DOCTOR OF PHILOSOPHY

Genetic diversity study in landraces of rice (Oryza sativa L.) by agro-morphological characters and microsatellite DNA markers

Bajracharya, Jwala

Award date:
2003

Awarding institution:
Bangor University

Genetic Diversity Study in Landraces of Rice (*Oryza sativa* L.) by Agro-morphological Characters and Microsatellite DNA Markers.

A thesis submitted in candidature for the degree of

Philosophiae Doctor

By

Jwala Bajracharya (Chitrakar)

M.Sc. (Botany, Seed Technology)

School of Agricultural and Forest Sciences

University of Wales, Bangor

April 2003

Dedicated to my mother **Pancha Kumari Bajracharya** *and also to my father* **Jwala Bajra Bajracharya**, *who is no longer able to share my happiness on this achievement.*

ACKNOWLEDGMENT

I gratefully acknowledge the scholarship awarded by the International Plant Genetic Resources Institute (IPGRI), Rome, Italy. This research study was a part of Nepal Country Component of the IPGRI's Global Project on "*Strengthening the scientific basis of* in-situ *conservation of agricultural biodiversity on-farm*" and was financially supported by Netherlands Ministry of Foreign Affairs Development Cooperation, DGIS (WW104801) and DFID Plant Sciences Research Programme for the benefit of developing countries (R30243), UK. I am grateful to IPGRI, DGIS and DFID for the grants.

I am highly obliged and greatly indebted to my supervisor Dr John R Witcombe and to Dr Katherine A Steele whose help was invaluable for all of the research involving SSR markers. I sincerely thank them for their academic guidance, encouragement and support in carrying the research works and in write-ups of this thesis, for their frequent visits and supervision of field and molecular works carried in Nepal and finally for the kindness and sympathy shown when I had very hard time in Bangor due to illness (aneurysm) of my mother. I gathered some knowledge and experiences on molecular techniques during the course of this research and it was Dr John R Witcombe who suggested to work on microsatellite markers instead of isozyme. I am grateful to him.

I am very much thankful to Dr Devra I Jarvis, the International Project Coordinator, IPGRI Global Project on *In-Situ* Conservation, Dr Bhuwon R Sthapit, the Regional Coordinator, IPGRI-APO, and Dr Madhusudan P Upadhyay, National Project Coordinator, *In- Situ* Conservation, Nepal for their valuable technical back-ups, supports, helps and kind considerations all the time.

I express my thanks to Mr Raghunath P Sapkota, the Executive Director of NARC, Mr Bimal K Baniya, the Chief of Agriculture Botany Division and Dr Anil Subedi, the Executive Director of LI-BIRD for providing institutional support in conduction of the field and molecular research works and in establishing the molecular facilities in Biotechnology Laboratory in NARC. I thank Mr Dhurba Joshi, the then Executive Director, NARC for institutional approval and support on this aspect. I greatly acknowledge the financial support provided by IPGRI-APO and assistance of NIAR, Japan for procurement of the molecular equipment in establishing the facilities in NARC and made this research finally possible.

I extend my sincere thanks and appreciation to the individuals from *in-situ* conservation project teams and National agriculture research systems who have diligently worked in the rice fields and helped in taking the morphological observations and made the field research work possible: Krishna P Ghimire and his sons; Puspa R Tiwari, Deepak Panday, Paras Mahat, Janaki D Neupane, Jivan Shrestha; Ram S Karki, Janani Shrestha, Bina Dongol, Gunbhai Bajracharya; and Devendra Chaudhary, and Ram V Yadav. I also thank the project site coordinators: Deepak K. Rijal; Pashupati Chaudhary; and Deep M Shakya and the Chiefs of Agriculture Research Centres: Bishnu H Adhikari; Yam R Panday; Shambhu P Khatiwada; and Mahesh Vaidya for their cooperation during the field trials in respective research stations. I would like to take the opportunity to thank Dr Toby Hodgkin for his technical advice and kind visit to the field trial in Kaski. Foremost, I gratefully express my gratitude to local farmer families in Jumla, Kaski and Bara sites for their valuable seed samples and information, and their participation in the field research.

I learnt the basic molecular handlings for the first time with Julian Bridges and I was assisted technically by Go Ping Sun in Bangor. I thank them for their technical assistance in molecular works carried in Bangor. I also acknowledge the cooperation and arrangement provided by Dr. Hari P Bimba for molecular works carried in Nepal.

I extend thanks to Bharat Adhikari and Dr. Daljit S Virk for the statistical assistance in data analysis and interpretation. In this regard I must thank Krishna Manandhar for her help in entering the raw data.

Finally but not the least I wish to thank all the personnel in the project for their helps, we worked in team and gathered information in relation to this research work. I particularly thank Ram B Rana and Devendra Gauchan for the survey information. Besides, I had personal and academic interactions with families, friends, and seniors belonging to different professions in Nepal and Bangor. I sincerely thank them for their cooperation and sharing at the hard times: Ashok Mudwari, Dr Krishna D Joshi, Dr Thakur P Tiwari, , Dr Phil Hollington, Julie Abberley, Sandra Hughes and Gwen Edwards. I wish to thank the gathers of Nepalese students and their families at Bangor for company and homely environment that really made the life and stay easier.

After all, it is all the beliefs, loves, inspirations, and understandings of my families-in-law that I was encouraged to undertake this research work. I am incredibly and specially thankful to my mother-in-law, Nani Chhori Chitrakar, who always inspires and sees insights of my profession and has taken care of households and my son all the time.

I express appreciation and love to my son Rojin Chitrakar. Finally, I express heart-felt thanks, obligation and gratitude to my husband Ranjan Chitrakar for his love, profound understanding, and great moral support to carry on this research.

ABSTRACT

The genetic diversity of 632 rice landraces was studied from three ecosites representing three agro-ecozones of Nepal. There were 147 landraces from Jumla (2240-3000 m altitude), 291 from Kaski (668-1206 m) and 194 from Bara (80-90 m). Rice diversity was assessed by agro-morphological traits and microsatellite (SSR) markers. This research was part of the Nepal Country Component of the IPGRI Global Project on "*Strengthening the scientific basis of* in-situ *conservation of agricultural biodiversity on farm*".

A broad range of landrace diversity was under cultivation under traditional management systems in these three ecosites. The landraces were grown for home consumption, sale, and for religious and cultural purposes and were adapted to a range of agroecological niches.

The rice accessions from the three ecosites were evaluated in field experiments in 2000 and 2001. Each set of landraces was grown in a single trial in Agriculture Research stations of the ecosites. Forty three agro-morphological traits (qualitative and quantitative) were measured including those that the farmers used in distinguishing these landraces. Using this morphological variability data, diversity measures such as the Shannon Weaver diversity index (H') for qualitative traits, coefficients of variation (CV) for quantitative traits, and three multivariate procedures were used to estimate diversity and relationships within and between the groups of rice landraces that shared the same name. There was a moderate to high variability for the studied traits and many of them were useful for assessing the level of genetic diversity. For most morphological traits the landraces from Kaski and Bara were more variable (H' 0.7-0.8 and CV for quantitative traits of 18-20%), compared with landraces from Jumla (H' 0.2 and CV 8%). Cluster analysis and discriminant function analysis based on significant qualitative and quantitative traits discriminated between groups of landraces of the same name in Kaski and Bara. However, Jumla rice landraces having different names were closely related and had a narrow genetic base.

Seventy rice accessions (21, 24 and 25) comprising of accessions with 10 different names from each ecosite and a few Nepalese and Indian PPB varieties were assayed for genetic diversity at 39 SSR marker loci using agarose gel electrophoresis. The number and frequency of alleles, the polymorphic information content (Nei's gene diversity, PIC) and multivariate analyses on allelic data were used to measure the diversity within and among the landraces and varieties of rice produced by participatory plant breeding (PPB). The patterns of genetic diversity among landraces revealed by the SSR polymorphisms varied between the study sites and the varieties. Landraces from Kaski and Bara showed a high genetic variation with about 0.34 average molecular genetic diversity (PIC) and about 88% of the markers were polymorphic. Landraces from Jumla hardly varied for the SSR markers tested - only one marker was polymorphic. Multivariate analysis (cluster analysis) corresponded to the pattern of variation found with the morphological traits. Both showed there was high genetic dissimilarity among groups of Kaski and Bara landraces and low dissimilarity among Jumla landraces.

The study showed that the most abundantly grown landrace: Kathegurdi and Laltenger had little within-cultivar variation. Landraces grown in a small area by only a few households and genotypes with culinary importance: Basmati, Jetho Budho, Rato Anadi, Jhinuwa, Nakhisaro and Sathi, had a comparatively large within- and between-cultivar variation. PPB varieties were as diverse as the landraces, when assessed by SSR markers, and PPB is a useful approach for increasing diversity in environments with low landrace diversity.

Overall, the study showed that the mid-hill ecosite with diverse agro-ecological environments and the lowland ecosite with most favourable rice growing environments conserved a high level of rice landrace diversity. This diversity in morphological and SSR variation is discussed in relation to management of rice genetic resource of Nepal.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF PLATES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

Abs	Absorbance
ADO	Agriculture Development Office
APP	Agricultural Perspective Plan
ARS	Agriculture Research Station
AVRDC	Asian Vegetable Research and Development Centre
BAU	Birsa Agricultural University
CAZS	Centre of Arid Zone Studies
CBD	Convention on Biological Diversity
CBS	Centre Bureo of Statistics
DF	dilution factor
DFID/PSRP	Department for International Development, Plant Sciences Research Programme
DOAD	Department of Agriculture Development
FUDS	Farmers' units of diversity
FYM	Farm yard manure
GVT	Gramin Vikas Trust
IBPGR	International Board of Plant Genetic Resources
IPGRI	International Plant Genetic Resource Institute
IRRI	International Rice Research Institute
LARC	Lumle Agriculture Research Centre
LI-BIRD	Local Intiatives for Biodiversity, Research and Development
LMDG	Local multidisciplinary group
MALDI-TOF MS	Matrix assisted laser desorption/ionization time of flight mass spectrometry
NARC	Nepal Agricultural Research Council
NMDG	National multidisciplenary group
PPB	Participatory plant breeding
PRA	Participatory rural appraisal
PVS	Participatory variety selection
SDW	Sterilized distilled water
VDCs	Village development committees

CHAPTER I

GENERAL INTRODUCTION

1.1 Introduction

Rice (*Oryza sativa* L.) is one of the ancient cultivated crops. It is the world's most important staple crop, feeding more than a third of the world's population (Khush, 1997) and is grown most extensively in latitudes ranging from 53°N to 35°S in tropical and subtropical regions of the world (Chang and Oka, 1976; and Lu and Chang, 1980). It has a broad adaptation to a wide range of growing environments and the highest altitude at which rice is grown is 2621 m above sea level in Nepal (Shahi and Heu, 1979).

South and Southeast Asia is believed to be a centre of diversity of rice, where large areas of marshy lands are suitable for its cultivation, and where many wild rice species are found growing naturally (Stoskopf, 1985; and Vavilov, 1930). More than 90% of the world's rice is grown and consumed in Asia, where 60% of the earth's people reside and depend upon it for their livelihood (Khush, 1997).

Rice is the most important cereal grain crop of Nepal and plays a significant role in Nepalese agriculture in terms of the area in which it is cultivated, its importance for employment generation, and its contribution to food security at both the national and household level. It has the highest priority for agricultural research in Nepal as envisaged in the Agricultural Perspective Plan (APP, 1995). Rice alone occupies over 1.5 million ha of total agricultural land and produces over 4 million t of grain, accounting for a total of 60% of the cultivated land, and 58% of the total grain production. Out of the total rice area of the country, 73% is located in terai (plain), 24% in hill and 3% in high

hill districts (CBS, 2000). Of these, 48% of the total rice area is irrigated, and on average 29% is covered by unimproved local rice (Table 1.1).

Table 1.1: Area, production and yield of rice by agro-ecological zones in Nepal, 1999/00.

	Irrigated	Unirrigated	Improved	Local	Total
Rice area (,000 ha)					
Mountains	14	32	26	20	46
Hills	167	211	227	150	377
Terai	559	568	850	277	1127
Nepal	740	811	1104	447	1551
Rice production (,000 t)					
Mountains	32	56	54	34	88
Hills	451	456	610	297	907
Terai	1697	1338	2450	585	3035
Nepal	2180	1850	3114	916	4030
Average yield (t ha^{-1})					
Mountains	2.1	1.8	2.1	1.7	1.9
Hill	2.5	2.2	2.7	2.0	2.4
Terai	2.7	2.3	2.9	2.1	2.7
Nepal	2.6	2.3	2.8	2.1	2.6

Source: Statistical information on Nepalese agriculture, 1999/2000, HMG/MOA (CBS, 2000)

The cultivation and diversity of rice in Nepal is unique with a fascinating diversity of rice growing environments. Rice is grown in a range of agro-ecozones ranging from the warm subtropical in the foothills (100-1000 m altitude) to temperate in the high mountains of the Himalayas up to a recorded altitude of 2621 m. Double rice cropping is possible in the plain regions up to 90 m. The five major rice growing environments in Nepal: are early rice with assured irrigation; main rice with partial or full irrigation; high-altitude rice with rainfed or partial irrigation; upland rice on totally rainfed tars; and deepwater rice (submerged rice areas of the *Terai*) (Sthapit, 1994).

The sustainability of agricultural ecosystems is dependent on the local landrace diversity (Tilman, 1996; and Tilman *et al.,* 1996). The maintenance, utilization and management of this diversity in the field are crucial to sustainable agriculture especially for the resource poor farmers who practice agriculture under low input marginal conditions (Worede *et. al.,* 2000). Landraces are passed from generation to generation of farmers as a part of their heritage. The landraces are exposed to conscious selection for preferred traits and also to natural selection for adaptation to the local environment (Loutte *et. al.*, 1997; and Teshome *et. al.,* 1997). Landraces constitute a conspicuous source of variation and provide useful genes and traits for crop improvement (Worede, 1991; and Zeven, 1998). Landrace diversity is more likely to be conserved by resource-poor farmers who cultivate rice in marginal growing environments and who have low adoption of new technologies (Shrivastava and Damania, 1989; and Byerlee and Moya, 1993). Fragmented land holdings, imperfect market conditions, and diverse cultures and preferences all promote the use of a diverse range of landraces (Brush, 1995; and Gurung and Vaidya, 1998). Traditional practices of seed introduction and exchange system among the farming community, contribute to maintaining crop landrace diversity on farm (Joshi *et al.*, 1997; and Rana *et. al.*, 2000a,b,c). On the other hand, participatory approaches such as participatory variety selection (PVS) and participatory plant breeding (PPB) have positively contributed to the diversity conservation by the exploitation of the existing variation in local landraces in the development of improved varieties (Sthapit *et. al.*, 1996; Witcombe *et. al.,* 1996; and Zhang *et. al.,* 1994). In Nepal, several released varieties have landraces in their ancestry (Table 1.2).

Table 1.2: Examples of utilization of local landraces of rice as parents for breeding improved varieties in Nepal. Landraces are indicated in italics.

Variety/Strain	Parents	Main characters
Khumal 2	*Jarneli*/Ku-16-361-BLK-2-8	Fine grain
Khumal 4	IR 28/*Pokhreli masino*	Fine grain
Palung 2	BG 94-2/*Pokhreli masino*	Fine grain
Chhomrong	Local selection	Cold tolerant
Khumal 5	*Pokhreli masino*/Ku-1B-361-BLR-2-6	Fine grain
Machhapuchhre 3	*Chhomrong*/Fuji 102	Medium grain

Source: Plant genetic resources, Nepalese perspectives, 1995

Nepal is within a centre of primary diversity of crop genetic resources (Zeven and Zhukovsky, 1975). The coexistence of wild rice species (*Oryza nivara,* Sharma and Shastry; *O. rufipogon,* Griffith; *O. granulata* Nees et AM ex Watt; and *O. officinalis* Wall ex. Watt) with natural populations of relatives of rice (*Hydrorhiza aristata* Nees; and *Leersia hexandra* L.) is part of the evidence that Nepal is a center of diversity for rice (Shrestha and Vaughan, 1989; Adhikary *et. al.*, 1995; and Upadhyay, 1995). This was the rationale of selecting Nepal as a partner country in the Global project on "Strengthening the Scientific Basis of *In Situ* Conservation of Agricultural Biodiversity on Farm" initiated by the International Plant Genetic Resources Institute (IPGRI). *In situ* conservation is the conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species of natural and cultivated crops in their habitat in contrast to *ex situ* conservation, the conservation of components of biological diversity outside their natural habitats (UNCED, 1992). *Ex-situ* conservation involves sampling germplasm from the field and its subsequent storage in gene banks. It cannot conserve full range of genetic diversity and nor can it conserve the dynamic process of crop evolution and the farmers' knowledge of crop selection and maintenance (FAO,

1996). *In situ* conservation - on-farm conservation, aims to maintain the genetic diversity of crops and their wild relatives in farmers' fields (Maxted *et al.*, 1997). However, *in situ* conservation techniques are less well developed than those of *ex situ* conservation and the two methods are complementary.

The objectives of the IPGRI *in situ* project are to

- support the development of a framework of knowledge on farmer decision-making processes that influence the *in situ* conservation of agricultural biodiversity
- strengthen national institutions for the planning and implementation of conservation programmes for agricultural biodiversity
- broaden the use of agricultural biodiversity and the participation of farming communities in its conservation and other groups in its conservation (Jarvis and Hodgkin, 1998).

This study is an activity of the IPGRI *In Situ* Conservation Project and forms part of the crop biology component.

Loss of genetic diversity is the main threat to the sustainable use of plant genetic resources to meet the present needs and aspiration of future generation (Chang, 1985). Loss of agricultural habitats, the genetic erosion by replacement of indigenous germplasm (landraces) by improved varieties, land degradation due to natural catastrophes, and commercialization of agriculture are some of the processes which contribute to the loss of landraces, their wild relatives and their habitats (Qualset *et al.*, 1997; and Pham *et al.*, 2001). The extensive displacement of rice landraces between 1970-1990 by varieties produced by IRRI research in South-East Asia was documented

by Perlas and Vellve, 1997. To an extent, this has also happened with the rice diversity in Nepal. For example, *Tauli, Marshi,* and *Thapachiniya,* were once popular landraces of the Kathmandu valley, but are no longer in cultivation in farmers' field (Upadhyay, 1995). It is, therefore, a prime concern to conserve the remaining diversity for utilization by future generations.

For the efficient utilization and conservation of diversity information is needed on the genetic variation. Agro-morphological characterisation of traits in the field, and laboratory-based biochemical and molecular markers are efficient tools for assessment of genetic diversity. Jackson *et al.*, (1999) pointed out the efficiency of these tools and these were supplementary to each other for measurement of diversity.

The most easy and common tool in gauging genetic diversity is that of measuring agro-morphological or phenotypic trait differences. Farmers use certain phenotypic features of plants for selection and identification. The resultant crop varieties (landraces) are named by farmers to distinguish them and to help in selecting the seed from true-to-type plants from season to season. Hence agro-morphological traits are linked to the genetic diversity through naming farmers' varieties, indicating that farmers have some understanding about the crop genetic diversity in their fields (Jarvis *et. al.*, 2000). The assessment of diversity in morphological traits therefore, is central to the on-farm conservation of crop varieties.

Diversity is also measured by differences in biochemical, protein and molecular (DNA) properties within and between plant populations. The assay of isozymes has been an efficient and predominant technique for examining genetic diversity. It has added new dimensions to the studies of many plant species and thereby provided insights into their

population genetics (Nielson, 1985). DNA markers are increasingly being recognised as useful tools for assessing genetic diversity amongst plant species since these are less influenced by the environment (Lee, 1995) and are more ubiquitous and more random than isozyme markers. These marker systems include restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeat polymorphisms (SSR or microsatellites) (Wang and Tanksley, 1989; Jena *et al.*, 1992; Newbury and Ford-Lloyd, 1993; Ford-Lloyd and Newbury, 1994; Mackill *et. al.*, 1996; Maughan *et. al.*, 1996; Ellis *et. al.,* 1997; Lanaud and Lebot, 1997; Virk, *et al.*, 2000; and Shrivastava *et. al*, 2001). Microsatellites or simple sequence repeats (SSRs) are locus-specific and co-dominant PCR-based markers. This technique has extensively been applied in variety identification, mapping and genetic diversity studies of rice (McCouch *et. al.*, 1988; Kurata *et. al.,* 1994; Yang *et. al.,* 1994).

The present study on rice diversity used both agro-morphological and microsatellite DNA marker techniques for measuring the genetic diversity, Nepal.

1.2 Aim of research

- To explore the extent and distribution of rice landrace diversity maintained by the farmers relative to the traditional and economic importance of landraces in terms of number of named varieties and area under cultivation.
- To examine the genetic variation in rice landraces in three ecosites selected by the project, by the use of agro-morphological traits
- To use (SSR) microsatellite DNA markers to study the genetic relationships between rice landraces named by farmers as distinct units (varieties) in three ecosites.
- To measure and compare the amount of diversity present within and between farmer-named rice varieties in each ecosite.
- To study and measure the level of genetic variability amongst different groups of rice germplasm comprising PPB varieties, PPB lines and populations, modern cultivars and landraces.
- To establish a molecular testing laboratory in Nepal Agricultural Research Council (NARC), Khumaltar, Nepal for undertaking research towards in characterising the important plant genetic resources of the country.

CHAPTER II

LITERATURE REVIEW

2.1 Concept and history of genetic diversity

The word biodiversity was evidently coined for the first time by Walter G. Rosen in 1985 in a meeting in Washington called the "National Forum on Biodiversity". Two years later, it came in to general use after the publication of the proceedings entitled "Biodiversity" edited by E.O.Wilson (Zhu, 1996). The definition of biodiversity, accepted and signed by 156 nations and European community, has been documented in an official document, "United Nations Environment Programme, 1992" (page 27) is as follows:

"Biological diversity" means the variability among living organisms from all sources including inter alia, terrestrial, marine and other aquatic systems or parts, this includes diversity within species, between species and of ecosystems.

The diversity of living organisms has been recognized and studied at different levels throughout human history. The importance of diversity, its utilization and the threat imposed on it have been knowingly and unknowingly recognized and understood as a part of human conciousness (Frankel, *et al.*, 1995). A knowledge and experience of diversity led Linnaeus to recognise taxonomically discrete entities and arrange them into a hierarchy of species, genera, and families. Vavilov, the great explorer, (1926) explored the world distribution of crop plants and identified the geographical centers of diversity of present day crop species.

Genetic diversity is the lowest level of diversity among three levels of biodiversity: genetic, species and community (agro-ecosystem) diversity. Genetic diversity is the amount of genetic variation present in a population, which is a group

of related individuals having different genotypes (genetic structures and genetic constituents). The ability of a species to respond adaptively to varied environments depends on the level of genetic diversity it contains (Ayala and Kiger, 1984). The diversity is the result of natural evolutionary processes such as natural selection and hybridisation, spontaneous mutation, migration (dispersion and geographical isolation) and genetic drift that, over several millennia, has led to the enormous range of today's adapted genotypes. Therefore, genetic diversity is essential for the long-term survival and continuing evolution of plants and animals (Templeton, 1994). It reduces genetic vulnerability and, in crop plants, provides an effective gene pool for crop improvement that can help feed the ever-growing population in the world. Genetic diversity in a crop gene pool can be classified into wild relatives of crop plants, landraces and the germplasm in formal breeding programmes. Wild relatives and landraces are considered to account for the bulk of the diversity within the system (Eyzaguirre and Iwanga, 1995). Because of their diversity, landraces have been conserved for their usefulness in agricultural production systems and serve as a source of genetic material in crop improvement.

2.2 Origin, biodiversity and adaptation of rice

2.2.1 Botany and genetics of rice

Rice is an autogamous plant belonging to genus *Oryza* of the family Poaceae. The genus *Oryza* is thought to have originated at least 130 million years ago and spread as a wild grass in Gondawanaland, the super continent that eventually broke up into Asia, Africa, the America, Australia and Antartica (Chang, 1976). *Oryza sativa* L. and *Oryza glaberrima* Steud. are the two cultivated species of rice. *Oryza sativa* L. is the Asian cultivated rice grown worldwide, hence is also known as common rice. It

includes all the cultivated varieties grown in America, Asia and Europe. *Oryza glaberrima*, the African rice, is cultivated on a limited scale in West Africa. *O. sativa* has 24 somatic chromosomes and is usually considered as, and behaves like, a diploid species. However, there are some genetic and cytological evidences indicating the existence of: tetraploid species with 48 somatic chromosomes; even haploid species with a set of 5 chromosomes (Nandi, 1936); and amphiploid species (Poehlman, 1979).

2.2.2 Interrelationships in species of *Oryza*

In the genus *Oryza*, there are a reported twenty-one wild species, as well as the two cultivated species, and these are distributed throughout Asia, Africa, America and Oceania (Khush, 1997). Roschevicz in 1931 classified these species into six sections including wild and cultivated species (Tateoka, 1963). However, Morishima and Oka (1960) studied the morphological variation in 16 of these species and suggested that these rice species can be divided into three main groups:

- *Oryza sativa* and its relatives (primary gene pool),
- *Oryza officinalis* and its relatives (secondary gene pool), and
- other more distantly related species (tertiary gene pool).

The classification of these 23 species of rice into groups, the geographical distribution of the species and their useful traits are given in Table 2.1. Species of wild and cultivated rice belonging to *O. sativa* complex constitute the primary gene pool with good ease of gene transfer among them. Species belonging to the *O. officinalis* complex constitute the secondary gene pool. Species belonging to the *O. meyeriana*, *O. ridleyi,* and *O. schlechteri* complexes constitute the tertiary gene pool. Crossing between *O. sativa* and the wild species of secondary gene pool is extremely

difficult. Crossing between *O. sativa* and the tertiary gene pool if it occurs at all, is very rare.

The two cultivated species of rice, *O. sativa* and *O. glaberrima,* are thought to have evolved in parallel from a common ancestor in Gondwanaland (Figure 2.1). The wild progenitor of common rice, *O. sativa,* is the Asian common wild rice, *O. rufipogon* Griffith. This wild species constitutes a range of variation from perennial to annual types and has a high out-crossing rate and low seed productivity. Annual forms of this wild species were domesticated as *O. nivara* to become *O. sativa.* The wild perennial progenitor of the African cultivated rice, *O. glaberrima,* was *O. longistaminata* and its annual wild form was *O. breviligulata.*

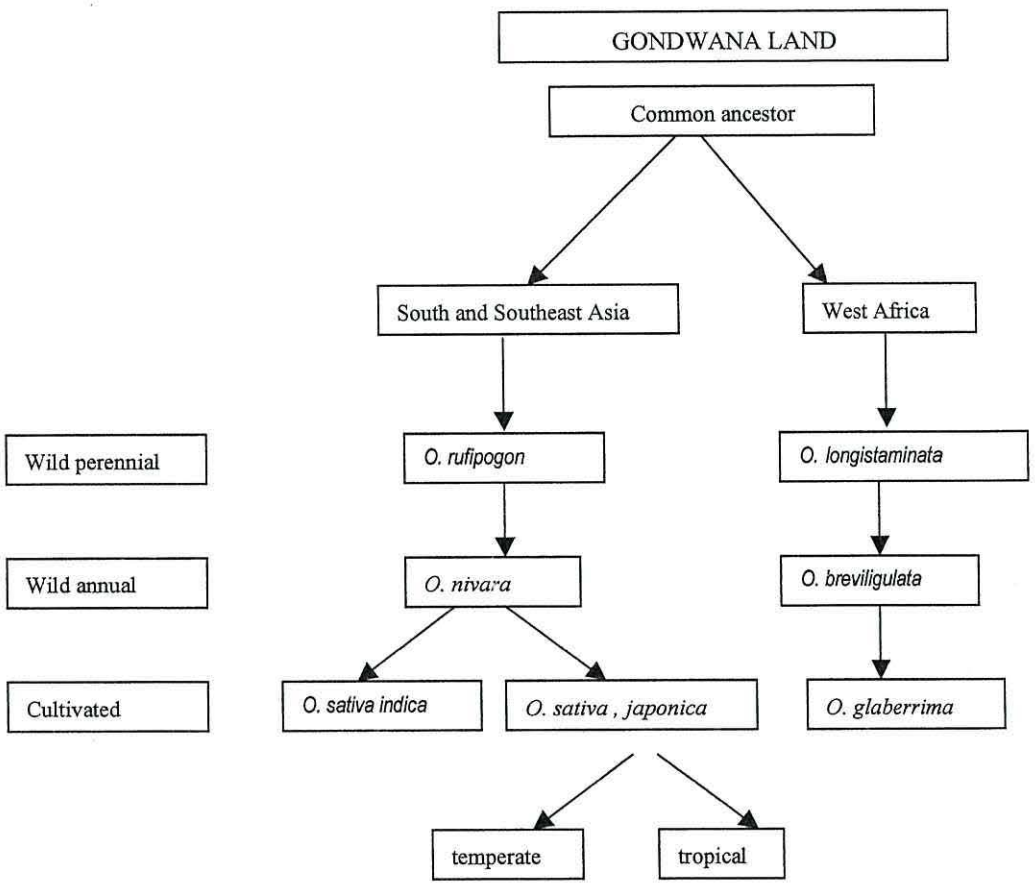


Figure 2.1: Evolutionary pathway of two cultivated species of rice (Khush, 1977).

