	Culm length	(cm)
FLL	Flag leaf length	(cm)
FLW	Flag leaf width	(cm)
PAL	Panicle length	(cm)
NDM	Number of days to maturity	(d)
GWT	100 grain weight	(g)
GYD	Grain yield per accession	(g)
GRW	Grain width	(mm)
GRL	Grain length	(mm)
GLW	Grain length-width ratio	(-)

Appendix 2: Data collected in rice diversity study

ACCNO	NDH	PHT	NPT	CUN	CUL	FLL	FLW	PAL	NDM	GWT	GRW	GRL	GLW
TZA2906	130	114	21	4	89	63.5	1.4	26	164	3.35	9.1	2.254	4.037
TZA2765	137	104.5	10	4	80	65.5	1.7	25	164	2.82	8.9	2.016	4.447
TZA1557	94	169	9	6	129	52	1.3	28	120	1.22	5.9	1.512	3.934
TZA2823	121	78	16	4	58	52	1.3	22	150	2.27	8.6	1.96	4.404
TZA2747	115	123.2	11	4	105	84.9	2.58	23.5	135	3.16	10.4	2.214	4.697
TZA3846	115	119.1	11	4	99	70.4	2.8	26.9	135	3.26	9.3	2.1	4.428
TZA4034	126	130.5	12	5	101.1	77	2.7	24.8	164	2.89	9.9	2.094	4.734
TZA3691	93	137	12.6	4	113.6	62	2.6	24.6	140	3.63	9.69	2.29	4.06
TZA3572	130	105	15	4	78	63	1.3	25.5	164	2.63	9.79	1.932	5.067
TZA1519	110	111.5	16	4	88	61.3	1.3	23	135	3.2	9.06	2.17	4.175
TZA2723	136	96	10	4	63	67	1.4	23.5	180	2.63	8.99	2.07	4.343
TZA3726	95	110	14	4	85	58.2	1.3	22.4	141	3.62	9.64	2.38	4.05
TZA4180	128	124.6	17	4	78.6	80.2	2.68	25.2	164	2.62	9.28	2.09	4.44
TZA2286	125	127.1	13	4	96.2	66	2.38	25.8	164	2.68	9.24	2.03	4.552
TZA2771	159	131	10	4	86	64	1.7	22.2	180	2.26	10.05	1.872	5.371
TZA2339	96	120	9	4	102	24	2.2	21.2	137	3.74	9.256	2.432	3.806
TZA2784	129	110	12	5	79	58	1.7	23	183	3.16	9.934	2.188	4.54
TZA3028	148	75	8	4	56	55	1.4	19	183	3.14	8.786	2.016	4.358
TZA2779	132	103	16	4	75	58	1.5	24	169	2.82	9.206	2.116	4.351
TZA2786	128	109.2	9	4	77	69.5	2.64	24.7	169	3.19	10.41	2.151	4.836
TZA2748	137	89.83	13	4	63.67	74.2	2.4	17.2	173	2.39	8.75	1.83	4.781
TZA2735	130	95	12	4	71	68	1.3	26	164	3.37	10.2	2.248	4.539
TZA2751	134	111	12	5	84	57	1.4	23	164	2.88	9.49	2.148	4.418
TZA2794	132	103	6	4	77	55	1.6	28	183	2.46	8.364	2.076	4.029
TZA2691	139	89	11	4	72	57	1.7	22	183	2.95	10.04	2.162	4.646
TZA4007	129	100	9	4	78	74	1.3	25	169	3.01	11	2.162	5.089
TZA2699	134	117	10	4	75	73	1.5	27	164	2.9	9.34	2.02	4.624
TZA1517	82	180	18	6	160	62	1.6	27.5	120	1.11	7.278	1.435	5.072
TZA2866	132	126	10	5	90	71	1.5	25.5	183	2	9.276	2.168	4.279
TZA3822	141	111	12	6	90	67	1.4	24	183	2.62	9.404	2.038	4.614
TZA3581	145	109	6	4	84	76	1.5	24.5	183	2.83	9.818	1.612	6.091
TZA2772	139	94	13	4	65	111	1.4	22	180	2.6	9.264	2.002	4.627
TZA2845	116	110	16	4	86	74	1.4	26	160	3.83	9.506	2.468	3.852
TZA2852	134	118	13	6	90	75	1.6	24	183	3.25	10.18	2.202	4.625
TZA2844	101	103	26	4	82	52	1.2	21.5	158	3.44	10.19	2.238	4.551
TZA1516	143	95	12	5	65	64	1.5	24.5	183	2.93	9.57	2.156	4.439
TZA1536	141	93	13	4	60	62	1.2	23	169	2.8	9.701	2.074	4.679
TZA1534	141	93.5	10	4	61	60	1.3	24.5	169	2.49	9.782	2.048	4.776
TZA2744	134	125	13	6	108	76	1.6	22	183	2.68	8.842	2.116	4.179
TZA4161	101	105	13	4	84	57	1.3	26	164	3.07	9.43	2.146	4.394
TZA2679	132	115	9	5	91	66	1.8	24	180	3.38	10.32	2.216	4.659
TZA4210	136	119	14	5	110	74	1.6	27	183	3.16	8.534	1.99	4.283
TZA2763	143	102	12	4	66	71	1.2	25	183	2.72	9.114	1.982	4.598
TZA2696	104	134.5	12	4	108	66	1.7	25.8	158	3.65	8.24	2.392	3.445
TZA2796	100	125.6	12.6	4	109.4	54	1.7	21.5	140	3.38	8.236	2.51	3.281
TZA2931	137	109	10	5	87	69	1.7	23	183	3.43	9.468	2.102	4.504