Table 2.1: Classification and distribution, chromosome number, genome and potential traits of *Oryza* species (Khush, 1997).

Species	2n	Genome	Distribution	Useful traits
***O. sativa* complex**				
O. sativa L.	24	AA	Worldwide	Cultigen
O. nivara Sharma and Shastry	24	AA	Tropical and subtropical Asia	Resistance to grassy stunt virus, blast, and drought avoidance
O. rufipogan Griff.	24	AA	Tropical and subtropical Asia, tropical Australia	Elongation ability, resistance to BB, source of CMS
O. breviligulata A. Chev. et Roehr.	24	A^gA^g	Africa	Resistance to GLH, BB, drought avoidance
O. glaberrima Steud.	24	A^gA^g	West Africa	Cultigen
O. longistaminata A. Chev. et Roehr.	24	A^gA^g	Africa	Resistance to BB, drought avoidance
O. meridionalis Ng	24	A^mA^m	Tropical Australia	Elongation ability, drought avoidance
O. glumaepatula Steud.	24	$A^{gp}A^{gp}$	South and Central America	Elongation ability, source of CMS
***O. officinalis* complex**				
O. puncatata Kotschy ex Steud.	24, 48	BB, BBCC	Africa	Resistance to BPH, zigzag leafhopper
O. minuta J. S. Pesl ex C. B. Presl.	48	BBCC	Philippine and Papua New Guinea	Resistance to sheath blight, BB, BPH, GLH
O. officinalis Wall ex Watt	24	CC	Tropical and sub tropical Asia	Resistance to thrips, BPH, GLH, WBPH
O. rhizomatis Vaughan	24	CC	Sri lanka	Drought avoidance, rhizomatous
O. eichingeri A. Peter	24	CC	South Asia and East Africa	Resistance to yellow mottle virus, BPH, WBPH, GLH
O. latifolia Desv.	48	CCDD	South and Central America	Resistance to BPH, high biomass production
O. alta Swallen	48	CCDD	South and Central America	Resistance to striped stemborer, leaf folder, high biomass production
O. grandiglumis (Doell) prod.	48	CCDD	South and Central America	High biomass production
O. australiensis Domin.	24	EE	Tropical Australia	Drought avoidance, resistance to BPH
O. brachyantha A. Chev. et Roehr.	24	FF	Africa	Resistance to yellow stem boree, leaf-folder, whorl maggot
***O. meyeriana* complex**				
O. granulata Nees et AM. Ex Watt	24	GG	South and South East Asia	Shade tolerance, adaptation to aerobic soil
O. meyeriana (Zoll et Mor. Ex Steud.) Baill	24	GG	South East Asia	Shade tolerance, adaptation to aerobic soil
***O. ridleyi* complex**				
O. longiglumis Jansen	48	HHJJ	Irian Java, Indonesia, and Papua New Guinea	Resistance to blast, BB
O. ridleyi Hook. f.	48	HHJJ	South Asia	Resistance to stem borer
Unknown genome				
O. schlechteri Pilger	48	Unknown	Papua New Guinea	Stoloniferous

BPH = brown plant hopper; GLH = green leaf hopper; WBPH = white-backed plant hopper; BB = bacterial blight; CMS = cytoplasmic male sterility.

Rice is a genetically diverse species having a broad adaptation to a wide range of growing environments and a worldwide distribution. The diversity of the rice crop has evolved over thousands of years, as Asian and African peasant farmers selected different types to suit local cultivation practices and needs. This process of selection has led to numerous rice varieties adapted to a wide range of agro-ecological conditions, and with resistance to insect pests and diseases. Cultivated rice, *Oryza sativa* L., consists of two widely grown subspecies: *indica* and *japonica* according to the classification introduced by Kato *et al.* (1928). The Chinese have also recognized these two rice varietal groups as '*Hsien*' and '*Keng*', since the Han dynasty. These two varietal groups are the specialized gene pools, which together make it possible to cultivate rice under diverse geographical and ecological habitats ranging from tropical to subtropical and to more temperate zones and from irrigated to rainfed environments (IRRI, 1993; and Yang *et al.,* 1994).

In the polyphyletic origin of *O. sativa* (Khush, 1997), a third group of rice as *javanica* has been described (Moringa, 1954). This group was proposed to include the 'bulu' and 'gundil' varieties of Indonesia. In a genetic affinity study using isozyme analysis, Glaszmann (1987) found that *javanica* varieties fall within the *japonica* group II and are referred as tropical *japonicas*. The temperate upland *japonica* group is typically *japonica*, and can be differentiated from other *japonicas* by taxonomic, biochemical means (by $KClO_3$ resistance, phenol reaction and isozyme analysis) and by their physiological cold resistance (Oka, 1958; and Glaszmann, 1987).

Rice is an ideal model plant for the study of genetic and genomic analysis due to:

- its diploid genetics (2n=24),
- relatively small genome size (C=0.45pg, Causse *et al.,* 1994),

- smallest genome size of all major crop plants of 430Mb (Arumuganathan and Earle, 1995),
- one of the most densely mapped crop species with an average DNA to cM ratio of 250-300Kb/cM resulting in a combined total of over 2000 mapped markers (Kurata *et al.*, 1994),
- significant level of genetic polymorphism at the molecular level (McCouch *et al.*, 1988; and Wang and Tanksley, 1989), and
- the availability and conservation of a large amount of genetically diverse materials.

There are >200,000 accessions of both domesticated and wild rice maintained by rice researchers worldwide (Jena *et al.*, 1992).

2.2.3 Domestication of cultivated rice

The domestication of wild rice to produce today's cultivated rice has a very long history. It probably started about 9000 years ago. In India, carbonized grains of rice were found that dated back to about 8040 years ago. In China, a source of evidence for the domestication and cultivation of rice was the intact carbonized rice grains found in Zhejiang province in China, which were estimated to be 7040 years old. The pottery shards bearing the imprints of grains and husk of *O. sativa* discovered by Welhelm G. Solheim II in 1966 at Non Nok Tha in the Korat area of Thailand, were another archaeological evidence for domestication of rice in Southeast Asia and dated back to at least 6000 years ago (Chang, 1976 and Oka, 1988). The antiquity of rice cultivation and

the linguistic language used for rice, and for foods derived from rice, in South and Southeast Asia reveal it as a centre of diversity of rice.

2.3 Plant genetic resources (PGR) conservation approaches

The importance of crop diversity, and its conservation for exploitation in improved agricultural and horticultural crops, is acknowledged by conservation scientists (curators), breeders, and policy makers. Conservation is the process that secures and holds actively the diversity of the gene pool and makes it readily available for utilization. *Ex situ* and *in situ* conservation are the two basic, complementary conservation strategies, each being composed of various techniques with the common objectives of capturing a broad genetic base and halting the rate of loss of species and genetic diversity for the direct, and indirect, benefit of humankind (Maxted *et al.*, 1997; and Prance, 1997). The Convention on Biological Diversity (CBD) at the Earth Summit in 1992 has highlighted the need for an approach to biodiversity conservation that uses both *ex situ* and *in situ* conservation in a complementary manner (Maxted *et al.*, 2002).

2.3.1 *Ex situ* conservation of PGR

The definition of *ex situ* conservation, as defined in Article 2 of the Convention on Biological Diversity (UNCED, 1992), is the conservation of components of biological diversity outside their natural habitat. The technique involves the maintenance of living collections of plants conserved in field gene banks or botanical gardens or the maintenance of samples of seed, tubers, tissues, pollen or DNA maintained under controlled artificial conditions. It is static, efficient, and retains the structure of the

original populations (Guldager, 1975) and makes the germplasm readily available for use by breeders and researchers (Ford-Lloyd and Jackson, 1986). The storage of seed in gene banks has been the principal strategy for the conservation of most crop genetic resources. It is the most favoured strategy for the conservation of rice germplasm (Bellon *et al.*, 1997), as rice has orthodox seeds, i.e., they can be dried to a relatively low moisture content and can be stored safely at subzero temperatures for short-, medium- or even long-term with limited loss of viability. However, the technique freezes the process of evolution so the germplasm conserved is not being adapted to changes in environment with respect to new races of pests, diseases, and major climatic changes (Maxted *et al.*, 1997).

2.3.2 *In situ* conservation of PGR

In situ techniques involve the maintenance of the genetic diversity in the location where it has been encountered either in wild or cultivated forms in traditional farming systems. According to the definition in Article 2 of the CBD (UNCED, 1992), it is the conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species in their natural surroundings and, in the case of domesticates or cultivated species, in the surroundings where they have developed their distinctive properties (Maxted *et al.*, 2002).

On-farm conservation is one of techniques of *in situ* conservation, which involves the on-farm conservation of local crop landraces with the active participation of farmers. Ford Lloyd (1990) has pointed out the advantages and limitations of *in situ* conservation especially with regard to the global warming and the position of world

germplasm of horticultural crops. However, less progress has been made in *in situ* conservation compared to *ex situ* conservation and the scientific understanding *in situ* conservation is also more limited. This is because the *in situ* conservation strategy has only lately come into practical use as a complementary method to *ex situ* conservation (Brush, 1991; and IPGRI, 1993).

The International Plant Genetic Resources Institute (IPGRI), Rome, Italy is committed to advancing the conservation and use of plant genetic resources for the benefit of current and future generation. IPGRI has initiated a project on "Strengthening the scientific basis of *in situ* conservation of agricultural biodiversity on-farm" in nine different countries. Nepal was the one of the participating countries from Asia in the project and the work described in this thesis is part of this project.

2.3.3 Participatory plant breeding (PPB) as a means of conserving biodiversity

Landraces were the only varietal choices for farmers in marginal rice growing environments in Nepal (Rana *et al.,* 2000a,b,c). This is true in most developing countries, as farmers in marginal areas are little exposed to improved cultivars, the alternative to the landraces. Participatory varietal selection (PVS) and participatory plant breeding (PPB) are two farmer participatory approaches for the breeding and identification of improved crop varieties acceptable to resource-poor farmers (Witcombe *et al.*, 1996; and Sthapit *et al.*, 1996). PVS and PPB can have a long-term impact on maintaining inter- and intra-varietal diversity on farm.

An approach for increasing the choice of varieties available to resource-poor farmers, participatory variety selection (PVS), has been tried by many authors.

Chambers (1998) described approaches to giving farmers choices of different genetic material. Maurya *et al.* (1988), Sperling *et al.* (1993), Weltzien *et al.* (1996), and Joshi *et al.* (1997) used participatory research in identifying advanced lines and cultivars of rice, beans and pearl millet in farmers' fields or on research stations. In PVS programmes in rice in India and Nepal, new varieties of rice were successfully adopted and spread very rapidly with a major impact on temporal[1], average[2] and weighted diversity[3] (Witcombe, et al., 2001).

Participatory plant breeding is the breeding and selection of genotypes by farmers from segregating generations to create improved new varieties (Witcombe, *et al.*, 1996). By using diverse parents in the crossing programme the products of PPB can contain new genes and diverse combinations of genes. There are evidences that genetic diversity of rice has been increased within high-altitude villages of Nepal by the adoption of PPB materials (Sthapit *et al.*, 1996). A Department for International Development, Plant Sciences Research Programme (DFID PSRP) project on PPB for high potential production systems in rice is underway in Nepal. Farmers have been successful in selecting a range of advanced lines with different ranges of maturity groups from

---

[1] Temporal diversity is a measure of change in the cultivar diversity over time. It indicates the replacement of the cultivars from time 1 to time 2.

[2] Average diversity is the diversity among cultivars growing in a specified region unweighted by the cultivated areas they occupy.

[3] Weighted diversity is the average diversity where each cultivar is weighted by the percentage area they occupy.

segregating progenies of a cross between Kalinga III x IR64, and some of this material has been included in the research described in this thesis.

2.4 Assessing genetic diversity

The genetic diversity of crop plants is broadly defined as the extent of dissimilarity among a set of materials. This information is necessary for the efficient selection of parental lines for new crosses and is essential for planning genetic conservation programmes and the use of conserved biodiversity (Tatineni *et al.,* 1996; Frankel *et al.*, 1995; and Virk *et al.*, 2000). Hence, detailed studies on the level and distribution of genetic variation have become a prerequisite for the efficient management of crop diversity (Hamrick *et. al.*, 1991).

Information on the level and distribution of genetic variation within a species can best be obtained empirically from differences in agro-morphological and physiological traits, biochemical (seed proteins and isozyme) traits and DNA sequences. Witcombe (1999) reviewed the methods of measuring genetic diversity that have been used in a range of crop species. In rice, markers such as isozyme, RFLP, RAPD, ISSR (inter simple sequence repeats) and AFLP have been applied and diversity indices and patterns have been assessed using different sets of germplasm. Cox and Wood (1999) summarised the methods for quantifying the diversity within and between populations or cultivars and listed their strengths and weaknesses (Table 2.2).

Diversity at the allele level can be explained as relative *richness* and *evenness* (Frankel *et al.,* 1995). *Richness* is the total number of different genotypes present in a population or sample and the *evenness* is the equality in frequency of the different types

in the population or sample. Witcombe (1999), and Widawsky and Rozelle (1998) reviewed the measures of diversity that are in use in estimating diversity of crop (See, for example, the footnotes on page 19 of this Chapter).

Table 2.2: Methods of quantifying genetic diversity (adapted from Cox and Wood, 1999).

Type of data	**Example**	**Genome coverage**	**Proximity to DNA diversity**	**Problems**
Pedigree	Coefficients of parentage; ancestral composition	Complete	Low	Inaccuracy of breeders' record; missing data, random selection and biases; intra-line variation disregarded
Phenotype	Polygenic traits	Low to high	Low	Expression of similar phenotypes by dissimilar genotypes; genotype x environment interaction
	Major gene traits	Low	High	Usually very few loci available; non-random sample of genome; loci may be subject to strong selection; dominance in non-inbred lines
Genetic marker	Protein level	Low to moderate	High	Number of loci limited; non-random sample of genome; some loci subject to selection; dominance in non-inbred lines
	DNA level	Moderate to high	Highest	'Identical' alleles not necessarily identical by descent; relationship, if any, between markers and phenotypes usually unknown; lack of polymorphism; possible biased sample of genome if non mapped markers are used.

2.4.1 Farmers' perceptions and knowledge of diversity

Farmers continue to grow landraces to meet their agronomic or cultural needs and they have a well-developed, indigenous knowledge of the crops and crop varieties they grow (Bellon 1991). Bellon *et al.* (1997) reviewed the factors affecting the maintenance of diversity in different folk communities across the world with examples of different crop species. He identified seed flows, variety selection, variety adaptation and seed selection and storage as the four components of farmers' diversity management of a crop and these were influenced by the agro-ecological, socio-economic and cultural conditions. For example, Conklin (1957); Lambert (1985) and Lando and Mak (1994) in diversity studies in rice reported that the farmers selection concerns are not homogeneous. Rich and poor farmers, farmers in productive or marginal areas and even within a farming household, male and female farmers of the same household, may have different needs in a crop variety. In the baseline survey of the IPGRI *in situ* conservation project in Nepal, many of these factors were also found to affect the traditional management of rice landrace diversity across and within the study ecosites (Rana *et. al.*, 2000a,b,c). Farmers' indigenous knowledge is thus linked to the maintenance and management of the genetic diversity of a crop (Eyzaguirre and Iwanga, 1995; and Jarvis *et al.,* 2000). The choice of a crop or landrace is first determined by adaptation to the agro-ecological domain and farm management practices under which it will be grown. Next farmers select for phenotypic features of plants that better meet their preferences or identification.

2.4.2 Agro-morphological variation

The most easily obtained and commonly used assessment of genetic variation in a population is a measure of the morphological differences. It is the classical way of assessing genetic diversity and is still the only way for some minor crops. The morphology of the plant is assessed in the field and does not require breeding and laboratory studies. However, assessment of agro-morphological variation can be very time-consuming, requires careful experimental design and data analysis, and is difficult to conduct on species with long growth duration.

Morphological markers and quantitative traits have been commonly and traditionally used to estimate genetic distances and classify landrace varieties (Goodman, 1972). Moreover, using morphology to estimate genetic variation is often the only practical and realistic way to rapidly estimate variability and it is the only possibility where biochemical and molecular analyses cannot be made and the resources and time are unavailable for quantitative genetic studies.

Asian rice varieties show an impressive range of variation in qualitative and quantitative characters. Chang and Bardens (1965) developed descriptors for rice based on morphological and biochemical characters of the rice plant. Variation in grain characters such as size, shape and colour has often been used in distinguishing and classifying rice varieties into coarse and fine types. Caldo *et al.*, (1997), in a morphology-based diversity analysis of 78 improved rice varieties and their ancestral lines, found 41 morphological traits which contributed total variation in ancestral lines and 33 traits in improved varieties. Traits such as basal leaf sheath colour, blade colour, collar colour, apiculus colour, stigma colour, culm length, and panicle length were important variables

distinguishing in improved rice varieties, whereas differences in ancestral lines were mostly for quantitative traits. Improved varieties had a lower phenotypic divergence than the ancestral parental lines.

2.4.3 A comparison of agro-morphological variation with other traits

The distinct advantage of studying morphological variation is that the phenotypic traits are often ecologically adaptive. Agro-morphological traits to an extent correlate with genotypic variation, and local differentiation or ecotypes (Schaal *et al.,* 1991; and Witcombe, 1999). Witcombe and Rao (1976) in their study on genecology of wheat collections of Nepal, found that collections differed for many agro-morphological characters and these differences were associated with altitude and geographical origin. Similar data on correspondence of geographical distribution and agro-morphological distances are available for many vegetable species in AVRDC (Chowdhury *et al.*, 2001).

A number of studies have been carried out to assess and characterize the germplasm diversity of different agricultural and forage crops using both agro-morphological, isozymes and DNA markers. Witcombe (1999) reviewed the results of a number of diversity studies that had measured both agro-morphological and DNA markers. He concluded that there was often a good correlation, but high correlation depended on how diverse were the genotypes that could often be the result of evaluating extensively diverse materials. Karihaloo (1996) used morphological variation, cytological variation, isozyme variations, and molecular variation (RAPD) to estimate the genetic diversity in the *Solanum melongena* complex. The results demonstrated that the level of diversity varied depending upon the type of marker used. The isozyme and RAPD were

found correlated and showed close genetic relationships even though an extensive morphological diversity existed among the tested samples.

2.4.4 Allozyme variation

The analysis of isozymes has been an efficient and predominant technique used in examining genetic variation in natural populations since the 1960s and is still a method chosen for certain applications. It is straight-forward, and allows an examination of allelic variation in genes encoding the soluble enzyme and provides an estimate of gene and genotypic frequencies within populations. Such data can be analyzed in numerous ways to measure genetic differentiation, and population subdivisions (Weir and Cockerham, 1984), genetic diversity (Nei, 1973) and gene flow (Slatkin, 1989). It has thus added new dimensions in genetic studies of many plant species and thereby provided insights into their population genetics (Nielson, 1985). Many reports have been documented concerning the methodologies, applications and limitations of isozyme electrophoretic spectra in cultivar identification and diversity studies (Almgard and Landegren, 1974; Brown, 1978; and Brown and Clegg 1983).

Isozymic variation provides a useful basis for the estimation of population structures and phylogenetic relationships, because the zymographic pattern directly reflects a particular gene system (Shahi, *et al.*, 1969). Isozyme polymorphism has received much attention in characterizing the Asian cultivated rice varieties and their wild progenitors. Glaszmann (1987) suggested a simplified starch gel electrophoresis of isozymes for the classification of rice varieties encoded on 5 diagnostic genes and compared it with the results of standard methods involving 15-21 loci encoding 10

enzymes. The resulting zymograms were identical for both methods and permitted the classification of 99% rice cultivars under the test. A study in peroxidase, acid phosphatase and esterase within and between species of *Oryza* revealed a large amount variation (Shahi *et al.,* 1969). Esterase isozymic variation has been found to be greatest in collections of rice (wild and cultivated species) in Nepal, Bhutan, Assam, Burma, Vietnam and Yunan in China. Nakagahra *et al.*, (1975) postulated these areas as the centre of genetic diversity for esterase isozymes in *Oryza sativa*. Gao *et al.,* (1999) found that there was a very low level of allozyme diversity within populations of *Oryza granulata*.

2.4.5 Molecular genetic variation

With the development of recombinant DNA technologies, it has become possible to examine variation in DNA sequences in any plant species. DNA variation is measured directly and therefore this method avoids environmental effects which can confound morphological evaluation of agronomic traits, and avoids the more biased allozyme estimates of genetic variation. Therefore molecular techniques are being used in many laboratories to undertake research on plant genetic resources for conservation and breeding. A brief introduction to the principles and procedures of some of these markers of relevance to variation in plant genetic resources and their variation is described on the internet (http://www.cgiar.org/ipgri/training). Technical innovations in molecular biology are occurring at a rapid pace and more than a dozen DNA-based marker systems have been developed. These have broadened the study of polymorphism to the level of DNA sequences that can be either of coding or non-coding, conservative or hyper-variable,

nuclear or organelle. These DNA sequences have different constraints. For example, ribosomal DNA shows little variation within a species (Schaal *et al.,* 1991), whereas hypervariable sequences show genotype-specific variation (Jeffreys *et al.,* 1985).

These techniques can be grouped into 3 Categories:

1: non PCR-based methods;

2: arbitrary or semi-arbitrary primed PCR-based methods; and

3: site targeted PCR methods (Karp *et al.*, 1997).

Therefore, these techniques can be either PCR-based, hybridization-based or both.

The choice of method depends on the type and level of diversity, availability of resources and consumables, and technical skill (Table 2.3).

Table 2.3. Comparison of molecular techniques used in relation to the conservation of genetic resources.

	RFLPs	RAPDs	AFLPs	SSRs
Principle of assay	Hybridisation, Fingerprinting with specific probes	Amplification of DNA fragments	Selective amplification of double-restricted fragments	Amplification of microsatellite (SSR) sequences
Category	Non PCR	Random PCR	Random PCR	Targeted PCR
Type of polymorphism	Point mutations, insertions, deletions, DNA rearrangements	Sequence differences	Point mutations, insertions, deletions, DNA rearrangements	Repeat length variations
Level of polymorphism	Medium	Medium	low	High
Abundance	High	High	Very high	Medium
Dominance	Codominant	Dominant	Dominant	Codominant
Repeatability	High	Low	Medium	High
Sequence information required	No	No	No	Yes
Automation possible	No	Yes	Yes	Yes
Level of skill required	High	Low	Medium to high	Medium
Cost	High	Low	Low to medium	Medium

2.4.5.1 Restriction Fragment Length polymorphism (RFLP)

The phenomenon of restriction fragment length polymorphism (RFLP) was first described in mutant strains of adenoviruses and later it was recognized as a powerful tool for constructing highly saturated linkage maps (Botstein *et al.* 1980). It is a non-PCR-based method and RFLP markers are co-dominant and mostly single copy. In RFLP analysis, the DNA (5-10μg) is digested with restriction enzymes, then fragments are separated by gel electrophoresis. Gels are blotted and hybridization is done with

florescently or radioactively labeled homologous probes. The resultant hybridized products are visualized using auto-radiography to identify the polymorphism among the individuals tested. The sites of digestion are determined by the presence or absence of a small sequence (4, 5, or 6 basepairs) of DNA that is specific to a particular restriction endonuclease. An endonuclease can only cut double stranded DNA wherever its specific recognition site occurs so any genetic polymorphism at this site or changes in the lengths of DNA sequences between restriction sites will create variation in fragment lengths. The resulting patterns of DNA fragments are compared to determine variation (Figure 2.2). RFLPs are widely used in genetic mapping and in the characterization and estimation of genetic diversity in eukaryotic species. The method may be targeted at analyzing nuclear or organellar (mtDNA, cpDNA) genomes. However, the technique is laborious, time consuming, technically difficult and is expensive.

2.4.5.2 Random Amplified Polymorphic DNA (RAPD)

Random amplified polymorphic DNA (William *et al.*, 1990; and Welsh and McClelland, 1990) is a PCR-based marker system, which involves the use of a single, usually 10-mer, arbitrary primers in a PCR reaction to amplify a DNA fragment over a million times in 1-2 hours.

In the polymerase chain reaction (PCR) a defined sequence of DNA is enzymatically synthesized *in vitro*. The reaction uses two oligonucleotide primers that hybridize the opposite strands (5' – 3') and flank the target genomic DNA sequence that is to be amplified. *Taq* DNA polymerase, a heat-stable DNA polymerase isolated from the thermophilic eubacterium *Thermus aquaticus* Bm, catalyses the extension of the DNA

fragment during amplification. PCR involves a series of repetitive cycles of template DNA denaturation, primer annealing and extension of the annealed primers by Taq. DNA polymerase results in a doubling of the concentration of template DNA in each cycle. It requires a target DNA template for amplification, one pair of primers complementary to the template region to be amplified, deoxynucleotide triphosphate (dNTP), a reaction buffer containing Mg^{2+} and a programmable heating block and suitable reaction vessels.

The 10-mer primers used in RAPDs bind throughout the genome wherever there are complementary sequences and anneal at low temperature. If two sites are on opposite strands within 2 kb of each other, the fragment between them will be amplified. As a result, a number of anonymous, but often reproducible, fragments up to a maximum of 2 kb are generated by amplification of the strand of template DNA and the primer. The resulting amplification products can then be visualized by ethidium bromide staining on agarose or polyacrylamide gels after electrophoresis (Figure 2.2). Because of several advantages such as a relatively unbiased portion of the genome sampled, simplicity of use, low cost, and the use of a small amount of template DNA, RAPD has widely been used in the detection of genetic diversity in closely related species, in different populations of a species, in individuals within populations, and for conservation genetics (Smith and Wayne, 1996; and Cruzan, 1998). It has successfully been employed in determining the genetic diversity in many species including rice (Yu and Nguyen, 1994; and Ge *et al.*, 1999), wheat (Fahima *et al.*, 1999), barley (Bustos *et al.,* 1998). However, the technique has some limitations. It is impossible to distinguish heterozygotes from homozygotes for the dominant allele, sometimes non-parental bands are seen in offspring of known parentage (Aman, 1997), and it is less reliable and reproducible compared to

most other methods. These disadvantages relate to the use of a low annealing temperature and the resulting reduction of primer-template specificity.

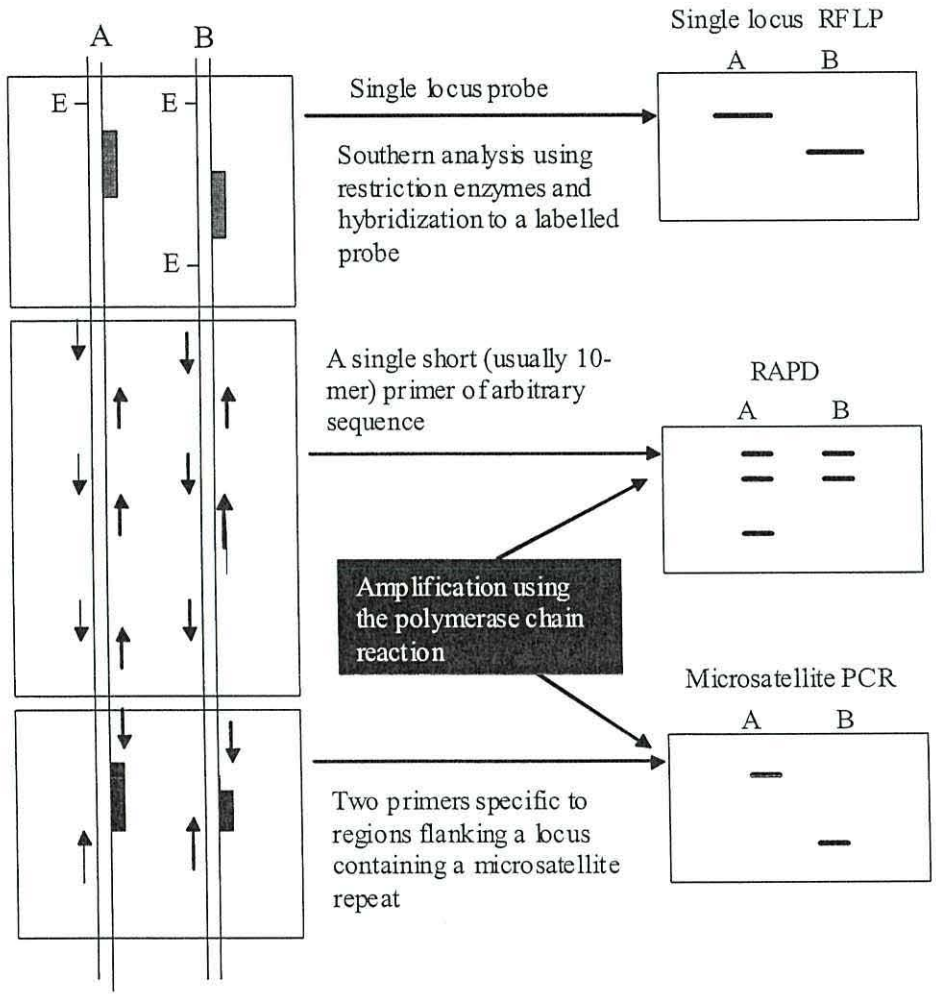


Figure 2.2: Diagrammatic sketch showing the processes by which molecular bands are produced using different molecular techniques (adapted from Newbury and Ford-Lloyd, 1997).

2.4.5.3 Amplified Fragment length Polymorphism (AFLP)

Amplified fragment length polymorphism (AFLP) analysis complements RFLP and RAPD marker systems as it combines restriction digestion and PCR. AFLPs are dominant markers equally applicable to all species and highly reproducible (Vos *et al.*, 1995) and provide an effective means of detecting several polymorphisms in a single assay thus increasing the possibility of identifying polymorphisms and expediting the construction of high-density linkage maps (Cho *et al.,* 1998). It involves the initial restriction digestion of the genomic template DNA with two specific enzymes: a 6-base cutter (rare, e.g. *EcoRI,*) and a 4-base cutter (frequent, e.g. *Mse*I). Specific, double-stranded adapters are then ligated (added) to the ends of the cut fragments to provide a known sequence for PCR primer annealing. It requires two rounds of PCR with specific primers. The first round involves the amplification of all fragments with non-selective primers. The amplified products are numerous, overlapping and would not be resolvable in a single gel (Karp *et al.*, 1997). The second round PCR amplifies only a subset of the fragments using selective primers, which have 1, 2 or 3 additional base pairs (A, G, C or T) at the 3' end of the adaptor sequence. In the second round, one of the selective primers must be radioactively, end labeled to visualize PCR products on denaturing polyacrylamide gels. Three kinds of fragments are produced. Type I are fragments with rare cutter ends only; Type II have one rare cutter end and one frequent cutter end; and Type III have two frequent cutter ends (Amin, 2002). The principle steps involved in AFLP are shown diagrammatically in Figure 2.3. The amplified products are separated on a polyacrylamide gel and visualized after exposure to X-ray film. The technique has been successfully applied in genetic mapping in a variety of plant species including rice

(Becker, *et al.,* 1995; Cho *et al.*, 1996; Mackill *et al.*, 1996; Pot *et al.,* 1996; Maheswaran *et al.,* 1997; Virk *et al.*, 1998 and Zhu *et al.*, 1998) and genetic diversity studies (Thomas *et al.*, 1995). Virk *et al.*, (2000) used the unmapped AFLP markers in diversity study of cultivated rice and revealed a pattern of diversity corresponding to that obtaining using other markers. However, the technique has some limitations. It is difficult to identify the same band across two different populations. Therefore, it is very difficult to use in comparative mapping. It is extremely difficult to detect co-dominant alleles and when mapping with AFLP, they often cluster close together (Price *et al.,* 2000).

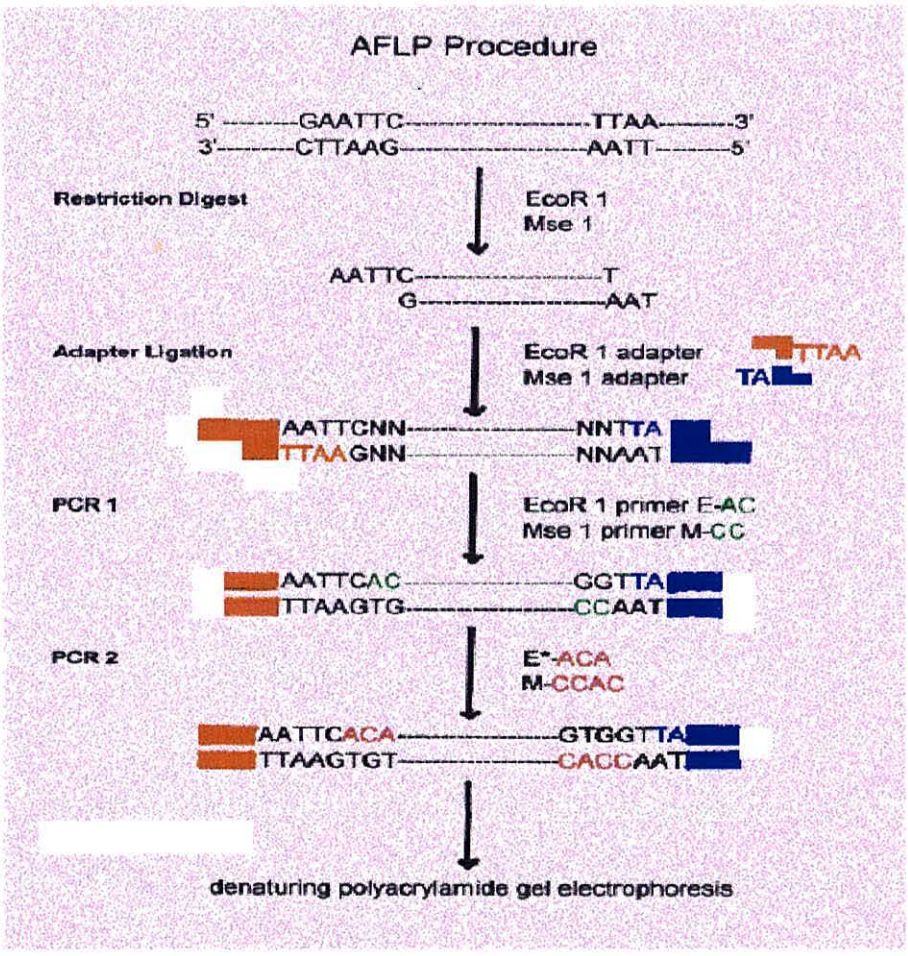


Figure 2.3: The principle steps involved in the AFLP protocol (Adapted from Internet).

2.4.5.4 Microsatellite marker (Simple sequence repeats SSRs)

Microsatellites (Litt and Luty, 1989) also called simple sequence repeats (SSRs) (Tautz, 1989 and Weber and May, 1989) are simple arrays of short nucleotide repeats from 1-6 base pairs per unit. Microsatellites were first studied in humans (Hamada *et al.,* 1982) but have now been found in a wide array of eukaryotic genomes of mammals (Love *et al.*, 1990; Serikawa *et al.*, 1992), birds (Cheng and Crittenden, 1994), fish (Estoup *et al.*, 1993), insects (Tautz and Ranz, 1984), and many monocot and dicot plant species (Wang *et al.*, 1994). Typically they may be dinucleotides e.g. $(AC)_n$, $(AG)_n$, $(AT)_n$, $(GA)_n$, $(GT)_n$; trinucleotides e.g. $(GGC)_n$, $(AAT)_n$, $(GAA)_n$; and tetranucleotides e.g. $(TATG)_n$, $(CCCG)_n$, where n is the number of repeats. The di-, tri-, and tetra-nucleotide repeats differ in how frequently they occur (Table 2.5). They are hypervariable and ubiquitously distributed throughout eukaryotic genomes. They are valuable as genetic markers because they are codominant, usually at a single locus, and easy to target and economic to use and are PCR based assay.

McCouch *et al.*, (1997) estimated that there are 5,700-10,000 microsatellite sequences distributed throughout the rice genome. Over 500 microsatellite markers have been developed for rice and their chromosomal location and level of polymorphism have been determined (Temnykh, *et al.*, 2000; and Blair *et al.*, 2002). These simple sequence repeats (SSRs) in rice were predominantly poly (GA) motifs isolated from two genomic libraries (Panaud *et. al.,* 1996; Chen *et al.*, 1997; Wu and Tanksley, 1993 and Akagi *et al.,* 1996). Microsatellites (SSRs) carry a 20-50 repeats, so they provide a high level of informative allelic diversity (Morgante and Oliveri, 1993) and are more variable than RFLPs (Akagi *et al.*, 1996). In general, SSRs with more repeats, and with dinucleotide

repeats, were found to be more polymorphic (Weber, 1990; Innan *et al*, 1997; and Schung *et al.*, 1998). It was also observed to be true in humans (Chakraborty *et al.*, 1997).

SSRs are analysed by PCR amplification of a short genomic region containing the repeated sequence, and size estimation of the repeat length referred as simple sequence length polymorphisms (SSLPs) is detected by gel separation on high resolution agarose or polyacrylamide gels staining either with ethidium bromide or silver, radio-labelled primers or florescence labelled primers (Figure 2.2). The base sequence information of the flanking regions of a SSR locus is determined by sequencing. The uniqueness and the conservation of the flanking regions provide choices to design and synthesize primers necessary for detection of microsatellite polymorphisms. A number of studies in rice have been undertaken to identify microsatellites (Akagi *et al*, 1996; Panaud *et al*, 1997; Wu and Tanksley, 1993; and Zhao and Kochert, 1992) and have made available the microsatellite markers to detect the polymorphism.

Highly saturated genetic maps based on microsatellite DNA markers in human and mouse genomes have been constructed (Dib *et al.*, 1996; and Dietrich *et al.*, 1996). Following these discoveries, these abundant and highly polymorphic markers were applied to several plant species, and the microsatellites made publicly available: rice (Akagi *et al.*, 1996; Panaud *et al.*, 1996; Wu and Tanksley, 1993; and Zhao and Kochert, 1992), barley (Becker and Heun, 1995), wheat (Roder *et al.*, 1995), maize (Senior and Heun, 1993), soybean (Akkaya *et al.*, 1995), tomato (Broun and Tanksley, 1996), grapevines (Thomas and Scott, 1993), sunflower (Brunel, 1994) and *Brassica* spp. (Bell and Ecker, 1994; Poulsen *et al.*, 1993). A map consisting of 121 microsatellite loci with a

genome wide coverage has been published for rice (Chen *et. al.,* 1997). A map consisting of total 312 microsatellite markers including 124 previously reported and 188 newly developed microsatellite loci has recently been developed, which is sufficient to be useful for both basic genetic studies and breeding applications (Temnykh *et al.*, 2000).

Zhao and Kochert (1992); Wu and Tanksley (1993); and Akagi *et al.* (1996) evaluated the rice microsatellites (SSRs) for their efficacy as genetic markers in terms of abundance, polymorphism and genome distribution (Table 2.4). $(CCG)_n$ was abundant, polymorphic and was found throughout the rice genome (Zhao and Kochert, 1992). $(GA)_n$ and $(GT)_n$ were moderately abundant in the rice genome (Wu and Tanksley, 1993).

Cho *et al.,* (2000) compared microsatellite markers, derived from genomic library screening, and expressed sequence tags (ESTs) using 14 diverse rice accessions. Microsatellites derived from genomic libraries detected a higher level of polymorphism and genetic variability than that of the ESTs. The highest overall degree of genetic diversity was seen in GA- and GT-dinucleotide containing microsatellites of genomic library origin. Markers with CCG- or CAG-trinucleotide repeats were found to be the most conserved.

Table 2.4: Frequencies of microsatellites (SSRs) in the rice sequences registered in the database (Akagi *et al.*, 1996).

2 Base		3 Base		4 Base	
Repeat units	**Number of sequences**	**Repeat units**	**Number of sequences**	**Repeat units**	**Number of sequences**
GA/CT	55	CCG/GCC	137	GATA/CTAT	2
AT/TA	5	GAG/CTC	67	GTAG/CATC	2
GT/CA	2	CTT/GAA	26	CTTT/GAAA	2
		CTG/GAC	25		
		ACG/TGC	20		
		TGG/ACC	18		
		ATC/TAG	3		
		CAG/GTC	2		
		TTG/AAC	2		
		ATT/TAA	1		

Diversity of microsatellite markers in rice varied according to the populations and combination of the markers selected for analysis. Garland *et al.*, (1999) investigated the microsatellite (SSR) polymorphisms in a range of rice cultivars of commercial and breeding interest to the Australian rice breeding programme. In this study, they observed 11 to 23 alleles for OSR prefixed markers (*Oryza* Sequence Repeats) and 5-9 for RM prefixed markers (Rice microsatellite). Akagi *et al.* (1997) found 5 to 10 alleles among 59 closely related *japonica* rice varieties.

2.4.5.5 Other molecular marker techniques

In addition to these marker systems, many other molecular marker techniques been developed that are applicable in crop breeding and genetic diversity study. Some of these that have been used in measuring genetic diversity are briefly described below.

***SCAR (Sequence-Characterised Amplified Regions)*:** The technique can be applied to both RAPDs and AFLPs. Markers are derived by the specific amplification of individual RAPD or AFLP bands. The technique identifies longer specific primers based on the DNA sequence of the RAPD or AFLP bands. These longer primers may reveal size polymorphism and hence can be co-dominant markers.

ISSR (Inter-Simple Sequence Repeats): ISSR is a new technique developed to solve the limitations of RAPD technique. It involves PCR amplification of DNA using a single primer composed of a microsatellite sequence anchored at the 3' end or 5' end by 2-4 arbitrary nucleotide. ISSR amplification reveals a much larger number of polymorphic

fragments per primer and a higher reproducibility than RAPD, and does not require a prior knowledge of DNA sequence for primer design (Yang *et al.,* 1994). The technique has been successfully employed in the determination of genetic variation in dent corn and pop corn (Kantety, *et al.,* 1995) and has important applications in the evaluation of genetic diversity. Qian *et al.*, (2001) detected the genetic variation within and among populations of a wild rice *Oryza granulata* from China and compared the RAPD and ISSR markers. It is a useful alternative technique to single-locus (SSR) or hybridization-based methods (Zietkiewicz *et al.*, 1994; and Goodwin *et al.*, 1997).

CAPS (Cleaved Amplified Polymorphic Sequences): In this technique, partial DNA sequence information of the locus of interest is used to amplify a segment of DNA at the locus from several different individuals and the amplified products (bands) are digested with a number of restriction enzymes and the products directly visualized on the agarose gel by ethidium bromoide staining and identify RFLPs between individuals (Bhat, 1996; and Karp *et al.,* 1997).

SNP (Single Nucleotide Polymorphism)

The field is fast evolving and new developments are continually emerging, e.g., SNP genotyping (single nucleotide polymorphism), which was developed to detect DNA polymorphism linked to disease in humans (Illig, 2002). It is a direct analysis of sequence difference between the individuals. The discovery of human SNP has stimulated progress in determining SNPs in plants. For example, genomic sequencing by Cereon in *Arabidopsis* has produced a very large SNP collection available to academic researchers

(Rafalski, 2002). Likewise, International Rice Genome Sequencing Project (IRGSP), Beijing Genomics institute (BGI), China and two public-private partnerships in genomics in Japan: Monsanto and Syngenta have made available the information on rice genome sequencing focused on the temperate *japonica* cultivar Nipponbare and on the tropical *indica* cultivar 93-11, a major variety grown China and Southeast Asia (Buell, 2002). SNPs between 2 varieties can be detected using the recent technology Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS).

2.5 Application of molecular markers

Molecular markers have been extensively and usefully employed in the support of long-term germplasm conservation and utilization. Hodgkin *et al.* (2001) reviewed the array of molecular markers that have been used in conservation and use of plant genetic resources. Ford-Lloyd *et al.*, (1997) also pointed out the importance and applications of molecular markers in management and organization of genetic resources in gene bank. They can be used to test if accessions or cultivars are true to type, and can be used to detect duplicate accessions, seed mixtures, inadvertent out-crossings and genetic drift. Molecular markers provide genetic information of direct value in key areas of both *in situ* and *ex situ* conservation. These are the determination of: accession identity; the degree of similarity among individuals in an accession or between accessions within a collection; the genetic structure of individuals, accessions, populations and species; and the detection of a particular allele of an accession, or an *in situ* population. This information will help in locating the population to be conserved, and for the management and use of the diversity (Karp *et al.*, 1997).

The genebank at the International Rice research Institute (IRRI) has a collection of more than 95 000 samples of rice germplasm and rice is one of the few crops that has such a huge and diverse collection in genebank (Virk *et al.*, 2000). The collection comprises landraces of *O. sativa* and *O. glaberrima*, breeding lines, commercial varieties of Asian cultivated rice, and all 20 wild species of the genus *Oryza* (Virk *et al.*, 1996). In rice, the diversity indices and pattern of diversity in sets of germplasm have been assessed using isozymes (Glaszmann, 1987); RFLPs (Ford-Lloyd and Newbury, 1994; Zhang *et al.*, (1992), RAPDs (Virk *et al.*, 1995; Martin *et al.*, 1997), and AFLPs (Mackill *et al.*, 1996; and Virk *et al.*, 1998).

2.5.1 SSRs

Microsatellites are considered appropriate for variety identification because of their ability to detect large numbers of discrete alleles repeatedly, accurately, and efficiently (Smith and Helentjaris, 1996). Wu and Tanksley (1993), Yang *et al.*, (1994), and Paunaud *et al.*, (1996) in their independent allelic diversity studies on cultivated rice varieties detected up to 25 alleles at a single microsatellite (SSR) locus and a high (0.69) polymorphism information content (PIC). The high levels of the genomic coverage and allelic diversity provided by SSRs in rice suggested that it is useful for variety identification (McCouch, *et al.*, 1997). The genetic profiles produced by a group of microsatellite (SSRs) markers can be used, together with pedigree and performance data, to document ownership and protect intellectual property rights. Ramakrishna *et al.* (1994) demonstrated that SSR-derived DNA fingerprints were ideally suited for the

identification of rice genotypes as all the primers they used detected a high level polymorphism.

Microsatellites (SSRs) have also been ideal for characterising genetic diversity in cultivated rice at both the inter-varietal (Xiao *et al.,* 1996) and intra-varietal (Olufowote *et al.,* 1997) levels. In addition, the markers have also been successful in amplifying the loci in a range of closely related non-domesticated rice species (Moncada *et al*, 2001; and Ishii and McCouch 2000). Yang *et al.,* (1994) employed microsatellite analysis to assess the extent of genetic variation in landraces and to evaluate the amount of genetic diversity that has been incorporated in modern elite rice cultivars. Similarly, in genetic heterogeneity among Australian breeding lines of rice, Garland *et al.*, (1999) found microsatellite as useful for variety identification and assessment of genetic relationships. Sebastian *et al.* (1998) and Sun *et al.* (2001) successfully used microsatellite markers to assess the diversity of farmers' rice varieties, wild rice and cultivated rice. Farmers' varieties and the wild rice populations exhibited high heterogeneity within and between the populations. However, it was observed low in cultivated rice as a result of natural and human selection. Similarly, Ni, *et al.* (2002) evaluated the genetic diversity and determined the genetic identity in a diverse collection of rice (*Oryza sativa*) including 8 modern rice cultivars from US, two subspecies *indica* and *japonica* and two wild species accessions of *O. rufipogon* and *O. nivara* using 111 microsatellite (SSR) markers distributed over the whole rice genome.

Furthermore, the fluorescent labelling of markers using different dye colours, automated detection on DNA sequencers and multiplexing of markers have increased the throughputs in diversity studies and allowed efficient genotyping (Ziegle *et al.,* 1992).

Blair *et al.* (2002) tested four multiplex panels (with 8-12 primer pairs) of fluorescent-labeled rice microsatellite markers (primers) on 72 accessions of cultivated rice represented a diversity of ecotypes and of isozyme groups and found them useful for fingerprinting and for clustering rice varieties.

2.5.2 AFLPs

AFLPs have also successfully been employed for DNA fingerprinting in genetic mapping and biodiversity studies in many plant species including rice (Cho *et al.*, 1997; Zhu, 1996 and Zhu *et al.*, 1998). Zhu *et al.* (1999) used map-based DNA fingerprinting with AFLP in scanning the rice genome using two mapping populations of *indica* and *japonica* and ten AFLP primer combinations. They pointed out the importance of genome scanning and assessed the variation between rice populations at the chromosomal level using DNA Fingerprint Linkage Blocks (DFLBs). The importance of this powerful technique in application to genome breeding and utilization of core collections of germplasm was discussed (Zhu *et al*, 1999).

2.5.3 Combination of the markers

Genetic variation within and between five populations of wild rice (*Oryza granulata*) from two regions of China was investigated using RAPD and ISSR markers. The study showed a low level of diversity among the populations of wild rice for both markers. However, ISSR was found to be superior to RAPD in terms of polymorphism detected and amplification reproducibility (Qian *et al*, 2001). Parsons *et al.* 1997 likewise employed RAPD and ISSR markers in diversity assessment of cultivated rice (*Oryza*

sativa) from Bangladesh and Bhutan. A contrasting genetic diversity was revealed and ISSR produced a slightely higher level of polymorphism than RAPD.

In a RFLP analysis, *indica* rice has been shown to be genetically more diverse than *japonica* rice, as a result of extensive genetic differentiation between major portions of the rice genome (Zhang *et al.,* 1992). Olufowote, *et al.* (1997) in a comparative evaluation of within cultivar variation in IR8 (a modern variety) and a landrace variety named *Latisail* with SSR and RFLP, distinguished all the accessions of landrace from each other with all 12 SSR markers tested and 4 out of 12 RFLP markers. It showed the efficiency of SSR for within accession/cultivar variation and detection of the level of heterogeneity in a landrace population.

2.6 Analysis of genetic diversity using morphological and molecular data

Effective use and conservation of germplasm is based on the available information on genetic diversity, genetic knowledge of the traits and their relative contribution to variability. So it is imperative to understand the types of the data generated by these different techniques and the methods analyzing the generated data.

The analysis of genetic relationships starts with the construction of a matrix specifying the state of an attribute for a sample. Commonly measured attributes in diversity analysis of genotypes are continuous phenotypic variables such as agro-morphological traits (based on measurement - maturity, height, phenology etc.); discrete phenotypic variables (usually multi-state - leaf colour, grain colour, panicle type etc.); and discrete genetic marker variables (normally binary - absence / presence). Similarities or distances between two samples are calculated on the basis of the differences between

them for the set of the attributes and construct a sample x sample matrix. Taxonomic distance, squared Euclidean distance and the Mahalanobis squared distances are some common indices that have been widely used in agro-morphological diversity data (Ezeaku *et al.,* 1999; Martinello *et al.*, 2001; Tatineni *et al.,* 1996; Ayana and Bekele, 1999). Likewise, Jaccard's similarity coefficient, simple matching coefficient and Nei's genetic distances are the common genetic indices used in analysis of genetic marker data (Autrique *et al.*, 1996; Parsons *et al.*, 1996; Riek *et al.*, 2001). These indices are based on the frequencies of alleles or states of traits and their sharing among the samples.

Two multivariate statistical methods (phenetic classificatory and ordination) have been widely used to study the patterns of genetic diversity expressed in multidimensional space defined by the marker character. These multivariate techniques have been used alone or in combination (Rohlf, 1992) to study various aspects of diversity within crop germplasm. Principal component analysis operates on correlations among all of the variables and the resultant genetic distances among the samples can be reflected in a 2 or 3 dimensional scatter plot (Manly, 1994 and Karp *et al.*, 1997). Another approach is the classification (clustering) based on the similarity, Euclidean distance or genetic distance between pairs of individuals or samples which sort out the coefficient values using one of several algorithms, e.g. single linkage or unweighted pair-group method of arithmetic averages (UPGMA), and the relationships between the samples will be depicted in the resulting dendrogram (Newbury and Ford-Lloyd, 1997). Both the approaches are based on the aggregation of the most genetically similar samples and reveal the relationships or distances among the samples. These techniques have been used to visualize diversity in germplasm across a range of characters or bands and to separate geographical or

ecogeographic pattern of diversity (Francisco-Ortega *et al.*, 1992; Crossa *et al.*, 1995; Virk *et al.*, 1995; and Weising *et al.*, 1995).

CHAPTER III

MATERIALS AND METHODS

This chapter describes the major materials and methods used for the rice diversity studies reported in the following chapters and only those are described that are common to more than one chapter. All of this work was part of the IPGRI project on "Strengthening the scientific basis of *in situ* conservation of agricultural biodiversity on-farm" in which the Nepal Agriculture Research Council (NARC), the Local Initiatives for Biodiversity, Research and Development (LIBIRD) and the International Plant Genetic Resources Institute (IPGRI) were partners. Some of this work was supported by the UK Department for International Development Plant Sciences Research Programme (DFID-PSRP).

3.1 Genetic materials

3.1.1 Collection of rice landraces for variability studies

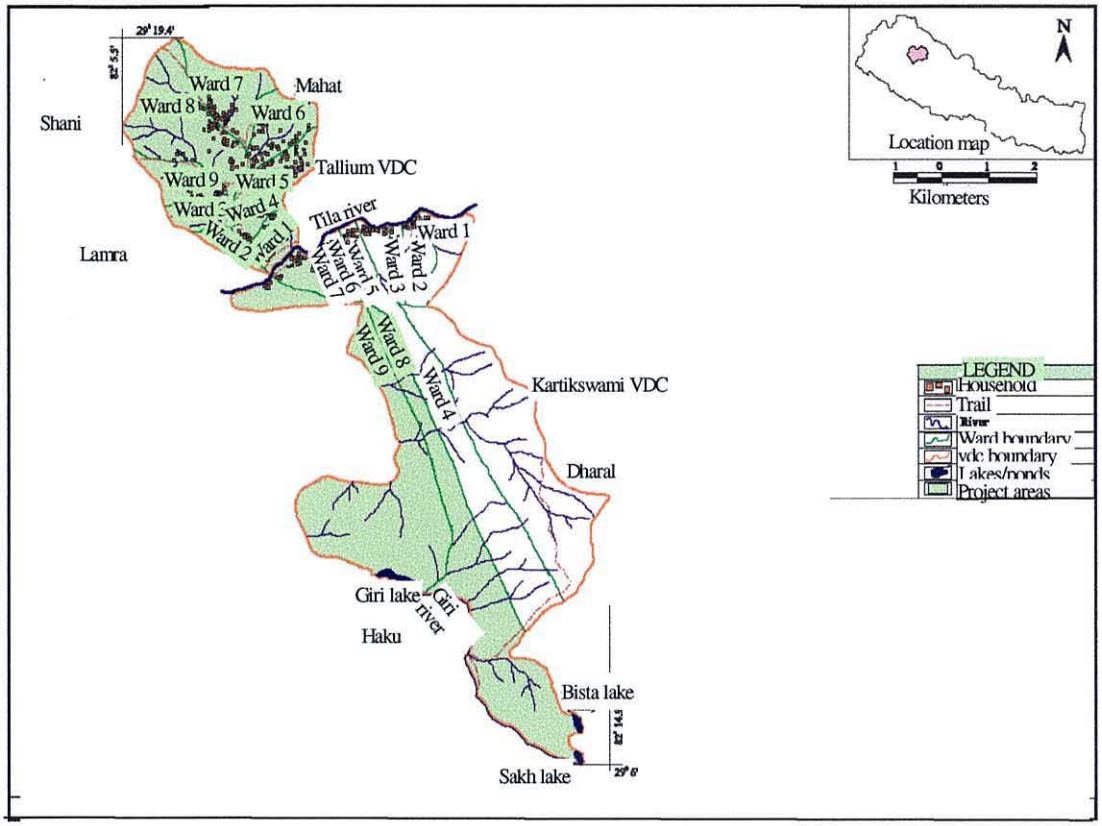
This work was carried out at three locations, termed 'eco-sites' in the IPGRI *in situ* project. They are in three districts of Nepal: Jumla, Kaski and Bara. These names are used for these three ecosites of the project throughout the thesis. The Jumla ecosite consisted of the Tallium and Kartikswami Village Development Committees (VDC), The Kaski ecosite of the Begnas and Rupakot VDCs, and the Bara ecosite of the Kachorwa VDC, (Figure 3.1a-c). These ecosites represent three major physiographic zones of the country and represent three different rice ecosystems: Jumla (high-hill), Kaski (mid-hill) and Bara (*Terai*-lowland). These terms for ecosystems have been used throughout the thesis, as appropriate, to describe the environments of the study sites. General descriptions of these sites are given in Appendix 3.1.