ACCNO	NDH	PHT	NPT	CUN	CUL	FLL	FLW	PAL	NDM	GWT	GRW	GRL	GLW
TZA1535	136	122	17	5	103	66	2.1	27	169	2.91	9.68	2.144	4.515
TZA2943	102	120	15	4	95	63.2	1.4	23.3	169	3.34	9.502	2.086	4.555
TZA2356	74	125	17	4	115	68	1.7	25	161	3.54	8.866	2.328	3.808
TZA2820	121	137	15	5	111	63	1.8	28	161	3.34	8.92	2.27	3.93
TZA1525	121	141	13	4	100.5	71	1.5	27.5	161	2.49	8.97	2.01	4.463
TZA2782	130	104	12	5	83	74	1.6	25	183	3.14	9.636	2.108	4.571
TZA2256	130	117	16	4	84	65	1.9	25	169	3.09	9.45	2.136	4.424
TZA2283	102	112	8	4	89.5	58	1.7	23	183	2.69	9.118	2.206	4.133
TZA2341	134	117.5	9	4	93	68	1.8	25	183	3.16	9.67	2.252	4.294
TZA2683	132	125	10	5	98	64	1.9	24	183	3.23	10.17	2.182	4.663
TZA2698	134	126.4	12	4	107.1	57.2	1.6	21.1	141	3.27	9.93	2.19	4.534
TZA3831	139	107	10	5	79	72	1.8	24	183	3.23	10.46	1.974	5.298
TZA1339	137	95	12	6	66	75	1.9	21.5	183	2.76	9.894	2.15	4.602
TZA3642	130	134	13	5	110	70	1.7	24	170	3.31	10.07	2.198	4.582
TZA2764	130	137	7	5	113	74	2	24.5	183	3.19	9.82	2.252	4.361
TZA4015	129	130	13	5	112	81	1.7	27	183	2.86	9.586	2.156	4.446
TZA3895	93	105	11	4	84	58	1.3	24	183	3.42	10.74	2.192	4.901
TZA2826	133	73	10	4	60	65	1.6	22	183	3.04	9.516	2.176	4.373
TZA2772	133	103	8	5	78	73	1.9	25	183	3.34	10.54	2.16	4.879
TZA4184	140	100	12	5	81	71	1.7	27	183	2.73	8.972	1.952	4.596
TZA4617	125	105	12	5	84	69	1.4	22	180	2.77	9.912	2.002	4.951
TZA2776	129	132	8	4	105	85	1.7	24	183	3.32	10.33	2.21	4.676
TZA2810	129	117	17	4	95	84	1.5	24	183	2.51	9.59	2.034	4.715
TZA2750	104	142.3	17	4	117.5	79	1.5	27	164	3.82	10.00	2.278	4.391
TZA3850	104	120	13	4	99	93	1.4	23	164	2.68	8.948	I.964	4.556
TZA4431	129	82	20	5	54	69	1.6	22.5	183	2.3	8.762	1.992	4.399
TZA2805	124	82	12	4	63	78	1.2	24.5	169	2.71	9.41	2.118	4.443
TZA2821	93	127.3	26	4	109.5	70.1	1.4	27.2	140	3.55	9.56	2.4	3.983
TZA2612	124	110	30	4	84	70	1.2	24	169	2.4	8.362	2.014	4.152
TZA2846	137	134	12	5	103	77	1.8	27	183	2.91	9.44	2.126	4.444
TZA4038	134	149	25	5	120	82	1.7	27	169	2.6	9.184	2.032	4.115
TZA2744	134	118	16	5	85	96	2.2	28	169	2.9	9.434	2.258	4.178