Plant materials for the study were the complete range of rice diversity from each site according to the names given to the landraces by farmers. A named landrace was taken as one farmers' unit of diversity (FUD) and a sample of seed from one household (HH) was taken as one accession. Seed samples were collected in 1998 during baseline and related survey studies, called *in situ* characterization by the project, in the three ecosites. Any population / cultivar / variety / landrace referred to in this study is a seed lot of a farmer or a farmers' unit of diversity (FUD) with a given name, which could be genetically the same or genetically different to another seed lot with the same name collected from a different farmer. In order to identify the farmer and his or her seed material, each collection was given an accession number. A total of 632 accessions of rice were used for the agro-morphological study, 147 from the Jumla ecosite, 291 from the Kaski ecosite, and 194 from the Bara ecosite. These accessions represented 10 different named landraces from Jumla, 75 from Kaski and 42 from Bara (Table 3.1a-c). Improved rice varieties recommended in each of the three agro-zones were included in the study as checks. These were: Jingling for Jumla; Ghaiya 2 and Sabitri for Bara; and Masuli for Kaski.

Table 3.1a: Rice landraces from Jumla, the high-hill eco-site.

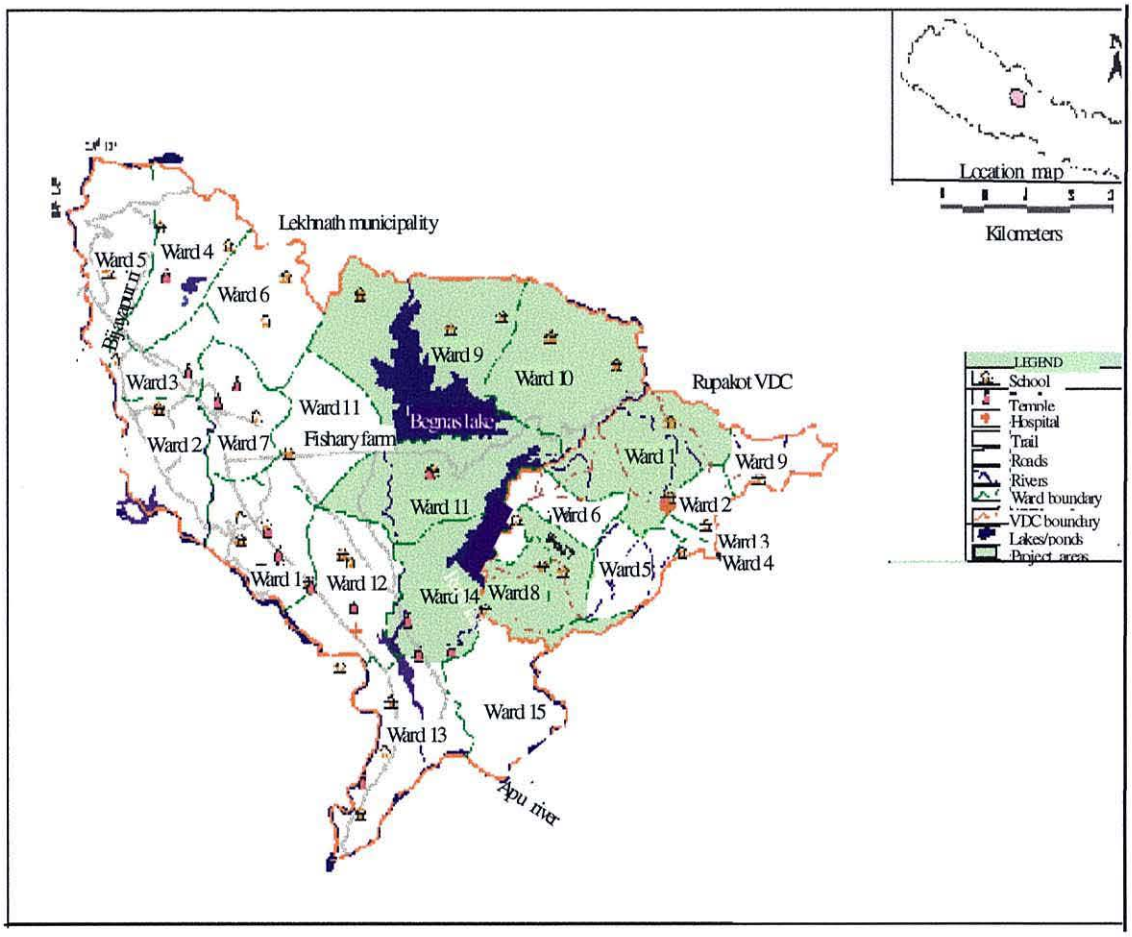
S.No.	Landraces	No. of accessions	Households grown
1	Seto Marshi	51	Large area† and many HH‡
2	Rato Marshi	52	Large area and many HH
3	Kalo Marshi	21	Large area and many HH
4	Mehele	10	Small area and few HH
5	Dhan	3	Small area and few HH
6	Palte Dhan	1	Small area and few HH
7	Darime	6	Small area and few HH
8	Rato Dhan	1	Small area and few HH
9	Seto Dhan (Seto Seed)	1	Small area and few HH
10	Jumli Rato Marshi	1	Small area and few HH
11	Jingling (MV)	1	Check variety
Total		148	

† = proportion of rice area in the ecosite; ‡HH = households



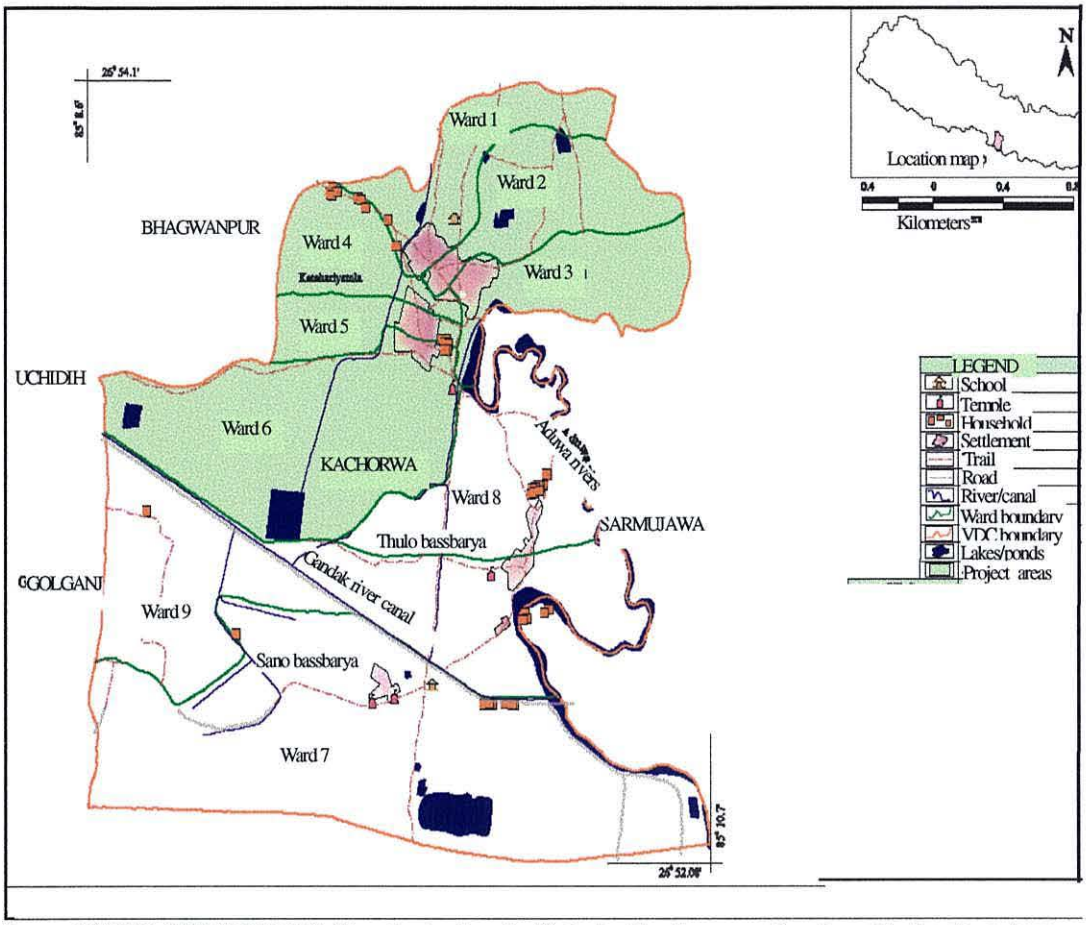
Source: NARC/LI-BIRD/IPGRI, 2001. Strengthening the scientific basis of *in situ* conservation agro-biodiversity on farm: Nepal contribution to the global project.

Figure 3.1a: Site map of Tallium and Kartikswami ecosite, Jumla District, Nepal.



Source: NARC/LI-BIRD/IPGRI, 2001. Strengthening the scientific basis of *in situ* conservation of agro-biodiversity on farm, Nepal's contribution to the global project.

Figure 3.1b: Site map of Begnas ecosite, Kaski district, Nepal.



Source: NARC/LI-BIRD/IPGRI, 2001. Strengthening the scientific basis of *in situ* conservation of agro-biodiversity on farm; Nepal's contribution to the global project.

Figure 3.1c: Site map of Kachorwa ecosite, Bara district, Nepal.

Table 3.1b: Rice landraces from Kaski, the mid-hill eco-site.

S.No.	Landraces	No. of accessions	Households grown
1	Kalo Jhinuwa	8	Small area† and few HH‡
2	Panhele Jhinuwa	3	Small area and few HH
3	Jhinuwa	2	Small area and few HH
4	Tunde Jhinuwa	2	Small area and few HH
5	Pakhe Jhinuwa	2	Small area and few HH
6	Lamcho Jhinuwa	1	Small area and few HH
7	Seto Jhinuwa	1	Small area and few HH
8	Masinho Jhinuwa	1	Small area and few HH
9	Tarkaya Jhinuwa	1	Small area and few HH
10	Juya Jhinuwa	1	Small area and few HH
11	Andheri Jhinuwa	1	Small area and few HH
12	Masino Dhade Jhinuwa	1	Small area and few HH
13	Kathe Gurdi	13	Large area and many HH
14	Lahare Gurdi	9	Small area and few HH
15	Thulo Gurdi	4	Large area and many HH
16	Seto Gurdi	3	Small area and few HH
17	Gajale Gurdi (Masino Lahare Gurdi)	2	Small area and few HH
18	Sano Gurdi	2	Small area and few HH
19	Gurdi	1	Large area and few HH
20	Thulo Kalo Gurdi	1	Small area and few HH
21	Bayarni	12	Small area and many HH
22	Kalo Bayarni	2	Small area and few HH
23	Seto Bayarni	2	Small area and few HH
24	Gajale Bayarni	1	Small area and few HH
25	Juya Bayarni	1	Small area and few HH
26	Seto Anadi	15	Small area and many HH
27	Rato Anadi	15	Small area and many HH
28	Sano Anadi	1	Small area and few HH
29	Dudhe Anadi	1	Small area and few HH
30	Madhese	12	Large area and many HH
31	Thulo Madhese	3	Small area and few HH
32	Naulo Madhese	2	Small area and few HH
33	Sano Madhese	1	Large area and few HH
34	Jarneli	13	Small area and many HH
35	Pakhe Jarneli	8	Small area and few HH
36	Tunde	8	Small area and few HH
37	Pakhe Tunde	1	Small area and few HH
38	Ramani	7	Small area and few HH
39	Pakhe Ramani	1	Small area and few HH
40	Kartike Marshi	1	Small area and few HH
41	Panhelo Marshi	1	Small area and few HH
42	Seto Marshi	1	Small area and few HH
43	Chiniya Marshi	1	Small area and few HH
44	Aanpjhutte	2	Small area and few HH
45	Sano Aampjhutte	1	Small area and few HH
46	Pakhe Gauriya	1	Small area and few HH
47	Gauriya	1	Small area and few HH
48	Ekle	15	Large area and many HH
49	Manasara	15	Large area and many HH
50	Jetho Budho	14	Large area and many HH
51	Aanga	10	Small area and many HH
52	Panhele	10	Large area and many HH

53	Naltumme	8	Small area and few HH
54	Biramphool	8	Small area and few HH
55	Basmati	7	Small area and few HH
56	Chobo	6	Small area and few HH
57	Palungtare	2	Small area and few HH
58	Jyamdikhole	2	Small area and few HH
59	Manamuri	2	Small area and few HH
60	Rate	2	Small area and few HH
61	Krishna bhau	2	Small area and few HH
62	Bhayare	2	Small area and few HH
63	Thapachini	1	Small area and few HH
64	Bhatte	1	Small area and few HH
65	Bale	1	Small area and few HH
66	Makai Khole	1	Small area and few HH
67	Dhabe Gaurama	1	Small area and few HH
68	Barmeli	1	Small area and few HH
69	Jadan	1	Small area and few HH
70	Masino	1	Small area and few HH
71	Battisara	1	Small area and few HH
72	Karna Jira	1	Small area and few HH
73	Tulasi	1	Small area and few HH
74	Pani Barmeli	1	Small area and few HH
75	Chhote	1	Small area and few HH
76	Masuli (MV)	1	Check variety
Total		292	

† = proportion of rice area in the ecosite; ‡ HH = households

Table 3.1c: Rice landraces from Bara, the *Terai* eco-site.

S.No	Landraces	No. of accessions	Households grown
1	Silhat	8	NA†
2	White Silhat	1	NA
3	Sathi	10	Small area‡ and many HH§
4	Karma	10	Small area and many HH
5	Basmati	14	Small area and many HH
6	Bhadaiya Basmati	4	NA
7	Sarihan Basmati	2	NA
8	Lajhi	4	NA
9	Dipahi	2	Small area and few HH
10	Mutmur	9	Large area and many HH
11	Muturi	6	Large area and many HH
12	Sotawa	10	Large area and many HH
13	Lalka Farm (Rato Faram, Lal Faram)	8	Small area and few HH
14	Aanga	9	Small area and few HH
15	Nakhisaro	9	Large area and many HH
16	Laltenger	7	Small area and few HH
17	Dudhisaro	3	Small area and few HH
18	Batí	7	Small area and few HH
19	Sikichan	3	NA
20	Manasara	2	Large area and many HH
21	Kataush	3	Small area and few HH
22	Rajala	5	NA
23	Khera	7	Small area and few HH
24	Rango	5	Small area and few HH
25	Budhidayan	2	Small area and few HH

26	Kariya Kamodh	1	NA
27	Madhumala	2	NA
28	Pakhad	4	NA
29	Gajargaul	2	Small area and few HH
30	Hattijhulan	5	Small area and few HH
31	Ujarka Faram	2	NA
32	Lal Khera	1	Small area and few HH
33	Seto Khera	2	Small area and few HH
34	Amaghaunch	2	Small area and few HH
35	Adalat	2	Small area and few HH
36	Brahmabhushi	1	Small area and few HH
37	Balamsar	1	Small area and few HH
38	Dudhraj	4	Small area and few HH
39	Harinkher	1	Small area and few HH
40	Chhataraj	2	Small area and few HH
41	Mansari	6	Small area and few HH
42	Sokan	6	Small area and few HH
43	Ghaiya 2 (MV)	1	Check – early rice
44	Sabitri (MV)	1	Check – normal rice
Total		196	

†NA = Not available; ‡ = proportion of rice area in the ecosite; § HH= households

3.1.2 Landraces for molecular variability study

A random stratifying sampling procedure was used to select 21 accessions from Jumla, 24 accessions from Kaski, and 25 accessions from Bara. In total, 10 differently named landraces from each of these sites were included and the proportion of accessions representing each named variety was proportional to the frequency of the four categories, shown in Tables 3.1a-c, for the area under cultivation and the number of households growing a particular named variety (Rana *et al.*, 2000a,b,c). These accessions were sampled from the complete sets of rice accessions for the agro-morphological study using a table of random numbers (Gomez and Gomez, 1976). These populations represented most of the rice growing environments found in these sites in the Baseline Survey Report of the IPGRI *in situ* conservation project (2000). Each accession consisted of bulk seeds of a particular rice population from farmers' grain stores. Three modern varieties, namely Kalinga III, IR36 and IR64 were included as control varieties throughout the SSR (microsatellite) diversity studies (Table 3.2). These three varieties were selected for the

study as they have been used in molecular marker diversity studies, or as parents of PPB-programme varieties studied in this thesis.

Table 3.2: Landraces and check modern varieties of rice included in the molecular diversity studies.

S.No.	Landrace/Variety	No. of accessions	Origin	Ecosite	Households grown
1	Seto Marshi	4	Nepal	Jumla	Large/many HH†
2	Rato Marshi	4	Nepal	Jumla	Large/many HH
3	Kalo Marshi	3	Nepal	Jumla	Large/many HH
4	Mehele	2	Nepal	Jumla	Small/few HH
5	Darime	2	Nepal	Jumla	Small/few HH
6	Dhan	2	Nepal	Jumla	Small/few HH
7	Rato Dhan	1	Nepal	Jumla	Small/few HH
8	Seto Dhan	1	Nepal	Jumla	Small/few HH
9	Palte Dhan	1	Nepal	Jumla	Small/few HH
10	Jumli Rato Marshi	1	Nepal	Jumla	Small/few HH
11	Kathe Gurdi	4	Nepal	Kaski	Large/many HH
12	Jetho Budho	4	Nepal	Kaski	Large/many HH
13	Gurdi	1	Nepal	Kaski	Large/Few HH
14	Rato Anadi	3	Nepal	Kaski	Small/Few HH
15	Aanga	2	Nepal	Kaski	Large/Few HH
16	Ramani	2	Nepal	Kaski	Small/Few HH
17	Seto Gurdi	2	Nepal	Kaski	Small/Few HH
18	Jhinuwa	2	Nepal	Kaski	Small/Few HH
19	Tunde Jhinuwa	2	Nepal	Kaski	Small/Few HH
20	Kalo Bayarni	2	Nepal	Kaski	Small/Few HH
21	Mutmur	3	Nepal	Bara	Large/many HH
22	Nakhisaro	3	Nepal	Bara	Large/many HH
23	Sathi	2	Nepal	Bara	Small/Many HH
24	Sokan	2	Nepal	Bara	Small/Few HH
25	Mansara	3	Nepal	Bara	Large/many HH
26	Basmati	3	Nepal	Bara	Large/many HH
27	Karma	3	Nepal	Bara	Small/Many HH
28	Lajhi	3	Nepal	Bara	Small/Many HH
29	Dudhraj	1	Nepal	Bara	Small/Few HH
30	Laltenger	2	Nepal	Bara	Small/Few HH
31	Kalinga III (MV) ‡	1	CRRI,§ Cuttack		Check variety
32	IR64 (MV)	1	IRRI,¶ Philippines		Check variety
33	IR36 (MV)	1	IRRI, ¶ Philippines		Check variety

† HH = Households; ‡ MV = Modern variety; § CRRI = Central Rice Research Institute, Cuttack; ¶ IRRI = International Rice Research Institute, Los Bannos, Philippines.

3.1.3 PPB bulk populations/varieties for molecular diversity studies

3.1.3.1 LI-BIRD

Seven morphologically different PPB bulk populations selected by the farmers in Chitwan, Nepal, under a research project: Participatory Plant breeding in rice, Nepal, executed by Li-BIRD and CAZS and funded by the UK Department for International Development (DFID) Plant Sciences Research Programme (PSRP) were used in this study (Table 3.3) (Witcombe *et al.*, 2001). All seven PPB bulk populations were derived from advanced generations of a single cross between Kalinga III and IR64. Kalinga III is early maturing and tall. It has a lower yield potential than IR64 but gives yield under low input conditions. It has a tendency to lodge because of weak straw. IR64 is later maturing, dwarf, high yielding and lodging resistant. It has multiple pest and disease resistance. These bulk populations were generated from a modified bulk population method. At the F_4 generation the bulk population from the cross was grouped into 6 bulks based on their plant height (tall or dwarf) and maturity classes (early <100 days seed to seed; medium 110-125 days; and late >125 days) and were named as early tall (ET); early dwarf (ED); medium dwarf (MD); medium tall (MT); late dwarf (LD); and late tall (LT) (Witcombe *et al.*, 2001). MT group was further divided into four based on the variability observed again in plant height and the maturity. These were MT1 (earlier shorter); MT2 (earlier taller); MT3 (later shorter); and MT4 (later taller). Only the satisfactory bulks which were accepted by the farmers, were included in this study. These were ET, MD, MT1, MT2 and MT3. During the course of the PPB programme, two other progenies with round grains were also identified as medium tall (plot # MT2-3 and plot # 128-3). They were used in optimization of the microsatellite (SSR) diversity analysis.

Table 3.3: PPB bulks and varieties included in molecular studies.

Genotypes	Lines/origin	Grain type
LI-BIRD materials		
1. KIII x IR64 early tall (ET) bulk	PPB bulk, Nepal	Fine grain
2. KIII x IR64 medium dwarf (MD) bulk	PPB bulk, Nepal	Fine grain
3. KII x IR64 medium tall-1 (MT1) bulk	PPB bulk, Nepal	Fine grain
4. KIII x IR64 medium tall-2 (MT2) bulk	PPB bulk, Nepal	Fine grain
5. KIII x IR64 medium tall-3 (MT3) bulk	PPB bulk, Nepal	Fine grain
6. KIIIxIR64 plot # MT2 – 3 (medium tall)	PPB progeny, Nepal	Round grains
7. KIIIxIR64 plot # 128 – 3 (medium tall)	PPB progeny, Nepal	Round grains
LARC / NARC material		
8. Machhapuchhre-3 (Chhomrong x Fuji 102)	PPB variety, Nepal	Medium
9. Machhapuchhre-9 (Chhomrong x Fuji 102)	PPB variety, Nepal	Medium
10. Chhomrong (selection from landrace)	Cross parent, Nepal	Medium
GVT/BAU/CAZS material		
11. Ashoka 200F (Kalinga III x IR64)	PPB variety, India	Medium
Cross parents		
12. Kalinga III (Upland variety)	Cross parent (India)	Fine grain
13. IR64 (Irrigated, lowland variety)	Cross parent (IRRI)	Fine grain

3.1.3.2 LARC / NARC

In addition there were two PPB varieties from the Lumle Agriculture Research Centre (LARC), NARC, Nepal and one parent, Chhomrong. Chhomrong is a pure line selection variety from a cold tolerant landrace population from Kaski, Nepal with resistance to sheath brown-rot disease (ShBR) and is well adapted to altitudes above 1500 m (Sthapit *et al.*, 1996). Machhapuchhre-3 (M3) and Machhapuchhre-9 are two cold tolerant and farmers' acceptable rice varieties developed from a cross between Fuji 102 and Chhomrong using the pedigree-bulk method with the participation of farmers and breeders (Sthapit *et al.*, 1996).

3.1.3.3 GVT / BAU/CAZS

Ashoka 200F was another PPB variety from India included in the study. It was a product of a cross between Kalinga III and IR64 and was developed by bulk selection. It is early maturing, slender grained and resiatant to lodging (Kumar *et al.*, 2001). It was used in a collaborative project between the Gramin Vikas Trust (GVT), Birsa Agricultural University (BAU) and Centre for Arid Zone Studies (CAZS).

3.1.3.4 Check and parent varieties

There were also included Kalinga III and IR64. These were the cross parents of seven LI-BIRD PPB bulk populations and PPB variety, Ashoka 200F. Kalinga III is an upland rice variety identified in western India in a PVS programme (Joshi and Witcombe, 1996). It is well suited to Terai under partially irrigated conditions in Nepal and is grown as chaite rice (Witcombe *et al.,* 2001). IR64 is a high yielding semi-dwarf modern variety for irrigated condition included in the study.

3.2 Agro-morphological variability studies

3.2.1 Field experiments

Field experiments to measure agro-morphological variability on the rice landraces from the Jumla, Kaski and Bara ecosites were carried out in Agriculture Research Station in each ecosystem in the main seasons of 2000 and 2001 (Table 3.4). These research stations were located close to the study sites and hence were within a similar agro-ecological zone. The meteorological information recorded during the rice crop growth is given in Table 3.5.

Table 3.4: Locations of field experiments, total number of rice accessions and year of study.

Ecosite	Location	Region	No. of accessions in experiment	Year of study
Jumla	ARS, Vijayanagar, Jumla	Far-western	147	2001
Kaski	ARS, Malepatan, Pokhara	Western	291	2000
Bara	NRRS, Hardinath, Janakpur	Central	194	2001

Table 3.5: Meteorological information of three field study ecosites.

Month	Average maximum temp (0°C)			Average minimum temp (0°C)			Rainfall (mm)		
	Jumla †	Kaski ‡	Bara †	Jumla †	Kaski ‡	Bara†	Jumla†	Kaski ‡	Bara†
May	NA	30.4	33.1	NA	12.8	23.8	NA	596	186
June	24.6	31.6	33.8	14.4	19.0	25.6	180	831	276
July	25.2	29.8	33.8	17.3	20.2	27.1	106	1166	418
August	25.2	29.3	31.8	15.5	20.3	27.0	129	1120	300
September	25.8	30.6	32.2	12.6	15.3	25.6	44	551	300
October	23.5	30.0	30.3	6.1	12.2	22.8	12	83	281
November	19.9	NA	27.3	2.2	7.5	16.5	1	9	0

† = recorded in 2001; ‡ = recorded in 2000; NA = Not available

The landraces, along with modern varieties as checks, were seeded at each research station on dry, clean and raised seedbeds in small blocks of 0.5 x 0.5 m on different dates following the traditional cultural practices of the sites (Table 3.6) that are described in chapter V. The seedbeds were adequately watered and covered with a thin layer of straw to protect the germinating seeds from birds and to maintain the temperature of seedbed. When seedlings were large enough, at about one month old, they were transplanted into flooded fields.

Table 3.6: Seeding and transplanting dates in 2000 and 2001.

Ecosite	Seeding date	Transplanting date	Plot size	Fertilization
Jumla	25 March, 2001	8 June, 2001 9 June, 2001	3 m row per plot with 25 x 15 cm spacing	None
Kaski	12 June, 2000	12 July, 2000 13 July, 2000 14 July, 2000	3 m row per plot with 25 x 20 cm spacing	8 t ha^{-1} FYM
Bara	13 June, 2001	10 July, 2001 11 July, 2001	2 rows of 2 m per plot 25 x 20 cm spacing	60:30:30 NPK kg ha^{-1} 30N kg ha^{-1} top dressed

The trials were laid out in randomized complete block designs (RCBD) with three replications using the plot sizes shown in Table 3.6. Fertilization was applied as per the local practices at each eco-site: green manure and farmyard manure (FYM) are used in farmers' fields in the mid-hill region, so only FYM was used on the Kaski materials; in the high-hills agriculture is practiced under low input conditions so no fertilizer was applied in Jumla; in Bara, the rice field was top dressed with urea after establishment of the seedlings. Across the sites, seedlings were transplanted in the plots at the rate of one seedling per hill across the sites.

3.2.2 Measurement of agro-morphological traits

Data on agro-morphological traits of seed, seedlings and plant stands were recorded using the IBPGR-IRRI Descriptors for Rice at various growth stages of the rice plant (IRRI-IBPGR, 1980). These data include qualitative and quantitative traits and phenological traits of rice. Depending upon the trait, measurements/observations were made either on

individual plants or on the entire plot (Table 3.7). For individual plant data, 15 plants were used for measurement at the suggested growth stages.

Qualitative traits were recorded visually on individual plants and classified according to the categories of a trait. However, some qualitative traits like seedling vigour, leaf blade colour, and leaf senescence were recorded on a plot basis. The methods of recording the data, appropriate growth stage for observation and the standard states (scales or ranks) used for discrete qualitative and quantitative traits are explained in Table 3.7. Culm length was measured from the ground to the base of the spike (panicle) of main tiller. Panicle length was measured from the base of the panicle to its tip excluding the awns. Exsertion was measured as the length of the culm from the flag leaf auricles to the base of panicle. Days to 50 % heading (DH), days to maturity (DM), 1000 grains weight (TGW) and yield per plot were recorded on a plot basis.

Table 3.7: Qualitative and quantitative traits of rice plants and their measurement scales or states (ranks).

Trait	Method	Growth stage	States/ranks or measurement unit
Seedling characteristics:			
Seedling vigour	Observation†	5 leaf stage	1= extra vigourous; 3 = vigourous; 5 = Normal; 7= weak; 9= very weak
Leaf characteristics:			
2nd leaf length	Measurement	Late vegetative	cm (rounded to one decimal place)
2nd leaf width	Measurement	Late vegetative	cm (rounded to one decimal place)
Blade pubescence	Observation	Late vegetative	1 = glaborous; 2 = intermediate; 3 = pubescent
Blade colour	Observation†	Late vegetative	1 = pale green; 2 = green; 3 = dark green; 4 = purple tips; 5 = purple margin; 6 = purple blotch; 7 = purple
Basal leaf sheath colour	Observation	Late vegetative	1 = green; 2 = purple lines; 3 = light purple; 4 = purple;
2nd leaf angle	Observation	Late vegetative	1 = erect; 2 = horizontal; 3 =

			drooping
Flag leaf angle	Observation	Late vegetative	1 = erect; 3 = intermediate; 5 = horizontal; 7 = descending
Ligule length	Measurement	Late vegetative	cm (rounded to one decimal place)
Ligule colour	Observation	Late vegetative	1 = white; 2 = purple lines; 3 = purple
Ligule shape	Observation	Late vegetative	1 = acute; 2 = 2-clefted
Collar colour	Observation	Late vegetative	1 = pale green; 2 = green; 3 = purple
Culm characteristics:			
Culm length	Measurement	After flowering	cm (rounded to one decimal place)
Culm number	Measurement	After flowering	Count in number
Culm angle	Observation†	After flowering	1 = erect; 3 = intermediate; 5 = open; 7 = spreading; 9 = procumbent
Internode colour	Observation	After flowering	1 = green; 2 = light gold; 3 = purple lines; 4: = purple
Strength	Observation†	Flowering to maturity	1 = strong; 3 = moderately strong; 5 = intermediate; 7 = weak (most plants flat); 9 = very weak (all plants flat)
Panicle characteristics:			
Length	Measurement	Maturity	cm (rounded to one decimal place)
Type	Observation	Maturity	1 = compact; 5 = intermediate; 9 = open
Secondary branching	Observation	Maturity	1=light; 2 = heavy; 3 = clustering
Exsertion	Measurement	Maturity	1 = well exserted; 3 = moderately well exserted; 5 = just exserted; 7 = partially exserted; 9 = enclosed
Axis	Observation	Maturity	1 = straight; 2 = droopy;
Shattering	Observation†	Maturity	1 = very low; 3 = low; 5 = moderate; 7 = moderately high; 9 = high
Threshability	Observation†	Maturity	1 = difficult; 5 = intermediate; 9 = easy
Grain characteristics:			
Awning	Observation	Maturity	0 = absent; 1 = present
Apiculus colour	Observation	Maturity	1 = white; 2: = straw; 3 = yellow; 4

			= red; 5 = purple
Stigma colour	Observation	Flowering	1 = white; 5 = purple
Lemma and palea colour	Observation	Maturity	0 = straw; 1 = gold furrows on straw; 2 = brown spots on straw; 3 = brown furrows on straw; 4 = tawny; 5 = reddish; 6 = purple spots on straw; 7 = purple furrows on straw; 8 = purple; 9 = black
Lemma and palea pubescence	Observation	Maturity	1 = glaborous; 5 = long hairs
Sterile lemma colour	Observation	Maturity	1 = straw; 2 = gold; 3 = purple; 4 = red
Sterile lemma length	Observation	Maturity	1 = short longer than 1.5 mm: 3 = medium (1.6 – 2.5 mm)
Post harvest characteristics:			
No. of grains per panicle	Measurement	Maturity	Count in continuous number
No. of empty hulls per panicle	Measurement	Maturity	Count in continuous number
Yield panicle^{-1}	Measurement	Maturity	g (rounded to one place decimal at 12 % m.c.)‡
Yield plot^{-1}	Measurement	Maturity	g (rounded to one place decimal at 12 % m.c.) ‡
1000 grains weight at	Measurement	Maturity	g (rounded to one place decimal at 12 % m.c.) ‡
Grain shape	Observation	Maturity	1 = round and small; 2 = medium; 3 = slender and fine
Grain size	Observation	Maturity	1 = extra long; 2 = long; 3 = medium; 4 = short
Phenology:			
Days to 50% heading	Measurement†	Maturity	Days from seeding
Days to 50% flowering	Measurement†	Maturity	Days from seeding
Days to maturity	Measurement†	Maturity	Days from seedlig

IRRI (1980); observation and measurement (†) = based on whole plot; ‡m.c. = moisture content

3.3 SSR (Microsatellite) diversity studies

Standard molecular biological chemicals and general techniques for preparing stock solutions, buffers, reagents and the equipment used are described in detail with brand names and protocols in Appendix 3.2a-c. These techniques were conducted according to

Sambrook *et al.* (1989). Molecular work was conducted at the University of Wales, Bangor, UK and at Agriculture Botany Division, NARC, Kathmandu, Nepal.

3.3.1 Preparation of seedlings for DNA isolation

Seeds of each accession under study were sown in John Innes No.2 compost in a plug tray and grown in a green house at Pen-y-Ffridd, University of Wales, Bangor, UK under daylight supplemented by halogen lamps at 150 μmol m^2 s^{-1} PAR (min. temp. 25°C).

3.3.2 DNA extraction

In this study, DNA was extracted using Qiagen DNeasy Kits (Qiagen DNeasy plant mini handbook, 1999). This extraction method constitutes three steps:

Lysis of cell wall and membranes to free DNA into solution,

Purification of DNA by precipitating proteins and polysaccharides, and

Purification of DNA and re-suspension in a buffer.

Young leaf tissue of 2-3 week old healthy seedlings was used for genomic DNA extraction. Fresh leaf material from each accession was weighed out to 100 mg for either bulk or individual plant DNA extraction and ground to a fine powder with a pre-chilled mortar and pestle in liquid nitrogen. For bulk DNA 10 seedlings were used. Further steps followed the Qiagen protocol (Appendix 3.2d).

3.3.3 Estimation of DNA concentration in the extract

DNA concentration in each extract was estimated by recording absorbance at 260nM wavelength in a Hewlett Packard 8453 spectrophotometer using a HP845x UV-visible

system at 190-400 nM display spectrum and 260 nM, 280 nM wavelengths. DNA concentration was calculated using the formula as follows (where DF = dilution factor):

$$Abs_{260nM} \times 50 \times DF$$

A visual estimate was also made by running the DNA samples on a 1 % (w/v) agarose minigel in 1 x TBE buffer (0.09 M Tris-borate and 0.5 M EDTA) for 1.5 hour at 80 V. For this 2 µl of sample was mixed with 2 µl of loading buffer and 6 µl of SDW (sterilized distilled water) to make 10 µl. Known concentrations of Lambda DNA diluted with 1 x TE were also run on the same gel. Visual recording of the fluroscence of the DNA band was done under UV light with ethidium bromide staining.

3.3.4 PCR amplification of SSR (microsatellites)

Specific primer pairs were used to amplify the simple sequence repeats. These were selected from the published primer sets of 323 microsatellite markers that are available in genomic libraries, RiceGene Database and the Genbank Database (Cho *et. al.*, 2000). The selected markers were with $(GA)_n$, $(CT)_n$, $(AT)_n$ and $(AAAT)_n$ repeats. Optimisation of PCR reaction conditions for amplification was conducted to obtain repeatable results with good yield. Each variable for optimization is described below. PCR reaction mixtures were then amplified in a MJ Research PTC – 100^{TM} Programmable Thermal Controller with Hot Bonnet (MJ Research, INC, Waltham, MA, USA) holding 96 x 0.2 ml - microfuge tubes or one 96-well, V-bottom plate.

(a) *Template DNA concentration*

Different concentrations of template DNA in the PCR amplification reaction were tested for optimal PCR amplification. 5 µl DNA extract (i.e. 4 ng, roughly of the same conc.)

per 1 µl of PCR reaction was found to be effective. This concentration was used for all the SSR (microsatellite) analyses.

(b) *Mg^{2+} ion concentration*

The concentration of Mg^{2+} ions in the PCR reaction is another important factor which affects the PCR amplification. Optimal $MgCl_2$ concentration may vary from approximately 0.5 mM to 5.0 mM Mg^{2+} and influences enzyme activity. In the method of McCouch *et al.*, 1997, a concentration of 2.5 mM was used for rice microsatellites. Two concentrations, 2.5 mM and 3.0 mM Mg^{2+}, were tested for optimum amplification with the DNA samples extracted from the rice landraces. The increase in Mg^{2+} concentration did not affect the production of amplified products.

(c) *Thermal cycling profile and cycle number*

A repetitive series of thermal cycling (heating and cooling) drives the three steps of PCR: denaturation, annealing and extension, and results in the accumulation of many copies of a specific DNA fragment. Primer annealing temperature, extension time and the number of cycles are critical factors for successful PCR. Five different thermal cycling profiles were tested for optimum PCR amplification with these landraces. These were as follows:

I Microsat 1 (From McCouch *et al.*, 1997)

Step1	Initial denaturation	:	94°C 5 mins
Step2	Denaturation	:	94°C 1 min
Step3	Primer annealing	:	55°C 1 min
Step4	Primer extension	:	72°C 1 min
Step5	Repeat from step2 to 4:		34 times

	Step6	Prolonged extension :	72°C 5 mins
	Step7	Refrigerated at 4°C for infinite hold	

II Microsat 2 (As I, but with increased extension time to 2 minutes per cycle)

III Microsat 3 (As I, but with reduced annealing temperature of 52° C)

IV Touchdown programme 1 (Annealing temperature reduces from 65° to 57° C, Cho *et al.*, 2000 with some modifications)

Step 1	Initial denaturation	: 94°C 5 mins.
Step 2	Denaturation	: 94°C 1 min.
Step 3	Primer annealing	: 65°C 1 min.
Step 4	primer extension	: 72°C 2 mins.
Step 5	Repeat step2 to 4	: 1 time.
Step 6	Denaturation	: 94°C 1 min.
Step 7	primer annealing	: 62°C 1 min.
Step 8	Primer extension	: 72°C 2 mins.
Step 9	Repeat step6 to 8	: 1 time
Step 10	Denaturation	: 94°C 1 min.
Step 11	Primer annealing	: 59°C 1 min.
Step 12	primer extension	: 72°C 2 mins.
Step 13	Repeat step10 to 12	: 4 times
Step 14	Denaturation	: 94°C 1 min.
Step 15	Primer annealing	: 57°C 1 min.
Step 16	Primer extension	: 72°C 2 mins.
Step 17	Repeat step 14 to 16	: 25 times

Step 18 prolonged extension : 72°C 5 mins.

Step 19 Infinite hold : 4°C

V Touchdown programme 2 (As IV, but with annealing temperature reducing through the cycles from 57° to 54° C)

The touchdown programme 1 (IV) with reducing annealing temperature profiles from 65° to 57° C was found to give optimum amplification with most primers and it was used for all the studies reported here.

(d) *PCR reaction and use of Taq DNA polymerase or Reddy Mix™ PCR Master mix*

Two methods were compared. The PCR protocols using *Taq* polymerase and Reddy Mix PCR master mix are given in appendix 3.2e. The protocol developed by McCouch *et. al.*, 1997 for SSR (microsatellite) analysis was followed with minor modifications. The PCR reaction conducted in a volume of 25 µl contained 5 ng of genomic DNA, 20 µM each primers, 2.5 mM each of dGTP, dATP, dTTP, and dCTP (Promega UK Ltd, Southampton, UK), 10 x PCR buffer, 3 mM $MgCl_2$ and 5 units of *Taq* DNA polymerase (Promega UK Ltd, Southampton, UK).

Reddy Mix™ PCR Master mix (ABgene) is an enzymatic mixture contained 3.0 mM $MgCl_2$, 10 x PCR buffer, *Taq* polymerase and blue dye. PCR reactions of 25 µl were conducted using Reddy Mix™ PCR Master mix and contained the same concentration of DNA sample and primers. PCR with Reddy Mix PCR master mix was found to produce discrete amplification products and was highly repeatable and more consistent than using Taq and other reagents separately. Reddy Mix was used continuously throughout the diversity analysis.

(e) *Selection of primers for diversity study*

Fifty pairs of selective (12-24)-mer primers, distributed in 12 chromosomes of rice were tested for polymorphism with 5 PPB bulk populations, two round-grained PPB progenies and two parental lines. Among these, 39 primer pairs gave good PCR amplification under optimal conditions. Details of microsatellite primers, their information content and approximate product size compared with IR36 are described in Appendix 3.3.

3.3.5 Gel Preparation, PCR product separation and detection of microsatellites

PCR products were separated by horizontal agarose gel electrophoresis. A 3 % (w/v) gel was made using 9 g Amresco SFR agarose (Anachem LTD) sprinkled slowly in 300 ml chilled 1 x TAE buffer in a conical flask and mixed uniformly. Then it was boiled in microwave with frequent swirling and boiling to dissolve agarose completely without leaving any clumps. Ethidium bromide (1 µl per 100 ml of distilled H_2O) was added and mixed uniformly and the gel was poured into gel casting tray and allowed to set and chilled before running the electrophoresis.

Samples were prepared using 2.5 µl of gel loading buffer (0.5% (w/v) bromophenol blue, 0.5M $Na_2EDTA.2H_2O$, 50% (v/v) glycerol, and 20% (w/v) Ficoll 400) added to 10 µl of the amplified PCR reaction to give 12.5 µl of each product for gel electrophoresis. For the samples using Ready Mix, the PCR product was run directly for electrophoresis. Gels were run under 1 x TAE buffer for 4-5 hours at 90 V. A DNA ladder of 1 kb was used as a standard for molecular weight estimation of PCR products. Gels were stained with ethidium bromide, visualized under UV illumination and photographed using a Kodak Digital Science (Macintosh version) or Polaroid Gel Cam

camera system. The approximate size of alleles between 100-300 basepairs (bp) was measured by comparison with the known standard.

3.4 Statistical analysis

3.4.1 Agromorphological data analysis

3.4.1.1 Analysis of variance

For traits that were measured on 15 plants, means were calculated for the plot. These plot mean values were used in subsequent statistical analyses. A univariate analysis was carried out for each trait to describe the range of variation for the accessions under study.

For qualitatively scored traits, the frequency (%) of each class was calculated across the landraces for each site. Likewise, for each quantitative trait, the distribution was checked for normality and histograms were plotted of the frequency of distribution. The distribution of most traits approximated to a normal distribution.

An analysis of variance (ANOVA) using a completely randomized block design was carried out for all traits in each of three sites (Table 3.1). Mean squares among landraces was tested against the error mean squares (landraces x replications) using an 'F' test. Traits with non-significant variation among the landraces were not included in further analysis.

Table 3.8: Analysis of variance for a randomized complete design at a single location.

Source	d.f	MS	Expected MS	F ratio
Among landraces, n	n-1	MS1	$\sigma_e^2 + r\sigma_g^2$	MS1/MSE
Replications, r	r-1	MSr	-	MSr/MSE
Error (landraces x replications)	(n-1) (r-1)	MSE	σ_e^2	

Where n = number of landraces; r = number of replications; σ_e^2 = error variance, which is MSE; σ_g^2 = genotypic variance. MS1 is the phenotypic variance among landraces.

3.4.1.2 Repeatability

The ratio of genotypic variance to the total phenotypic variation is the heritability. Heritability refers to a population from which the sample has been drawn at random. However, in the present case the accessions (landraces) were non-random and did not represent a base-population. In such a situation the ratio of the genotypic (σ_g^2) and phenotypic (σ_p^2) variation cannot be called heritability. Fehr (1987) defined such a measure as repeatability (r). This repeatability is different from that of Falconer (1989) and Santos (1999) where it refers to the analysis of repeated measurements on the same individuals. However, repeatability equals the intra-class correlation of Falconer (1989). Repeatability for landrace means was computed as:

$$\frac{r\sigma^2_g}{\sigma^2_e + r\sigma^2_g} = \frac{(MS1 - MSE)}{MS1}$$

The landrace mean error variance is σ^2_e / r since the error variance for a mean over 'r' replication will be r times smaller; the phenotypic variance among landrace means is $\sigma^2_p = \sigma^2_g + \sigma^2_e / r$ (Table 3.8). The landrace mean repeatability is then σ^2_g/σ^2_p. The genotypic variance (σ_g^2) among landrace means is computed as (MS1-MSE/r).

However, the estimates of repeatability in the present study were based on data from one location in one year. Landrace mean phenotypic variance (σ^2_p) is equal to $\sigma^2_e / r + \sigma^2_{ge} / t + \sigma^2_g$ (Fehr, 1987) where r = number of replications and t = number of locations, σ^2_{ge} variance caused by the genotype x environment interaction of different locations or years and σ^2_g is the genetic variance. These components of phenotypic variance cannot be estimated in the present study.

Repeatability was calculated for those traits that had significant F values. Traits with < 0.2 repeatability were not included in further multivariate analysis, as they are more likely to contribute random, rather than genetic effects to the analysis.

3.4.1.3 Correlations between traits

The accepted traits with >0.2 repeatability were further analysed to measure degree of linear relationships between them. For this Pearson's simple correlation analysis was performed between pairs of accepted traits to calculate Pearson's product correlation coefficient as:

$$r_i = \frac{Cov_{xy}}{\sqrt{(Var_x)(Var_Y)}}$$

where Cov_{xy} is the covariance between x and y traits; Var_x is the variance of x trait and Var_y is the variance of y trait. The r^2 values were tested for significance against the tabulated values for n-2 degree of freedom.

3.4.1.4 Coefficient of variation

Coefficient of variation (CV) is also taken as a measure of diversity, which estimates sample variation for quantitative traits. The variation is expressed relative to the mean of the sample so that it is also called a measure of relative variability or relative dispersion (Zhong and Qualset, 1995; Jarvis *et al.,* 2000). It was calculated as:

$$CV\ (\%) = \frac{\text{Standard deviation}}{\text{Mean}} \times 100\ \%$$

It can be used to compare variation across traits or populations that have very different means and hence very different standard deviations. The mean, range, standard deviation (SD) and coefficient of variation (CV) were used to generally describe the variation in the quantitative traits.

3.4.1.5 Shannon-Weaver Diversity Index

Shannon Weaver diversity index measures the proportional abundance, combining richness, the number of varieties or traits of a given variety, and the frequency of their occurrence (Magurran, 1991). For the categorical data for qualitative traits, such as panicle type, leaf colour, pubescence, awning, and lemma and palea colour, the Shannon Weaver Diversity Index was therefore used as a measure of diversity. It was calculated as follows:

$$H' = -\sum_{i=1}^{k} p_i \log p_i$$

Where k is the number of categories and p_i is the proportion of the observations encountered in category i (Shannon, 1948). The value H' is based on the distribution and

number of the categories of a trait over the observations (number of landraces in the present study). It therefore ranges from 0 value for a non variable trait to any arbitrary number greater than 0 for a variable trait.

3.4.1.6 Multivariate analysis

The data of the accepted traits from above mentioned statistical analyses, were standardized to a mean of zero and a variance of unity to avoid differences in scales used in recording data on the different characters to ensure to have equal weight in the analysis (Sneath and Sokal, 1973). These normalized data were used in following series of multivariate analyses.

(a) Measures of distance and cluster analysis

A number of measures can be used for measuring the similarity and distance values between two landrace populations based on the means of multivariate information (traits). Some multivariate distance measures used in the study comparing among the landrace populations and among locations are described below:

Average taxonomic distance

Average taxonomic distance was a measure of diversity used as measure of pair-wise dissimilarity among landraces under study for agro-morphological variation. It is calculated as follows:

$$E_{ij} = \sqrt{\sum_k 1/n\ (k_i - k_j)^2}$$

Where, k is the vector (trait data) and i and j are the factors corresponding to the distance for n size of the factors and to angular difference between the vectors.

Nei's genetic distance

Nei's genetic distance was another common measure of diversity used as distance or dissimilarity between two landraces. In this case it was based on presences or absences of alleles for SSR (microsatellite) marker diversity. It was calculated as follows:

$$d_{ij} = -1n \frac{\sum_k (x_{ki} x_{kj})}{\sqrt{\sum_k x^2_{ki} x^2_{kj}}}$$

where, k is the allelic information and i and j are two landrace populations.

The agro-morphological and the marker data were carried for the cluster analysis to express the relationships between landraces putting the results into a meaningful structure of dendrogram. The method used was agglomerative hierarchical clustering which sorted out the coefficient values using Ward's linkage or UPGMA (unweighted pair-group arithmetic averaging). The general hierarchical agglomerative clustering analysis using Ward's linkage was employed for the agro-morphological traits. In Ward's linkage, the distance between two clusters is the sum of squared deviations and it minimizes the within cluster sum of squares (Ward, 1963). There could be possible of the negative value of the similarity matrix, if the distance between two clusters is greater than the original values of distance matrix. The equation is as follows:

$$Dmj = \{(N_j+N_k)d_{kj} + (N_j+N_l)-N_j d_{kl})l\ (N_j+N_m)$$

where N_j, N_k, N_l, and N_m are the number of observations. The details of the parameters that were carried with different sets of molecular and agro-morphological data were explained in the respective chapters.

(b) Principal component analysis

In ordination, similar accessions are tied close together and dissimilar ones are dispersed far apart reflecting the relationships in higher dimensions with minimum distortion. Principal component analysis (PCA) is the simplest and most widely used ordination method designed to reduce the number of variables that need to be considered to a small number of indices (called principal components) that are linear combinations of the original variables (Manly, 1994; and Beuningen and Busch, 1997). In the present study, the PC analysis was carried out using different combinations of the measured agro-morphological traits for each site and across the sites to reflect the traits relationships. Results were plotted as scatter diagrams using the traits in two axes of the PCA.

PCAs were performed for each site on agro-morphological and molecular marker traits:

- All agro-morphological qualitative and quantitative traits
- All qualitative traits
- All quantitative traits
- All significant agro-morphological qualitative and quantitative traits
- All significant qualitative traits (in one case, where four traits were perfectly correlated only one of them was included)
- All significant quantitative traits
- Alleles of all polymorphic markers

PCA was also carried out across the three sites using all common agro-morphological traits and molecular marker traits (Chapter VII). The datasets were:

- All agro-morphological qualitative and quantitative traits
- All qualitative traits
- All quantitative traits
- All significant agro-morphological qualitative and quantitative traits
- All significant qualitative traits (in one case, where four traits were perfectly correlated only one of them was included)
- All significant quantitative traits
- Alleles of all polymorphic markers

(c) Linear discriminant function analysis

Linear discriminant analysis is another widely used classification method in diversity studies of plant genotypes. This analysis assigns the set of distinct accessions to one or several groups based on linear combinations of observations or measurements (traits). Discriminant function analysis in the present study was therefore carried to assess the conformity of accessions to the named farmers' unit of diversity (FUD), agro-morphological clusters formed in cluster analysis and the rice domains as the membership groups for each site. For this analysis, the same set of the significant agro-morphological traits were used. The principle of the analysis is described along with the respective results in Chapter V.

3.4.2 Molecular marker analysis

The amplified fragments seen as bands on gels after staining were scored as alleles and assigned 1 for presence and 0 for absence. The approximate size (in nucleotides as base pairs) of each band was determined based on its migration relative to the 1 kb DNA

ladder and the band which moved fastest, was taken as the smallest sized band. For subsequent numerical analysis, a rectangular binary data matrix was generated. The following diversity indices were calculated for each set of rice landraces from the three sites. These were the average number of alleles (A), the average number of polymorphic alleles (Ap), the percentage of polymorphic loci (PPL) and the percentage of polymorphic alleles (PPA).

3.4.2.1 Polymorphic Information Content (PIC)

Polymorphic information content (PIC) values were calculated for each microsatellite based on the allelic frequency detected in the studied accessions. This gives a relative value to each locus on the diversity it reveals according to the number of detectable alleles and their frequency (Botstein, *et al.*, 1980; and Powell *et al.*, 1996). A locus with low information has few alleles at low frequency. PIC values range from 0 to 1 and increase as the number of alleles increase and alleles occur at more or less equal frequency. The markers with highest values are best used to distinguish variety and markers with lowest values indicate the rare allele and are useful in structuring the varieties (Luce *et al.*, 2001). Anderson, *et al.*, (1993) later on produced a simplified version for self-pollinated species. It is calculated for each marker as follows:

$$PIC_i = 1 - \sum_{J=1}^{n} p^2_{ij}$$

where p_{ij} is the frequency of the *j*th allele for marker I, and summation extends over n alleles. This PIC value calculated for each marker over accessions is same as for Nei's Gene Diversity measure H_s (Nei, 1973).

3.4.2.2 Other molecular diversity parameters

Genetic diversity of the entries/populations based on set of measured molecular data, can be estimated using different diversity parameters other than PIC (Sun *et al.*, 2001). These are calculated as follows:

(a) Percentage of polymorphic loci (PPL):

$$P = (k/n) \times 100\%,$$

where k is the number of polymorphic loci, n is the total number of loci investigated.

(b) Average number of alleles per locus (A):

$$A = \Sigma A_i / n,$$

where A_i is the number of alleles at the ith locus, n is the total number of loci investigated.

(c) Average number of alleles per polymorphic locus (Ap):

$$Ap = \Sigma Ap_i / n_p,$$

where Ap_i is the number alleles at a certain polymorphic locus, np is the total number of polymorphic loci investigated.

(d) Percentage of polymorphic alleles (PPA)

$$PPA = (\Sigma Ap_i / \Sigma A_i) \times 100\ \%$$

3.4.3 Measuring agreement between agro-morphological and molecular variation

Mantel (1967) developed a test for matrix correspondence that takes two matrices and plots one against the other, element by element. This test yields a product-moment correlation (r) that is the measure of relatedness between two matrices. It is a useful

statistical test used to compare two independent matrices and is interpreted in terms of the relationship or agreement or goodness of fit between the two matrices. Thus the value of r lies in the range of -1 to +1, with r = -1 indicating a perfect negative correlation, and r = +1 indicating a perfect positive correlation (Manly, 1994). Mantel's test was used to compare the genetic distance matrix and taxonomic distance matrix based on molecular and agro-morphological traits for all the combinations within and between the sites. For two independently derived matrices, a correlation value greater than value 0.5 is significant at the 0.01 probability level (Tatineni *et al.* 1996).

3.5 Software

For the calculation of the statistic information in the analysis of multivariate data of agromorphological traits and SSR markers using classification and ordination methods, MINITAB-12, and NTSYS pc version 1.8 (Rohlf, 1992) were used. The Microsoft excel spreadsheet programme was used for original data preparation and for some preliminary calculations. MINITAB-12 and NTSYS statistical packages are designed for multivariate methods and are provided with options for calculation of the above-mentioned multivariate analyses and statistical parameters. NTSYS is provided with an option (MXCOMP) to conduct Mantel's test on the distance and similarity of matrices.

CHAPTER IV
ON FARM DIVERSITY OF RICE

4.1 Introduction

This chapter describes the activities carried out by groups of staff in the IPGRI project on *in situ* conservation to get preliminary socio-economic data and a general picture of the crop genetic diversity. All data presented here are from the group exercises.

In Nepal, the extreme variation in altitude, topography, physical and climatic conditions and the antiquity of agriculture have enriched the country with an immense diversity of crop species in the form of local cultivars i.e. landraces. About 95 % of the farming community in Nepal depends on landraces for their crop production (Sthapit, 2001). However, very little has been done to understand the diversity of landraces in traditional ecosystems by learning about the indigenous knowledge of indigenous farming communities.

Jumla, Kaski and Bara are the three districts of the country where the *in situ* conservation project ecosites are located. These districts were selected in 1996 jointly by the scientists from the project partner institutions: the Nepal Agriculture Research Council (NARC); the non-governmental organization the Local Initiatives for Biodiversity, Research Development (LI-BIRD); and the International Plant Genetic Resources Institute (IPGRI). They did this on the basis of the degree of agro-biodiversity and status of on-farm conservation (Upadhyay and Sthapit, 1995), and to represent the three major physiographic agro-ecozones. The three selected districts are located in three administrative regions of the country: the Far Western development region; the Western development region and the Central development region (Figure 4.1 and Table 4.1).

Figure 4.1: Map of Nepal showing the location of the *in situ* project ecosites (shaded areas).

Table 4.1: Description and characteristics of the three IPGRI *in situ* conservation project study sites (1998).

Ecosite	*In situ* village boundaries	Eco-physiographic zones	Administrative zones	Climatic range	Level of crop diversity	Access and intervention
Jumla	Tallium and Kartikeswami	High hill (2240-3000 m)	Mid western region	Cool temperate to alpine	Moderate to high	Low
Kaski	Begnas and Rupakot	Mid hill (668-1206 m)	Western region	Sub-tropical	Very high	Slight
Bara	Kachorwa	Terai (100-150 m)	Central region	Sub-tropical to tropical	Moderate to low	High

Paudel *et al.*, (1998); Rijal *et al.*, (1998); and Sherchand *et al.*, (1998).

This chapter is based on the survey reports and describes the amount and distribution of rice diversity maintained by the farmers in the three study sites. Each rice landrace has been given a name, which could be specific to either its morphological traits, or to its adaptation to agro-environments (Pham *et al.*, 2001). These names and landraces, along with indigenous knowledge on their cultivation and uses, are passed from generation to generation. They are continuously subjected to selection pressures to adapt to the local environments and this contributes to the conservation of diversity (Brush, 1991).

4.2 Materials and methods

The amount (extent) of genetic diversity in rice landraces in farmers' fields was measured by the number of named landraces, number of farming households growing each landrace and the area covered by each of them. Within each site, the relative importance of each landrace and diversity of domains and landrace distribution was determined by the surveys at the household level.

4.2.1 Site descriptions

The general descriptions and physiographic location of the ecosites are given in Table 4.1, Appendix 3.1, and Figure 4.1

4.2.2 PRA survey

Site selection was undertaken as a first activity of the project in 1998 in these three districts. For site selection, participatory rural appraisals (PRA) were employed and were carried out by a team represented by the researchers from the National Multi-Disciplinary

Group (NMDG) and the Local Multi-Disciplinary Group (LMDG), two groups formed in the project for planning and implementing the project activities. The groups had researchers from NARC, LI-BIRD, Department of Agriculture Development (DOAD) and Agriculture Development Offices (ADO) of Jumla, Kaski and Bara. The NMDG was based at the national level and the LMDG was based in the field. By group consensus the following were studied: the amount of crop diversity and genetic diversity at the species level of rice, barley, finger-millet, buckwheat, sponge-gourd, cucumber, pigeonpea and taro; the agro-ecological diversity; the socio-cultural diversity; the market access and the farmers' interest and willingness to collaborate in the project activities. The PRA processes included group discussions and the interviews with individuals to collect information. Based on the information collected and a comparative assessment of the considered sites, Tallium and Kartikswami in Jumla; Begnas and Rupakot in Kaski and Kachorwa in Bara were selected for detailed studies in the *in situ* conservation project (Rijal *et al.*, 1998, Paudel *et al.*, 1998 and Sherchand *et al.*, 1998). The diversity surveys concentrated on rice, as it was the most important crop. The information was collected at the village level and based on the farmers' knowledge on their landraces and a given vernacular name for a landrace was taken as the farmer's unit of diversity (FUD).

4.2.3 Diversity fairs

Diversity fairs were organized at each site in 1998: on 5th June in Begnas (Kaski); 24th November in Tallium (Jumla); and 23rd December in Kachorwa (Bara). Based on the agro-ecological and administrative boundaries of the villages, the total farming households were grouped into 16 farmers groups in Kaski, 20 in Jumla, and 22 in Bara.

There were 21-85 households per group. Packaging material and an information sheet on the varieties for each display were provided a day before the diversity fair. The groups were asked to fill the information sheet with general information such as: which characteristics of the variety were used by farmers for its identification; the right domain of the variety; its social, religious and cultural importance; and the source of the seed. All the groups from each site took part in the fair and displayed seeds of the rice landraces which were then kept by the project for further studies.

4.2.4 Baseline survey

A baseline survey was carried out in each site at the household level to obtain broad information relating to the maintenance of genetic diversity. The processes adopted for baseline study and the dates in which it was undertaken are summarized in Figure 4.2 and Table 4.2. The social science group of the project designed and developed the questionnaire. The questionnaire was tested on a small sample, and changes were then incorporated before a final version was approved for the full survey.

The study employed a proportionate, stratified random-sampling design. Wealth category was used for stratifying the households and was determined, at each site, by a group of key informants (3-9 farmers) who had previously been identified during the PRA exercises for site selection. They agreed the parameters for ranking each household (Table 4.2) and then categorized all the households into three wealth categories: 'resource-rich', 'resource-medium' and 'resource-poor'. A proportionate (22-23% of total households) random sample of households from each category were identified for field implementation of the questionnaire. The key person in the household who was

responsible in making decisions in relation to agricultural matters was interviewed using the questionnaire. The questions were on: the major source of livelihood; farmer resources (land holding, type of land and fertility); availability of irrigation; the cropping pattern; the varieties grown (the number, area, and its trend; productivity and its trend; reasons of growing each cultivar; the characters used to distinguish them; the preferred traits and specific adaptive traits of each cultivar). The survey was carried by the social group and field staff of the project. Only the diversity-related information from this survey was extracted and analysed here.

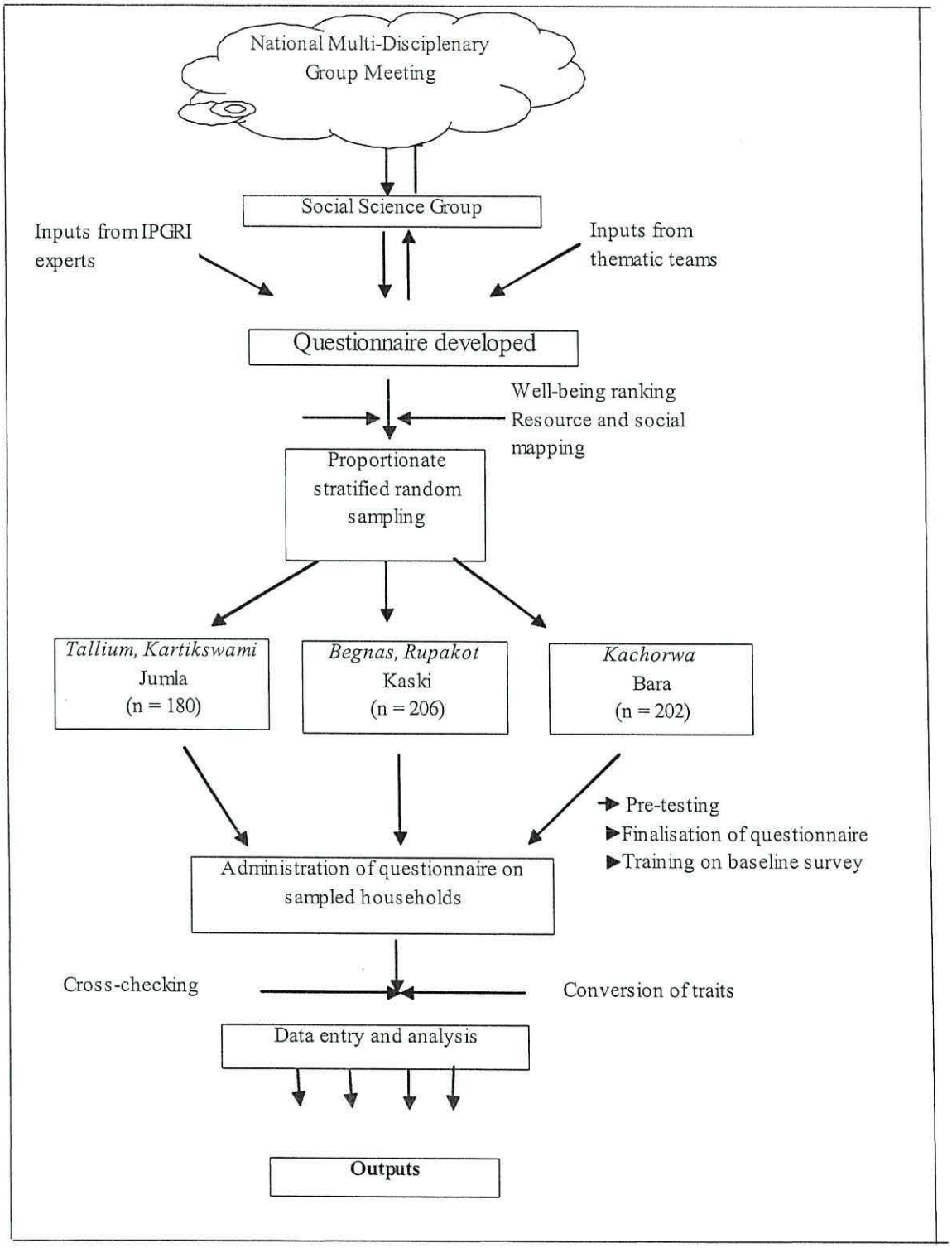


Figure 4.2: Methodology adopted in baseline study (Rana *et al.*, 2000a,b,c).(n = no. of households)

Table 4.2: Dates and processes undertaken in baseline survey and criteria used in wealth ranking the households in each site.

S.No	Ecosite	Dates	Activities	Criteria for wealth ranking
1	Tallium-Kartikswami, Jumla	31/10/98	Orientation on baseline exercise to staff	Land holding Food sufficiency
		01/11/98	Pre-testing of questionnaire	Orchard
		01/11/98	Incorporation of changes in questionnaire	Livestock
		02/12/98	Wealth ranking	Other source of income (services, wages, out income)
		02-12/11/98	Field exercise and implementation	
2	Begnas-Rupakot, Kaski	19/01/99	Orientation on baseline exercise to staff	Not available
		20/01/99	Pre-testing of questionnaire	
		21/01/99	Incorporation of changes in questionnaire	
		22/01/99	Duplication of questionnaire	
		NA	Wealth ranking	
		23/01/99 to 03/02/99	Field exercise and implementation	
3	Kachorwa, Bara	27/12/98	Orientation on baseline exercise to staff	Not available
		28/12/98	Pre-testing of questionnaire	
		28/12/98	Incorporation of changes in questionnaire	
		29-30/12/98	Wealth ranking	
		02-15/01/99	Field exercise and implementation	

Rana *et al.*, (2000a,b,c).

4.2.5 Farmers' indegenous knowledge

The characterization of the rice landraces in the standing crop in farmers' field was carried out in the Begnas and Rupakot (Kaski) and the Kachorwa (Bara) ecosites to verify the diversity. This activity could not be carried out in Jumla because of a problem of access. The study was accomplished by a general visit of the field staff to rice plots during the rice season. Farmers were interviewed about the indigenous knowledge concerning particular varieties, and asked which characters they used to distinguish a

variety and its useful characters and other adaptive traits. From the complete information collected in the surveys, all the data related to rice diversity were extracted and analysed.

4.3 Results

4.3.1 Total number of rice varieties (richness of landraces) by ecosite

A range of rice landraces, each with a farmer-given name was grown in all three ecosites. The sampled households in the mid-hills and *terai* (lowland) ecosites grew more landraces and modern varieties (MV) of rice than those in Jumla (high-hill) (Figure 4.3). The highest number of variety names were encountered (65-75) in the mid-hill eco-site (Kaski) and the lowest (10-21) in the high-hill ecosite (Jumla). No modern varieties were grown in the high-hills. The number of named varieties varied by the survey methods used. The diversity fair gave the most landraces in Kaski and Bara and PRA survey gave the fewest, or equal fewest, landraces (Table 4.3). The baseline survey and the characterization in the farmers field were found reliable methods with detailed information and could verify the landraces in the standing crops. The number of landraces generally decreased from low to high altitude in all the survey methods.

Table 4.3: Number of rice landraces with given names documented in various methods in three eco-sites, Nepal (1998-2000).

Methods	Tallium and Kartikswami, Jumla	Begnas and Rupakot, Kaski	Kachorwa, Bara
PRA survey	10	38	49
Diversity fair	11	75	79
Baseline survey	17	69	55
Characterisation in farmers fields	-	65	55

Paudel *et al.*, (1998); Rijal *et al.*, (1998); Rana *et al.*, (2000a,b,c); and Sherchand *et al.*, (1998).

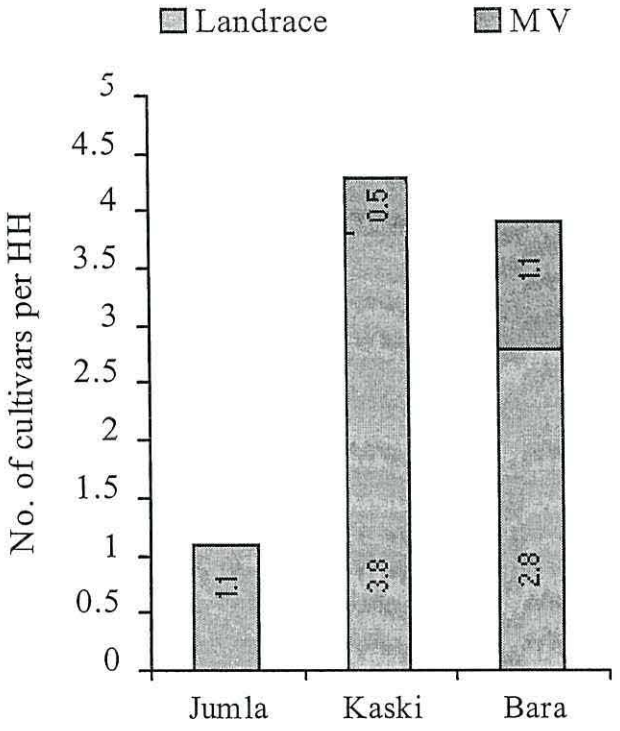


Figure 4.3: Distribution of genetic diversity at household (HH) level in three sites

4.3.2 Average area covered by each rice landrace (frequency/evenness)

Rice landrace diversity at the household level was measured in terms of the relative area covered by these landraces and the number of households growing each landrace. The majority of landraces across the sites were grown in a small area by few farmers (i. e. a small proportion of the total number of sampled households) (Table 3.1 from Chapter III and Figure 4.4). Only a few landraces were grown in larger than average areas by many households in each site. Landraces in this category were highly preferred for their quality traits; had wide adaptation to adjacent domains, and had value for local consumption and were more productive (Appendix 4.2). Examples were:

- Ekle, Mansara, Kathegurdi, Jethobudho from the Kaski site,
- Sotawa, Nakhisaro, Mutmur from the Bara site, and
- Marshi group from the Jumla site

which were mostly known for their grain quality and were grown by many farmers for their social prestige and for commercial values in larger areas than average (Sthapit *et al.*, 2001). Gurdi, in Kaski, and Mansara, in Bara, were the landraces, which were grown in a larger area than average by a few farmers mainly for local consumption.

Across the ecosites, the majority of the landraces were grown in a smaller than average area by a small proportion of the total sampled households (few households). Based on the total number of landraces reported by the sampled households, these accounted for 71% of landraces in Jumla, 76% in Kaski and 54% in Bara (Figure 4.4). Most landraces in this category were reported to possess varied use values. For example, in Kaski the landraces Naltumme and Tunde were grown in marginal environments. In Bara, landraces Bhatti and Silhat and in Jumla, landrace Darime were adapted to the stress environment. Landraces such as Ramani and Biramphool from the Kaski ecosite and Dudhisaro, Dudhraj and Chhatraj from the Bara ecosite were valued for their cooking quality. The landraces multiple use value and their specific adaptation to marginal environments helped to maintain diversity (Appendix 4.2).

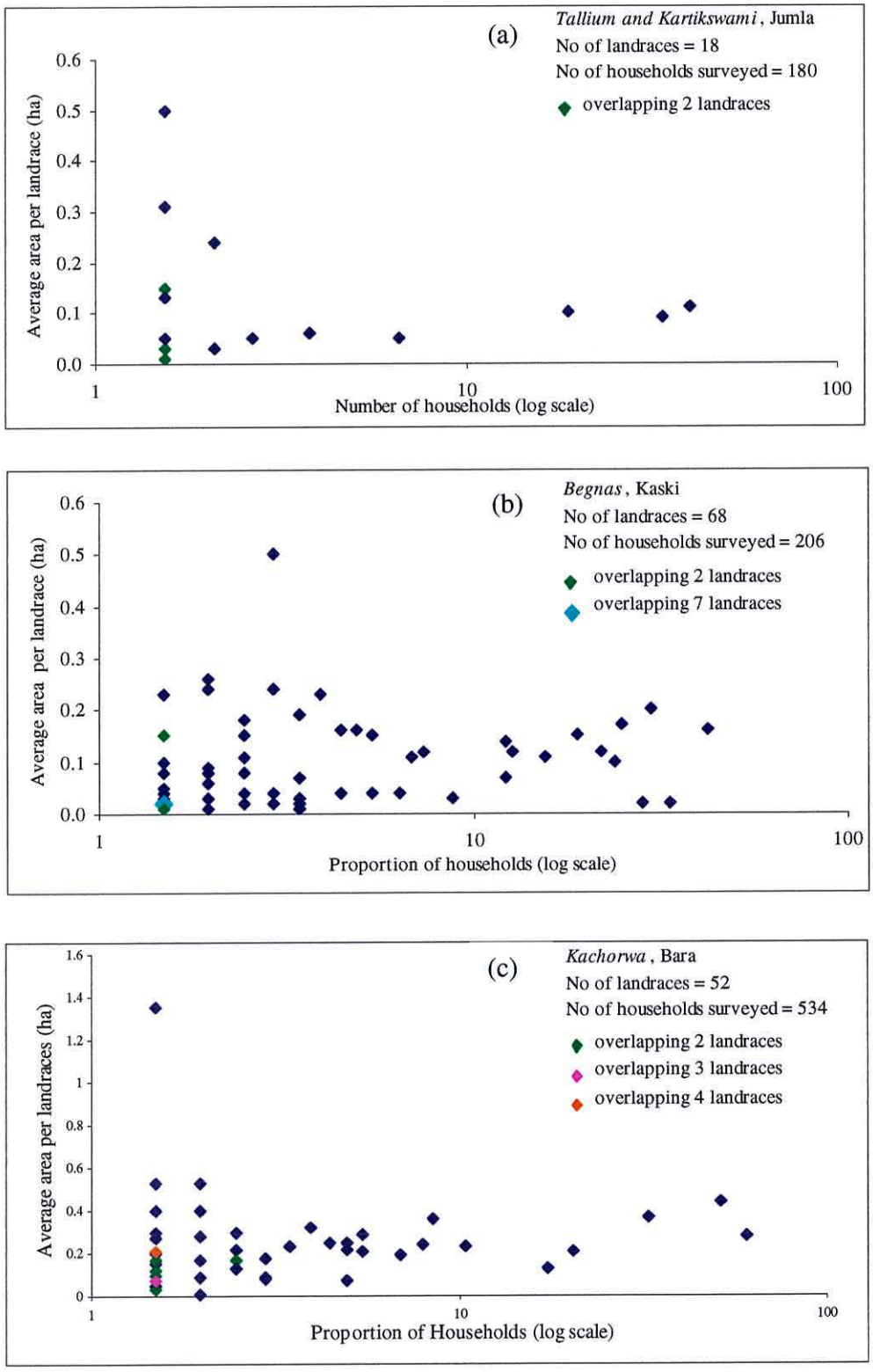


Figure 4.4: Landrace diversity relative to the average area and the number of households growing each landrace in (a) high-hills, (b) mid-hills, and (c) *terai* (lowland).

4.3.3 The social environment – Wealth affects the number of landraces grown by farming households

The average diversity of rice at the household level was highest in the mid-hill ecosite (Kaski) and least in the high-hill ecosite (Jumla) (Figure 4.5 and Table 4.4). The farmers in Kaski maintained the highest average number of landraces (3.8) and a single household could grow up to 22 landraces. A lower diversity was observed in Bara where the households grew an average of 2.7 landraces and up to 12 landraces were grown by a single household. However, in the high-hill ecosite 92% of the sampled households maintained just a single variety on their farm (average 1.1) and 3 was the highest number of landraces grown by a single household (Table 4.4).

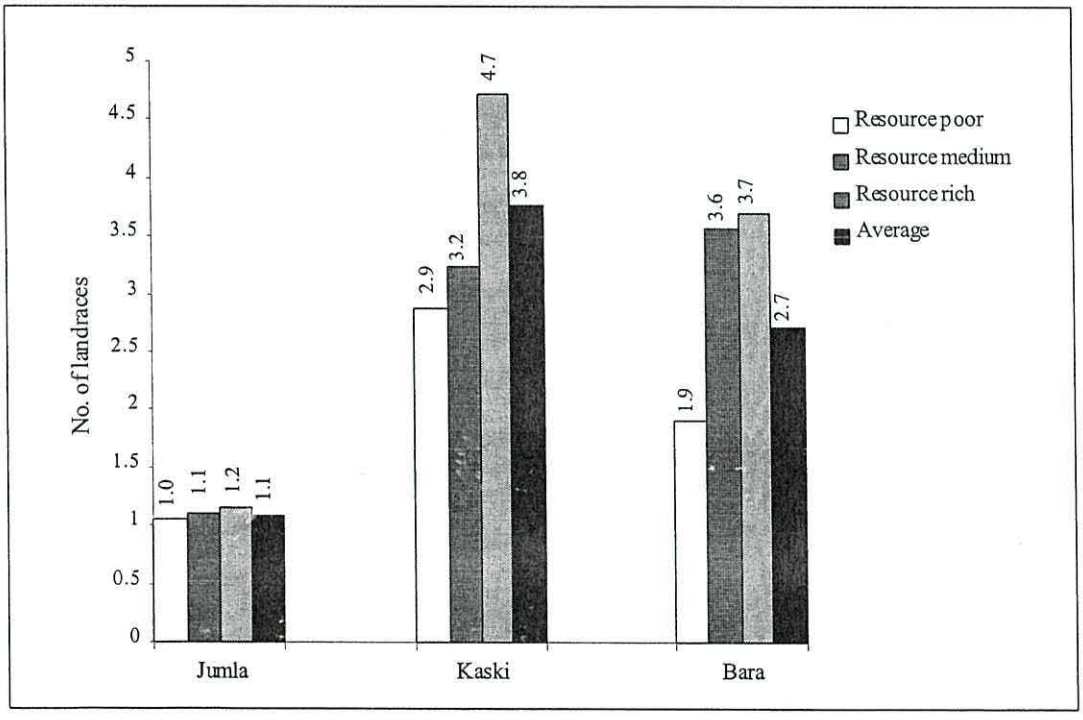


Figure 4.5: Comparative distribution of rice diversity by wealth rank at ecosite level.

Wealth affected the number of landraces that were grown on farm (Figure 4.5). The study showed that resource-poor farmers grew fewer landraces than the resource-rich farmers in Kaski and Bara. In Jumla, however, no significant differences were observed among the wealth categories (a range of 1.1 to 1.2 among the wealth classes). Surprisingly, the only household in Jumla that grew 3 landraces was resource-poor. In Kaski and Bara, although rich farmers grew more landraces there were differences. In Kaski, the resource-rich grew more landraces than the other two wealth categories, in Bara the resource-rich and resource-medium grew about the same, higher number of landraces than the resource-poor (Figure 4.5). Although, across the sites, resource-rich farmers grew and conserved more diversity than the resource-poor farmers, the resource-poor farmers grew specific landraces adapted to their marginal lands.

Table 4.4: Number of rice landraces/households by wealth category at the three ecosites.

Site	Number of landraces	Wealth categories			Total	
		Rich	Medium	Poor	Count	Count (%)
Tallium, Kartikswami	1	34	53	78	165	91.7
Jumla	2	6	6	2	14	7.8
	3	0	0	1	1	0.5
	Total	40	59	81	180	100.0
	Average	1.2	1.1	1.1	1.1	
	Stand. error	0.05	0.04	0.03	0.02	
	Range	1 - 2	1 - 2	1 - 3	1 - 3	
Begnas, Rupakot	1	2	13	12	27	15.5
Kaski	2	7	8	8	23	13.2
	3	9	12	10	31	17.8
	4	10	12	5	27	15.5
	5	13	8	1	22	12.6
	6	7	8	2	17	9.8
	7	5	4	1	10	5.7
	8	2	2	0	4	2.3
	9	2	0	1	3	1.7
	10	3	0	0	3	1.7
	11	2	0	0	2	1.1
	12	1	0	0	1	0.6
	13	2	0	0	2	1.1
	15	1	0	0	1	0.6
	22	1	0	0	1	0.6
	Total	67	67	40	174	100.0
	Average	4.7	3.2	2.9	3.8	
	Stand. error	0.44	0.22	0.33	0.22	
	Range	1 - 22	1 - 8	1 - 9	1 - 22	
Kachorwa	1	2	0	46	58	28.4
Bara	2	5	14	34	53	26.9
	3	7	16	15	38	19.3
	4	2	19	4	25	12.7
	5	3	4	2	9	4.6
	6	1	5	2	8	4.5
	7	2	1	0	3	1.5
	8	0	3	0	3	1.5
	9	1	0	0	1	0.5
	12	0	1	0	1	0.5
	Total	23	73	103	197	
	Average	3.7	3.6	1.9	2.7	
	Stand. error	0.43	0.24	0.11	0.13	
	Range	1 - 9	1-12	1 - 6	1-12	

Rana *et al.*, (2000a,b,c)

4.3.4 The physical environment – landraces are adapted to agro-ecological domains.

The agricultural land was classified according to altitudinal variation, topography, moisture regimes and soil fertility across the study sites.

4.3.4.1 Jumla – the high hill ecosite

In Jumla, farmers classified agro-ecological domains of rice based on the sources of irrigation. These were *Sim khet* (water logged marshy land with poor drainage), *Gadkule khet* (irrigated from snow melted river) and *Kholapani khet* (irrigated with water from stream). Seto marshi, Rato Marshi and Kalo Marshi were three most common varieties and they were grown by most farmers across all three domains. A classification of landraces according to domains was therefore not possible.

4.3.4.2 Kaski – the mid hill ecosite

Farmers in Kaski identified four different domains for rice. These were *Mule khet/Kule khet* (irrigation by seasonal canal), *Sim khet* (marshy wet land), *Tari khet* (rainfed good fertile land) and *Pakho tari* (completely rainfed marginal uplands). *Kule khet* ranked the first in production potential and followed by *Sim khet*, *Tari khet* and *Pakho tari* (Rana *et al.*, 2000b). The distribution and diversity of landraces varied greatly between these domains (Appendix 4.3a, Figure 4.6). *Kule khet* and *Sim khet* were the most favourable domains for rice, and had the greatest diversity of rice. *Tari* and *Pakho tari* were two domains where water was limiting and they were both less productive and less diverse. Out of the 69 landraces, 38% (26 landraces) were specific to a particular domain while 62% were grown in two or more adjacent domains. Jhinuwa, a small-

grained, aromatic rice, was the only one reported to be grown in three domains: *Tari*, *Mule khet* and *Sim khet*.

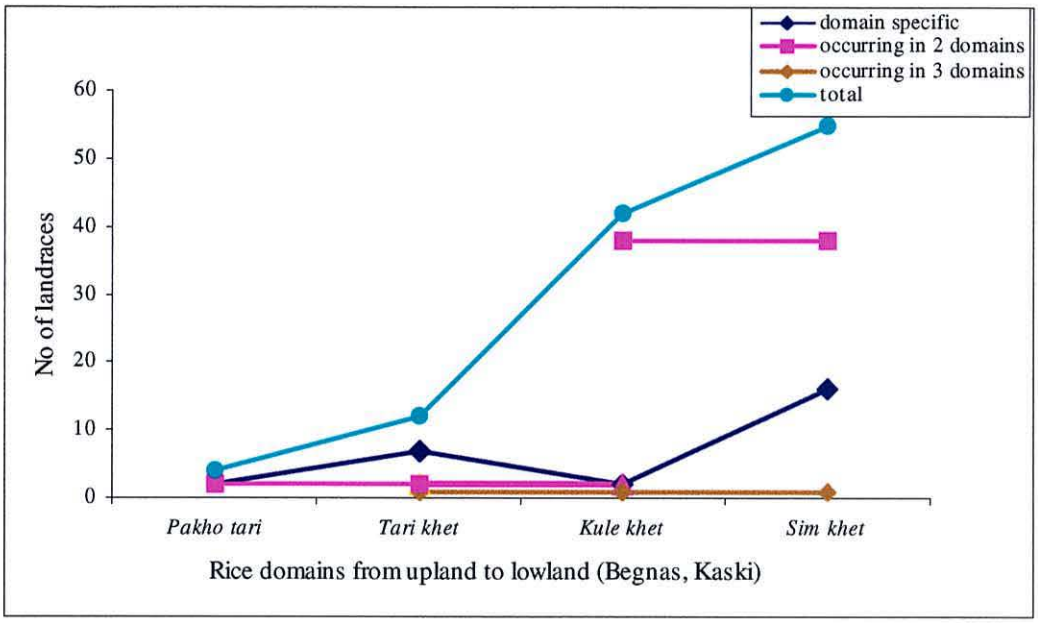
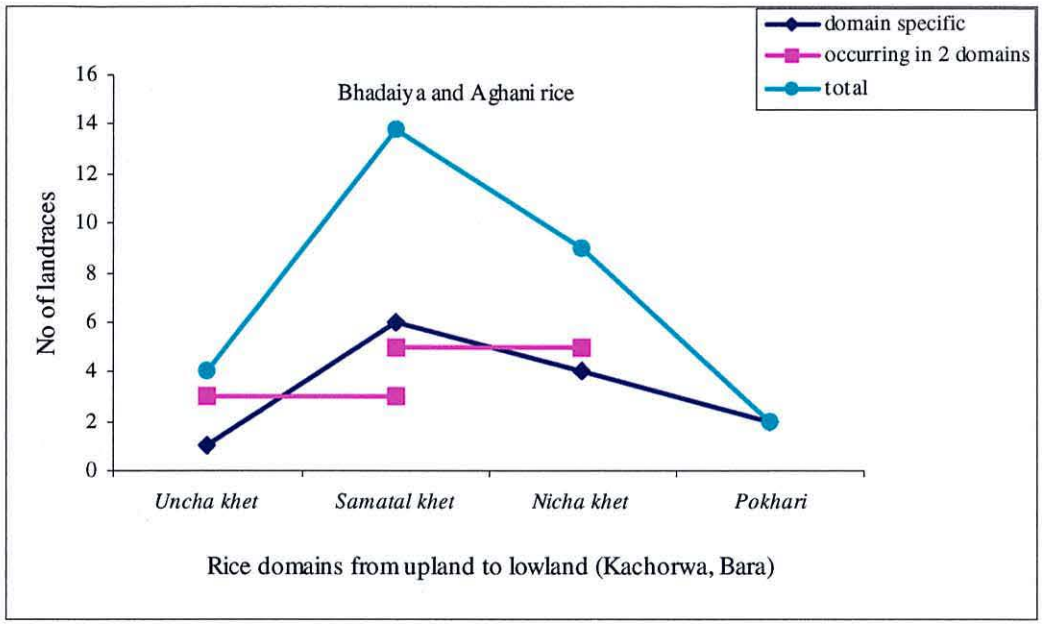


Figure 4.6: Agro-ecological domains and distribution of rice diversity from upland to lowland in Begnas, Kaski.

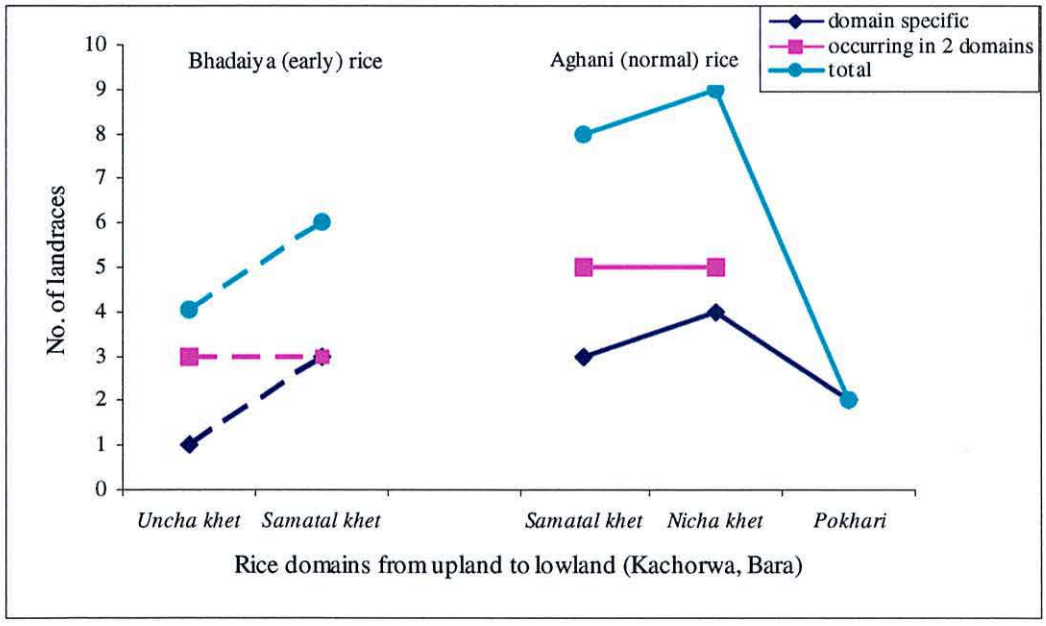
4.3.4.3 Bara - the lowland (*Terai*)

As in Begnas, farmers in Bara also delineated the rice domains mainly on the basis of moisture regime and the fertility status (Rana *et al.*, 2000c). Farmers classified the rice fields into four different domains: *Ucha khet* (rainfed agricultural land), *Samatal khet* (flat land with possible irrigation), *Nicha/khalar khet* (irrigated/wet land) and *Pokhari/Man* (accumulated water as a pond). *Samatal khet* and *Nicha khet* were irrigated rice fields, and were the most productive and the most common domains of the site. On the other hand, *Ucha* and *Pokhari* were marginal domains representing the two extremes of water availability from droughted to flooded land. The distribution and type of rice landraces in these domains were different, and were associated with the adaptive traits of

the landraces (Appendix 4.3b; Figure 4.7). In *Ucha khet*, only Bhadaiya (early-maturing) landraces were cultivated whereas in *Pokhari* only deep root rice varieties were grown. In these two domains, the number of landraces under cultivation was quite limited (Figure 4.7). *Samtal* and *Nicha khet* were the most favourable rice domains and had the greatest diversity of landraces. *Samatal* represents the domain where both Bhadaiya and Aghani rice (early-maturing and normal rice) were grown and was most diverse. However, for normal rice the most favourable domain, *Nicha khet*, had the greatest diversity. Out of 21 landraces reported in the survey, 13 (62%) were specific to domains while 38% were grown in adjacent domains (Appendix 4.3b and Figure 4.7).



(a)



(b)

Figure 4.7: Agro-ecological domains and rice diversity from upland to lowland in Kachorwa, Bara: (a) total rice landraces; (b) Bhadaiya (early) rice and Aghani (normal) rice plotted separately.

4.4 Discussion

The amount and distribution of landrace diversity in rice across the study eco-sites and rice domains within each ecosite varied greatly. The physical environment (agro-ecological domains), social environments, and diversity of uses of the landraces all helped to maintain the genetic diversity of rice.

Among the three ecosites, rice landrace diversity was greatest in the mid-hill Kaski ecosite. It is a mountainous site and is well known for its high quality rice in the Western hills of Nepal (Sthapit *et al.,* 2000). Because of the range in altitude, there is great environmental heterogeneity and diverse agro-ecosystems, and great diversity in the socio-economic structure of the farming communities. At the other extreme, the high-hill Jumla ecosite is the least favourable for rice growing and has the lowest rice diversity with a limited number of named landraces. Chilling temperature is the limiting factor for rice cultivation and Jumli Marshi is the only well known cold-tolerant landrace and is the dominant variety. Bara in the *Terai* is the most fertile and favourable site, lying on the fertile low-altitude strip of the Indo-Gangetic plain. The region is known as the granary of the country with high production potential and is famous for its aromatic rice and its diversity. The environment in Bara is comparatively homogenous lacking altitudinal variability and there has been replacement of landraces by modern varieties.

The study showed on-farm diversity is affected by the physical and social structure of a local environment. Across the ecosites, the physically and socio-economically favourable environments (favourable rice domains and resource-rich farming households) conserved the greatest diversity of rice landraces. The irrigated rice domains: *Kule khet* and *Sim Khet* in Kaski and *Nicha* and *Samatal khet* in Bara had the

most landraces, the majority of which had wide adaptation to adjacent domains. Marginal environments (stress prone domains) had fewer landraces and these had adaptation specific to that particular environment.

Many studies of on-farm conservation have shown that diversity is high in marginal environments and subsistence farmers have maintained diversity to their agro-ecological nitch areas (Harlan. 1975; Brush, 1995; and Brown, 1999). Favourable environments are more homogeneous and have shifted to monoculture or landraces have been replaced by the less diverse modern varieties. However, the results of the present study showed a general agreement with the ecological principle that when environments are more favourable greater diversity is maintained (Witcombe, 1999). This parallels examples of species diversity in natural habitats. Witcombe (1999) also argued that farmers in favourable environments have more options in choosing varieties than farmers in marginal areas. This could be seen in Bara, where farmers had the options to grow varieties with different growth durations where favourable environments (lack of chilling temperature and high water availability) permit farmers to grow two crops a year.

4.5 Conclusions

The physical and socio-economic settings of the traditional farming communities have a great effect on the extent and distribution of rice diversity on farm. The value of the landraces for multiple uses has helped the maintenance of their diversity. Favourable environments conserve a high level of diversity both at the ecosite and the domain levels. Among the ecosites, the mid-hill ecosite had the highest rice diversity with diverse agro-ecosystems. In two of the three sites, the resource-rich farmers were the most important

custodians of the rice varietal diversity on farm, although resource-poor farmers maintained landraces adapted to the most marginal environments.

CHAPTER V

AGRO-MORPHOLOGICAL VARIABILITY STUDY

5.1 Introduction

Characterisation of germplasm helps to estimate the level of genetic variation present in crop varieties and facilitates their further evaluation and use. A number of characters have been used to determine the relationships and diversity present in rice varieties including morphological traits and biochemical and molecular markers.

Agro-morphological characterization is a simple, inexpensive method in which differences in quantitative and qualitative characters are measured and analysed. There are many examples of studies on morphological and phenological variation in the germplasm of many crop species and their relatives, including wild rice, and they have been used to determine the centers of diversity of crop species (Witcombe and Gilani, 1979; Juliano *et al.*, 1998; Furman *et al.,* 1997; Tranquilli *et al.,* 2000; Ruiz *et al.,* 1997; and Autrique *et al.*, 1996).

Agro-morphological characters, including economically important traits such as grain yield, taste, aroma, and maturity, are used by farmers to describe and identify their traditional landraces. The present study includes the systematic measurement of agro-morphological characteristics of the rice landraces from the three ecosites. The variability in qualitative and quantitative agro-morphological traits was related to farmers' given names and to rice agro-ecosystems within the three sites.

5.2 Materials and methods

The accessions included in the morphological study are described in detail in Chapter III, Section 3.1.1. The field observations were carried out individually in three different

environments similar to the environment of their collection sites. The details of experimentation and observations measured were described in Chapter III, Section 3.2.1 and 3.2.2. Twenty eight qualitative traits in the Jumla landraces, 29 in Kaski and 22 in Bara rice landraces were recorded (Table 3.7). Fifteen quantitative traits were measured in all the landraces from all the sites. These quantitative and qualitative traits were recorded and measured following the IRRI-IPGRI descriptors for rice (IRRI/IBPGR, 1980).

The traditional practice of seeding was followed in Jumla. The seeds for sowing were presoaked in warm water and wrapped in moist cotton cloth and were laid somewhere in the kitchen for 4 days to a week-long period. The farmers in Jumla start this on the 12th Chaitra (25 March) every year. The sprouts were then broadcasted in the seedbed. This is the traditional practice of breaking dormancy in rice seeds in the high-hills.

The observed morphological variations were compared among rice accessions within each site, including the respective check varieties, using univariate and multivariate analyses. The comparisons of landraces across the ecosites were not made in this chapter because of the confounding effect of environment with genotype. The observations were analysed separately and together for quantitative and qualitative traits by site: i.e. at Jumla, Kaski or Bara using principal component analysis, cluster analysis and discriminant function analysis. Only the most informative results are presented for discriminant and cluster anlyses. Table 5.1 below shows the traits used for these analyses for each ecosite.

Table 5.1: Summary of the traits used for all of three multivariate analyses across the ecosites.

Traits	Jumla		Kaski		Bara	
	1 and 2	3	1	2 and 3	1	2 and 3
(a) Quantitative traits						
Leaf length	-	-	+	+	+	+
Leaf width	+	+	+	+	-	-
Length width ratio	-	-	+	+	+	+
Ligule length	-	-	+	+	+	+
Culm length	-	-	+	+	+	+
Tillers per plant	-	-	+	+	+	+
Panicle length	-	-	+	+	+	+
Panicle exsertion	-	-	+	+	+	+
Days to heading	-	-	+	+	+	+
Days to maturity	-	-	+	+	+	+
1000 grain weight	-	-	+	+	+	+
Grains per panicle	+	+	+	+	+	+
Empty hulls per panicle	+	+	+	+	+	+
Grain weight per panicle	+	+	+	+	+	+
Yield per plot	-	-	-	-	+	+
(b) Qualitative traits						
Seedling vigour	-	-	+	+	-	-
Leaf blade colour	-	-	-	-	+	+
Leaf blade pubescence	-	-	+	+	+	+
Flag leaf angle	-	-	+	+	+	+
Collar colour	-	-	-	-	+	-
Node colour	-	-	-	-	+	-
Internode colour	-	-	+	+	-	-
Culm angle	-	-	-	-	+	+
Culm strength	-	-	+	+	+	+
Panicle type	-	-	-	-	+	+
Panicle axis	-	-	+	+	+	-
Secondary branch	+	+	+	+	+	+
Shattering	+	+	+	+	-	-
Threshability	-	-	+	+	-	-
Leaf senescence	+	-	-	-	-	-
Awning	+	+	+	+	+	-
Stigma colour	+	+	-	-	-	-
Apiculus colour	-	-	+	+	+	-
Lemma and palea colour	+	+	+	+	+	-
Lemma and palea pubes.	-	-	-	+	+	+
Sterile lemma colour	+	+	+	+	+	-
Sterile lemma length	-	-	+	+	-	-
Grain size	-	-	+	+	+	+
Grain shape	-	-	+	+	+	-
Grain type	-	-	+	-	+	-
Bran colour	+	+	-	-	+	-

1 = Principal component analysis; 2 = Cluster analysis; 3 = Discriminant function analysis
+ = traits included; and - = traits not included.

Results

5.3.1 Distribution and expression of quantitative traits

For each of the 15 quantitative traits the mean, range (maximum and minimum), standard deviation, and coefficient of variation (CV) were calculated (Tables 5.2). All the quantitative traits showed a range of continuous variation approximated to a normal distribution with few exceptions (Figure 5.1-5.2). The exceptional traits were the phenological traits: days to heading (DH) and days to maturity (DM) in the Kaski and Bara ecosites (Figure 5.2). At Kaski, DH and DM were skewed towards lateness, and the same skewness was found in the normal and late-maturing groups in Bara. At Jumla, fewer class intervals were observed due to a lower range in DM.

Most of the quantitative traits were highly variable among the landraces in Bara and Kaski with a wide range between minimum and maximum values (Table 5.2). However, the landraces in the high-hill site (Jumla) were morphologically less diverse and had lower CVs for most of these traits

There were significant differences among the landraces, as measured by reapitability within each site for most of the quantitative traits: 0.02 – 0.6 repeatability for Jumla landraces; 0.3 – 1.0 for Kaski and 0.2 – 1.0 for Bara (Table 5.3). However, in Jumla, only few traits were significant including even phenological ones.

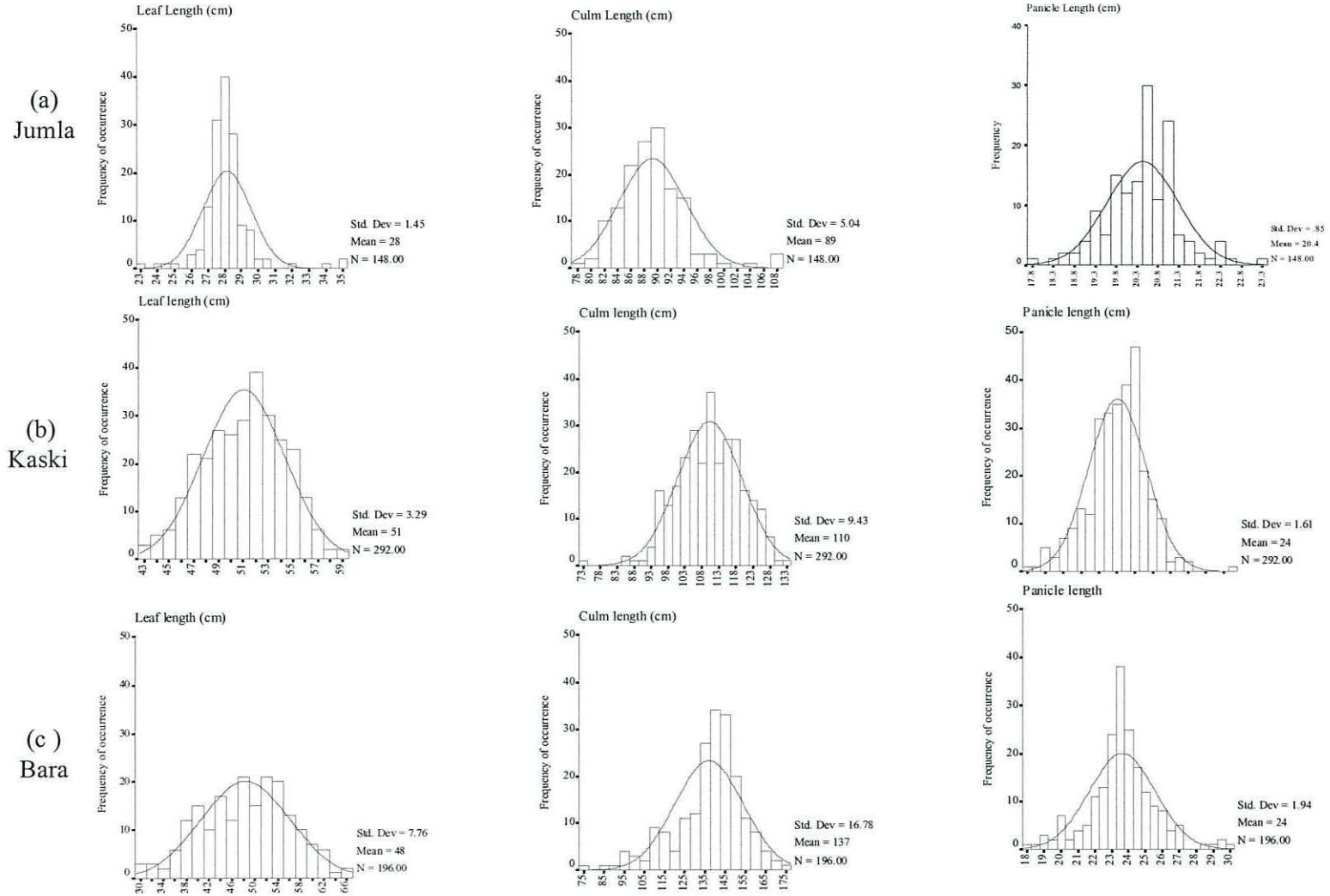


Figure 5.1: Frequency distribution of quantitative traits in rice landraces of three study sites: (a)Jumla; (b) Kaski and (c) Bara

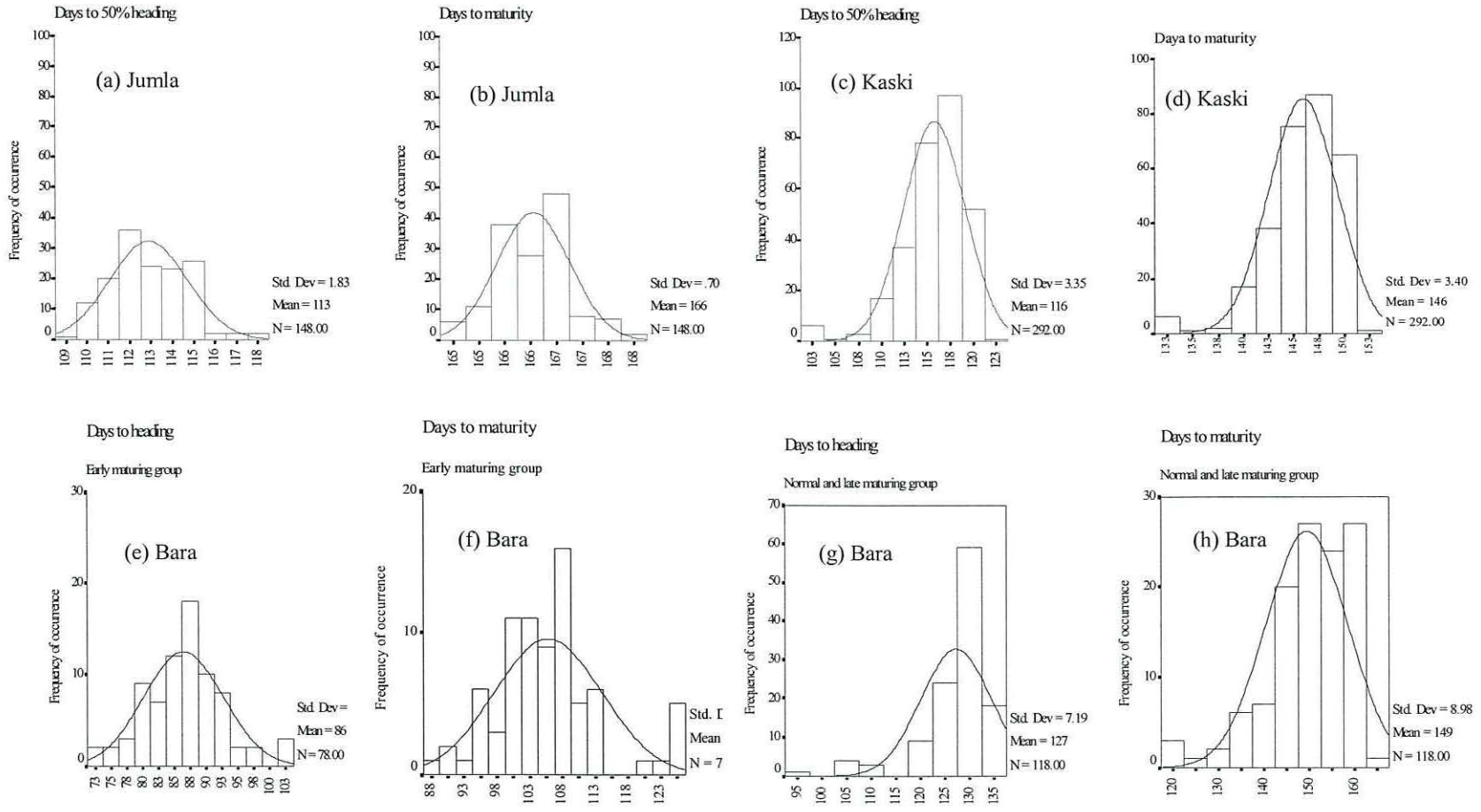


Figure 5.2: Frequency distribution of phenological traits in rice landraces showing (a) and (b) for Jumla, (c) and (d) for Kaski and (e), (f), (g), and (h) for Bara.

Table 5.2: Mean, minimum, maximum, standard deviation and coefficient of variation for each quantitative traits in 148, 292 and 196 rice accessions under the field trial in ARS†, Jumla; ARS, Kaski and NRRP‡, Janakpur in 2000-2001 during normal season of rice.

Traits		Jumla	Kaski	Bara
Leaf length (cm)	mean	27.6	51	48
	standard deviation	0.1	3	8
	minimum	22.7	43	29
	maximum	34.7	59	66
	CV %	0.43	6	16
Leaf width (cm)	mean	1.1	1.1	1.1
	standard deviation	0.01	0.1	0.01
	minimum	0.92	0.8	0.69
	maximum	1.6	1.5	1.6
	CV %	0.91	8.9	0.89
Length/width ratio	mean	26	46	45
	standard deviation	2	5	10
	minimum	21	35	23
	maximum	34	58	79
	CV %	7	10	23
Ligule length (cm)	mean	1.0	1.7	1.8
	standard deviation	0.01	0.2	0.03
	minimum	0.9	1.0	1.1
	maximum	1.2	2.3	3.0
	CV %	1.0	12.6	1.7
Culm length (cm)	mean	89	110	137
	standard deviation	5	9	17
	minimum	79	73	74
	maximum	108	133	173
	CV %	6	9	12
Tillers per plant	mean	8	5	8
	standard deviation	1	1	2
	minimum	7	4	5
	maximum	16	10	16
	CV %	14	21	24
Panicle length (cm)	mean	20	24	24
	standard deviation	0.9	2	2
	minimum	18	19	18
	maximum	23	31	30
	CV %	4	7	8
Panicle exsertion (cm)	mean	10	-	4
	standard deviation	1	-	1
	minimum	8	-	2
	maximum	13	-	10
	CV %	11	-	35
Well filled grains per panicle	mean	70	120	105
	standard deviation	10	33	24
	minimum	52	54	47
	maximum	156	231	159
	CV %	15	28	22
Empty hulls per panicle	mean	17	32	15
	standard deviation	2	16	6
	minimum	11	10	4
	maximum	29	174	40
	CV %	14	59	43

Yield per panicle (g)	mean	2.0	2.3	2.7
	standard deviation	0.3	0.5	0.8
	minimum	1.4	1.0	1.2
	maximum	4.1	3.7	5.3
	CV %	15.9	22.8	29.4
1000 grains weight (g)	mean	27	21	25
	standard deviation	2	6	4
	minimum	22	13	15
	maximum	31	38	37
	CV %	7	27	17
Yield per plot (g)	mean	228	110	222
	standard deviation	37	42	78
	minimum	142	34	84
	maximum	334	313	531
	CV %	16	40	37
Days to heading	mean	113	116	111
	standard deviation	2	3	21
	minimum	109	102	72
	maximum	118	123	136
	CV %	2	3	17
Days to maturity	mean	166	146	132
	standard deviation	0.7	3	23
	minimum	165	132	87
	maximum	168	153	165
	CV %	0.4	2	16

† ARS = Agriculture research station;
‡ NRRP = National rice research programme;

Table 5.3: F- values and repeatabilities of the quantitative traits of the rice accessions from the three ecosites.

Traits	Jumla		Kaski		Bara	
	F-value	R†	F-value	R	F-value	R
Leaf characteristics:						
2nd leaf lengtrh (LLENGTH)	ns	0.07	***	0.46	***	0.64
2nd leaf width (LWIDTH)	**	0.24	***	0.56	ns	0.16
Length/width ratio (LWRATIO)	ns	0.11	***	0.53	***	0.51
Ligule characteristics:						
Ligule length (LIGLEN)	ns	0.02	***	0.38	***	0.57
Culm characteristics:						
Culm length (CULMLEN)	ns	0.00	***	0.51	***	0.76
Tillers per plant (TILLERS)	ns	0.14	***	0.60	***	0.55
Panicle characteristics:						
Panicle length (PANLEN)	ns	0.05	***	0.27	***	0.67
Panicle exsertion length (PANEXSE)	ns	0.13	***	0.63	***	0.72
Post harvest characteristics:						
Grains per panicle (GRAINP)	***	0.60	***	0.58	***	0.59
Empty hulls per panicle (HULLSP)	***	0.41	***	0.35	***	0.42
Grain yield per panicle (GWP)	***	0.50	***	0.36	***	0.73
1000 grains weight (TGW)	ns	0.17	***	0.99	***	0.99
Yield per plot (YIELDP)	ns	0.15	***	0.90	*	0.23
Phenological traits:						
Days to 50% heading (DH)	ns	0.12	***	0.78	***	0.92
Days to maturity (DM)	ns	0.03	***	0.77	***	0.91

† R = repeatability; ns = not significant; * p = >0.05; ** p = >0.01; *** p = >0.001.

5.3.2 Distribution and expression of qualitative traits

The phenotypic traits like colour, pubescence, panicle type and awning were considered as qualitative traits and were scored using multistate classification. In genetic terms, although these traits may be controlled by more than one gene, their phenotypic expression is little influenced by the environment. These are listed in Tables 5.4. Traits like shattering, threshability, culm angle, panicle axis were also treated as qualitative traits with different character states as shown in Table 3.7 in Chapter III and are separately listed in Table 5.5.

In Jumla, most of these qualitative traits were non-variable i.e. were represented by only one category, but this category was different from the check variety. There were seven non-variable qualitative traits in Jumla landraces and one each in Kaski and Bara (Table 5.3). Even in variable traits, there were fewer classes in Jumla and most accessions had the phenotype of one class. For example, the leaf blade colour, ligule colour, apiculus colour, lemma and palea colour and sterile lemma colour were mostly represented by one class (Table 5.4). In Kaski and Bara, these traits together with grain characters were diverse among the landraces and were represented by two to many classes (Table 5.4). The analysis of variance determined the significant differences among the landraces for qualitative traits (Table 5.5). The true qualitative traits determined by major genes with high stability of expression had a repeatability of 1, or nearly 1 (Table 5.5). Such traits included apiculus colour, sterile lemma colour, grain shape, and bran colour. But the values were very low at Jumla. However, the repeatability for culm angle was 1 at Jumla and 0.98 at Bara but very low at Kaski. Some of these very low value for repeatability may be because of recording errors or sampling errors between replicates. Other traits, such as shattering, and threshability are likely to be quantitative traits, that have been recorded in a qualitative (categorical) fashion so they have repeatabilities less than 1 at all sites where they were measured.

Table 5.4: Frequencies, calculated as percentages of the phenotypic classes, for 14 qualitative traits in the rice accessions that are common across three ecosites.

Traits	Class	Jumla	Kaski	Bara
Leaf blade colour	1 Pale green	87	07	71
	2 Green	13	82	20
	3 Dark green	-	11	09
Flag leaf angle	1 Erect	-	28	34
	3 Intermediate	-	37	19
	5 Horizontal	100	31	16
	7 Descending	-	04	31
Ligule colour	1 White	26	88	100
	2 Purple	74	12	-
Culm angle	1 Erect	-	84	48
	3 Intermediate	100	14	27
	5 Open	-	01	20
	7 Spreading	-	01	05
Culm strength	1 Strong	-	51	04
	3	-	28	47
	5	100	12	43
	7	-	06	06
	9 very weak	-	03	-
Panicle type	1 Compact	-	100	48
	5 Intermediate	100	-	44
	9 Open	-	-	08
Panicle axis	1 Straight	-	07	30
	2 Droopy	100	93	70
Awning	0 Awnless	94	74	60
	1 Long awned	06	17	36
	2 Spiculed	-	09	04
Apiculus colour	1 White	-	06	07
	2 Straw	06	28	36
	3 Tawny	-	16	22
	4 Red	-	06	13
	6 Purple	94	39	05
	7 Black	-	05	13
Lemma and palea colour	0 Straw	01	38	23
	2 Brown spots	-	02	-
	3 Brown furrows	-	02	11
	4 Tawny	-	15	11
	5 Reddish	-	02	30
	7 Purple furrows	15	19	-
	8 Purple	84	-	-
	9 Black	-	06	15
	99 White	-	16	10
				10

Sterile lemma colour	1 Straw	08	72	63
	2 White	-	01	31
	3 Red	-	01	-
	4 Purple	92	21	06
	5 Black	-	05	-
Grain size (length)	1 Short	-	38	30
	2 Medium	100	49	40
	3 Long	-	11	28
	4 Ex. Long	-	02	02
Grain shape (quality)	1 Round (coarse)	-	30	23
	2 Bold (medium)	100	41	53
	3 Slemder (fine)	-	29	24
Bran colour	1 White	52	87	43
	5 Red	45	04	57
	8 Greenish	03	09	-

Table 5.5: F-values and repeatabilities of 28 qualitative traits of the rice accessions from the three ecosites.

Traits	**Jumla**		**Kaski**		**Bara**	
	F-value	**R†**	**F-value**	**R**	**F-value**	**R**
(a) Traits in Table 5.4						
Leaf blade colour (BLACOL)	ns	0.01	ns	0.03	***	0.98
Flag leaf angle (FLANGLE)	ns	0.15	***	0.57	***	0.99
Ligule colour (LIGCOL)	ns	0.01	ns	0.10	ns	0.00
Culm angle (CANGLE)	***	1.00	*	0.15	***	0.98
Culm strength (CULMST)	***	1.00	***	0.40	***	0.98
Panicle type (PANTYPE)	***	1.00	ns	0.07	***	0.97
Panicle axis (PANAXIS)	ns	0.00	***	0.77	***	0.92
Awning (AWNING)	***	0.69	***	1.00	***	1.00
Apiculus colour (APICOL)	ns	0.00	***	1.00	***	1.00
Lemma and palea colour (LPCOL)	***	0.72	***	1.00	***	1.00
Sterile lemma colour (STLCOL)	***	1.00	***	1.00	***	1.00
Grain size (GSIZE)	ns	0.00	***	0.99	***	1.00
Grain shape (GSHAPE)	ns	0.00	***	1.00	***	1.00
Bran colour (BRANCOL)	***	1.00	***	1.00	***	1.00
(b) Traits not in Table 5.4						
Seedling vigour	-	-	***	1.00	-	-
Leaf blade pubescence (BLAPUB)	ns	0.03	***	0.52	***	0.98
Basal leaf sheath colour (SHCOL)	*	0.20	ns	0.03	-	-
Leaf angle (LANGLE)	ns	0.01	ns	0.07	-	-
Ligule shape (LIGSHA)	ns	0.02	ns	0.00	ns	0.00
Collar colour (COLCOL)	ns	0.16	ns	0.03	***	1.00
Node colour (NODCOL)	ns	0.11	ns	0.00	***	0.99
Internode colour (INTDCOL)	-	-	***	0.40	-	-
Secondary branching (SECBRAN)	***	1.00	***	0.53	***	0.93
Shattering (SHATTER)	***	0.38	***	0.38	-	-
Threshability (THRESHA)	***	0.81	***	0.50	-	-
Stigma colour (STGCOL)	***	1.00	-	-	-	-
Sterile lemma length (STLLEN)	ns	0.00	***	1.00	***	1.00
Leaf senescence (LEAFSEN)	***	1.00	-	-	-	-

† R = repeatability; ns = not significant; * p = >0.05; ** p = >0.01; *** p = >0.001.

The estimates of Shannon-Weaver diversity indices (H') for each qualitative trait and the average diversity pooled over all the traits showed the diversity in the characters and landraces of the ecosites (Table 5.6). High values indicate high morphological diversity for a trait and the average diversity value (mean H') shows the average level of diversity across the traits within each of the three ecosites. The Jumla landraces were the least diverse as they had low values for most of the qualitative traits (Figure 5.3). The landraces from Bara and Kaski were highly diverse for apiculus colour, lemma and palea colour and grain characteristics. The average diversity for qualitative traits tended to decrease from the high-hills to the lowlands (Table 5.6).

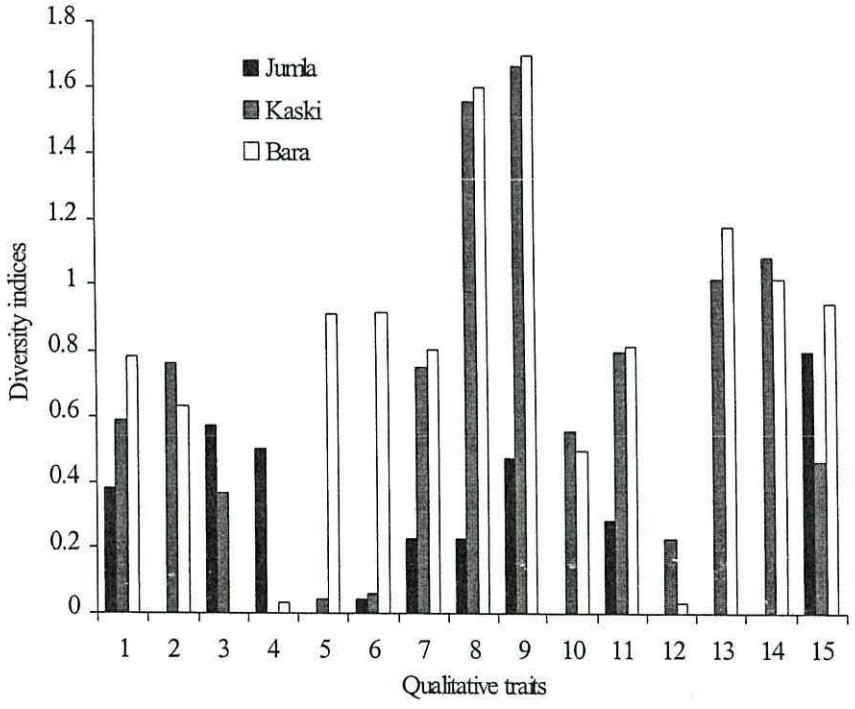


Figure 5.3: Shannon-Weaver diversity indices for selective qualitative traits (colour, type and pubescence) and their comparison among the rice landraces from three ecosites in the same order as in Table 5.6.

Table 5.6: Shannon-Weaver diversity indices (H') for selective qualitative traits (colour, pubescence and type) which were common across the three ecosites.

	Traits	Jumla n = 148	Kaski n = 292	Bara n = 196
1	Leaf blade colour	0.38	0.59	0.78
2	Leaf blade pubescence	0.00	0.76	0.63
3	Ligule colour	0.57	0.36	0.00
4	Collar colour	0.50	0.00	0.03
5	Node colour	0.00	0.04	0.91
6	Panicle type	0.04	0.81	0.92
7	Awning	0.23	0.75	0.80
8	Apiculus colour	0.23	1.56	1.60
9	Lemma and palea colour	0.47	1.67	1.70
10	Lemma and palea pubes.	0.00	0.55	0.49
11	Sterile lemma colour	0.28	0.80	0.82
12	Sterile lemma length	0.00	0.23	0.03
13	Grain size (length)	0.00	1.02	1.18
14	Grain shape (quality)	0.00	1.09	1.02
15	Bran colour	0.80	0.46	0.95
	H' (Mean) †	**0.23**	**0.72**	**0.79**

† = Mean of Shannon Weaver diversity index as average diversity (n = number of accessions)

5.3.3 Correlation between morphological traits

There was a significant correlation between many of the traits when measured separately for the three ecosites (Appendix 5.1). Across the ecosites, generally the lowest correlations were between leaf traits with culm and grain traits. This correlation analysis was primarily carried out as a procedure to reduce the data for subsequent multivariate analysis. For example, in the correlations among the 16 qualitative and quantitative traits measured on rice accessions from Jumla, the highest correlations were between the traits: culm angle, culm strength, panicle type and secondary branching (Table 5.7). Since they were perfectly correlated, out of these four traits only secondary branching was included for further analyses.

Table 5.7: Phenotypic correlations among 16 qualitative and quantitative traits measured on 148 rice accessions from the Jumla ecosite.

Traits	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Leaf width	1.00															
2. Culm angle	-0.02	1.00														
3. Culm strength	-0.02	1.00	1.00													
4. Panicle type	-0.02	1.00	1.00	1.00												
5. Secondary branching	0.02	-1.00	-1.00	-1.00	-1.00											
6. Shattering	0.23	0.62	0.62	0.62	0.62	1.00										
7. Threshability	-0.02	1.00	1.00	1.00	1.00	0.62	1.00									
8. Leaf senescence	0.02	-1.00	-1.00	-1.00	-1.00	-0.62	-1.00	1.00								
9. Awning	-0.12	0.02	0.02	0.02	0.02	-0.16	0.02	-0.02	1.00							
10. Stigma colour	0.09	0.34	0.34	0.34	0.34	0.39	0.34	-0.34	-0.90	1.00						
11. Lemma and palea colour	0.09	0.60	0.60	0.60	0.60	0.54	0.60	-0.60	-0.69	0.92	1.00					
12. St. lemma colour	0.02	0.28	0.28	0.28	0.28	0.28	0.28	-0.28	-0.72	0.80	0.74	1.00				
13. Bran colour	-0.15	0.08	0.08	0.08	0.08	-0.01	0.08	-0.08	0.22	-0.17	-0.11	-0.19	1.00			
14. Grains panicle $^{-1}$	0.17	-0.68	-0.68	-0.68	-0.68	-0.42	-0.68	0.68	0.01	-0.26	-0.45	-0.21	-0.09	1.00		
15. Empty hulls $^{-1}$	0.07	-0.70	-0.70	-0.70	-0.70	-0.26	-0.70	0.70	-0.07	-0.16	-0.33	-0.16	-0.09	0.42	1.00	
16. Grains weight $^{-1}$	0.17	-0.59	-0.59	-0.59	-0.59	-0.39	-0.59	0.59	0.18	-0.38	-0.51	-0.34	-0.10	0.90	0.36	1.00

Note: the correlation coefficients > 0.17 were significant at p = 0.05 and >0.28 at p = 0.001 levels.

5.3.4 Multivariate analysis for Jumla – high-hill ecosite

5.3.4.1 Principal component analysis

A total of 12 qualitative and quantitative traits were analysed. Only significant traits and only one trait of the set of perfectly correlated traits were included, The principal component analysis (PCA) was made for several combinations of traits: PCA of qualitatively scored traits alone (8 traits); PCA of quantitative traits alone (4 traits); and PCA of qualitative and quantitative traits together. The outputs of the anlyses are presented in the Appendix 5.2. The first two principal components (PC) explained 82.1% of total variation for PCA of qualitative traits alone, 75.4% for quantitative traits alone, and 65.7% for qualitative and quantitative traits together. Hence, a high proportion of the variability was accounted for in the PCA by qualitative traits alone. In the PCA of qualitative and quantitative traits together, the variability was mostly contributed by the qualitative traits. These traits were the awning, stigma colour and sterile lemma colour with high loading values in first two axes of PCA (Appendix 5.2).

Scatter plots were produced for the first two PCs of each PCA analysis (Figure 5.4a, b,c). Most rice accessions, though they had been collected with different names, were found to be morphologically similar and clustered close together with few exceptional accessions, which were dispersed in each scatter plot. The Jumla landraces were very distinct from the standard check variety and were discriminated by the first principal component, in all three analyses. All the awned accessions in the study formed two small but distinct groups of seven and two accessions that differed from each other for stigma colour, and they were separated by second principal component from the major awnless group of accessions (Figure 5.4b and data shown in Appendix 5.2). There

was a strong correlation between the absence of awns and the colour of the stigma and sterile lemma. All the awnless accessions had coloured (purple) stigma and coloured sterile lemma whereas the awned accessions had white stigmas apart from two accessions, 1035-1 and 1036-1, which grouped separately in the plot (Figure 5.4a and b).

Three accessions 1024-1, 1064-1 and 2017-1 were dispersed from rest of the accessions and were also dispersed apart from each others. These accessions were awnless accessions which differed for lemma and palea colour, sterile lemma colour and bran colour. 1024-1 and 2017-1 had an uncoloured sterile lemma, whereas 1064-1 had a purple sterile lemma. 1024-1 differed from 2017-1 as the latter had red rice bran rather than white. A low morphological variability was observed in Jumla landraces and awning, stigma colour and sterile lemma colour were the important traits in discriminating the variation. Nearly all the variation was due to the qualitative trait variation and indeed, adding quantitative data simply made the relationships less clear (Figure 5.4b compared to Figures 5.4a and c).

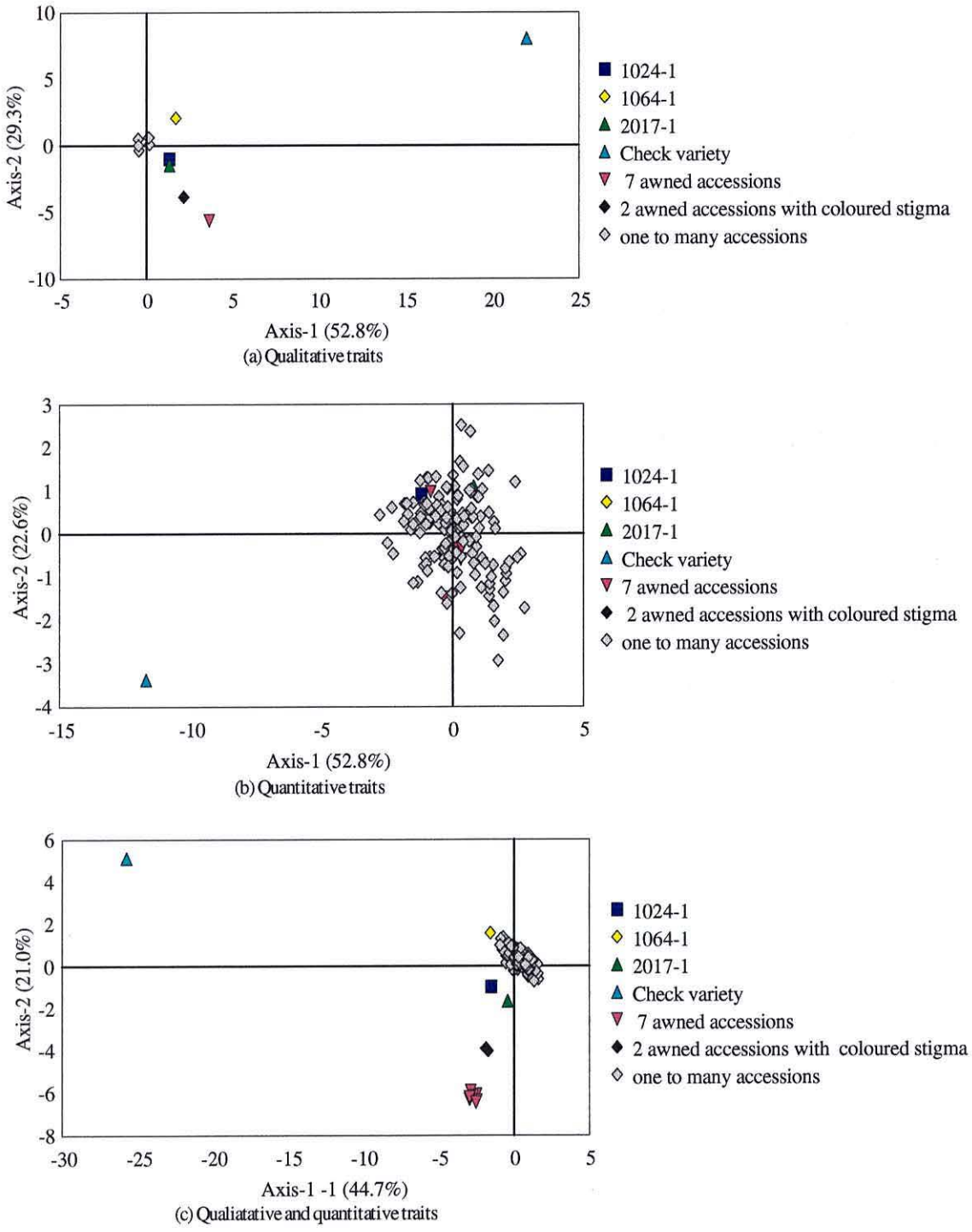


Figure 5.4: Scatter diagram of 148 rice accessions from the Jumla ecosite classified along the axes 1 and 2 of the ordination for (a): 8 qualitative traits; (b): 4 quantitative traits; and (c): 12 qualitative and quantitative traits combined.

5.3.4.2 Cluster analysis

The cluster analysis, based on 12 significant qualitative and quantitative traits together grouped the 148 Jumla accessions into 4 morphological clusters at a dissimilarity level of 3.5 (Figure 5.5). Exactly the similar groupings of the accessions into respective 4 clusters were observed in cluster analysis using qualitative traits alone. The number of accessions per cluster varied from the single check variety in cluster IV, to 72 in cluster II (Table 5.8). Cluster III mostly contained the 9 awned accessions and accessions 1024-1, 1064-1 and 2017-1, which were also dispersed from the major group of accession in the PCA using qualitative and quantitative traits together and qualitative traits alone (Figure 5.7). On the whole, the rice landraces showed a close similarity in their clustering pattern but were distinct from the check variety: this can be seen from the mean values of the clusters and their relative distance from each other and from the centroid (Appendix 5.3).

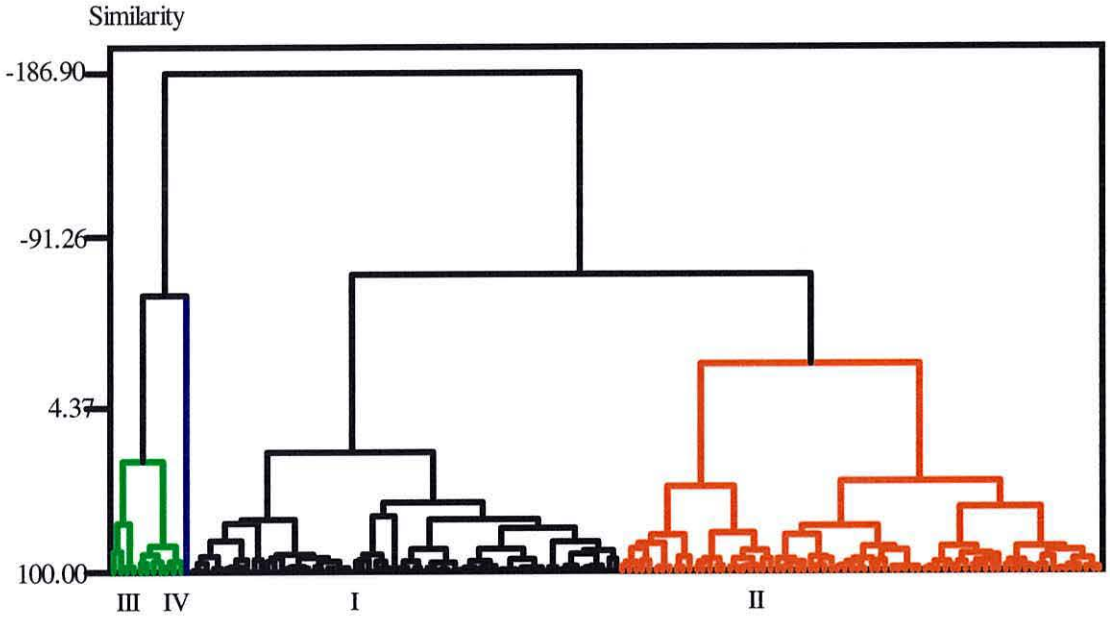


Figure 5.5: Dendrogram showing the clustering pattern of 148 rice accessions from the Jumla ecosite based on qualitative and quantitative traits. (Each colour represents a cluster.)

Table 5.8: Distribution in the cluster analysis of the Jumla rice accessions in 4 clusters by FUDs (local names).

Landrace	Total accessions	I	II	III	IV
Seto Marshi	52	28	20	4	-
Rato Marshi	49	24	20	5	-
Kalo Marshi	21	4	17	-	-
Mehele	12	3	8	1	-
Darime	6	1	4	1	-
Dhan	3	2	1	-	-
Palte Dhan	1	1	-	-	-
Seto Dhan	1	-	1	-	-
Rato Dhan	1	-	1	-	-
Jumli Rato Marshi	1	1	-	-	-
Check variety	1	-	-	-	1
Total of clusters	148	64	72	11	1

5.3.4.3 Discriminant function analysis

The results of the three way discriminant function analyses showed that the analysis using qualitative and quantitative traits together and the qualitative traits alone were most informative explaining high variation among the landraces. 36% accessions were correctly classified to their respective FUDs and 26% after cross validation in the analysis based on qualitative and quantitative traits together.

Table 5.9: Summary of discriminant function analysis of the Jumla rice accessions using 12 qualitative and quantitative traits.

Group	No of accessions per group	Prop. of discriminant	Prop. of cross validation
By landrace group:			
Darime	6	67	0
Kalo Marshi	21	62	57
Mehele	12	25	8
Rato Marshi	49	27	22
Seto Marshi	52	35	23
Total	**140†**	**36**	**26**
By agro-morphological cluster groups:			
Cluster I	64	98	NA
Cluster II	72	88	NA
Cluster III	11	82	NA
Cluster IV	1	NA	NA
Total	**148**	**92**	**NA**

† = The landraces represented by one accession were excluded.

5.3.5 Multivariate analysis for Kaski – mid-hill ecosite

5.3.5.1 Principal component analysis

In the principal component analyses on qualitative and quantitative traits carried out separately and together. The quantitative traits were the most informative (45% of the variation), the qualitative traits alone the next most informative (40% of the variation), and the combined qualitative and quantitative traits the least informative (36% of the variation). The most important traits were panicle exsertion, thousand grain weight, leaf

width, leaf length width ratio, grain number per panicle and yield per panicle. Lemma and palea colour and grain types contributed most in the PCA that used qualitative traits alone (Appendix 5.4). There was no clear discrimination of the landraces either by agro ecosystem in any of the analyses (Figure 5.6). However, the landraces specific to the upland domain (*Pakho tari*) formed a more or less distinct group in all three analyses (Figure 5.6). Similarly, the accessions of Anadi and Aanga formed distinct groups from the rest of the accessions. They were morphologically distinct with tall and thick culms, broad leaves, and bold grains and were adapted to the *Sim* and *Kule khet* domains. Since distinction by agro-ecosystem was unclear, discrimination by FUD was even less so.

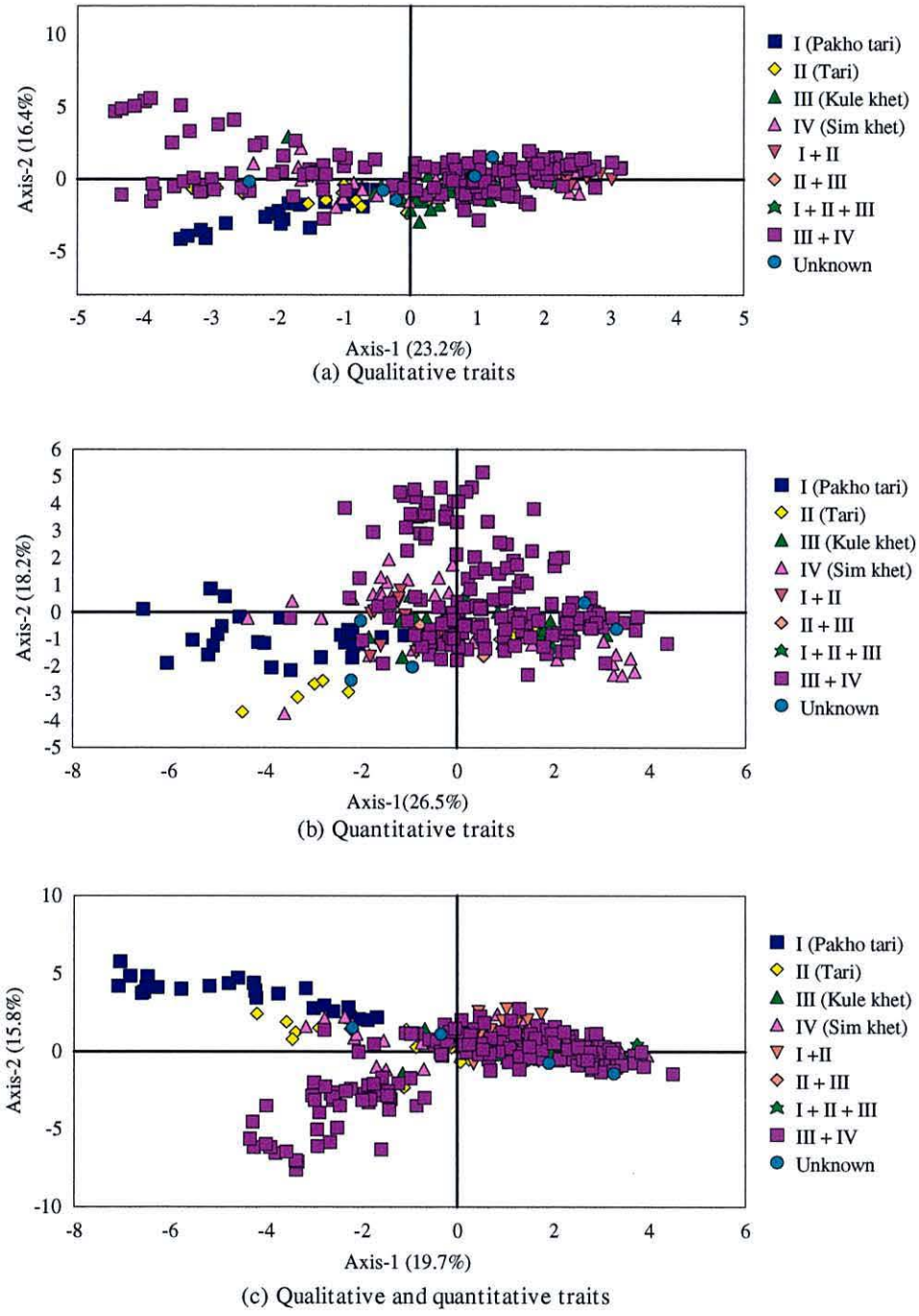


Figure 5.6: Scatter diagram of 292 rice accessions from the Kaski ecosite classified along the axes 1 and 2 of the ordination for (a): 17 qualitative traits; (b): 15 quantitative traits; and (c): 32 qualitative and quantitative traits combined. (Roman numerals denote rice domains.)

5.3.5.2 Cluster analysis

The cluster analyses showed similar results whether on qualitative or quantitative traits alone or a combined analysis of both of them. Only the results of analysis based on qualitative and quantitative traits were described here which grouped the 292 rice accessions into 6 distinct clusters (Figure 5.7; Tables 5.10 and 5.11). The mean values for these clusters and their relative distance from each other and from the centroid are given in Appendix 5.5.

The landraces specific to particular agro-ecological domains were grouped into different clusters (Table 5.10). The landraces adapted specific to marshy land (*Sim khet*) and those common to irrigated (*Kule khet*) and the marshy land (*Sim khet*) were evenly distributed over cluster I – V. However, cluster VI mostly had upland landraces (*Pakho tari*) and all the upland landraces were found in clusters V and VI. Hence, the results of the cluster analysis for agro-morphological group were very similar to those of the PCAs. The cluster IV was found to comprise all the Anadi groups growing common in *Kule khet* and *Sim khet*. These results were also similar to the results of the PCA.

Contrary to the PCA results, the rice accessions clustered by FUDs. Most of the landraces with the same name were grouped together in the same cluster with only a few exceptions (Table 5.11). Cluster I encompassed mostly the small and fine grain and aromatic rice landraces including the check variety Masuli. This cluster was most diverse in distribution of the landraces by domains (Table 5.10). The landraces under this cluster I were mostly from the Jhinuwa group, Bayarni group, Tunde, Jethobudho and Panhele. It is noteworthy that the check variety, Masuli, a fine grain variety was found to be included in this cluster I and showed consistent clustering pattern.

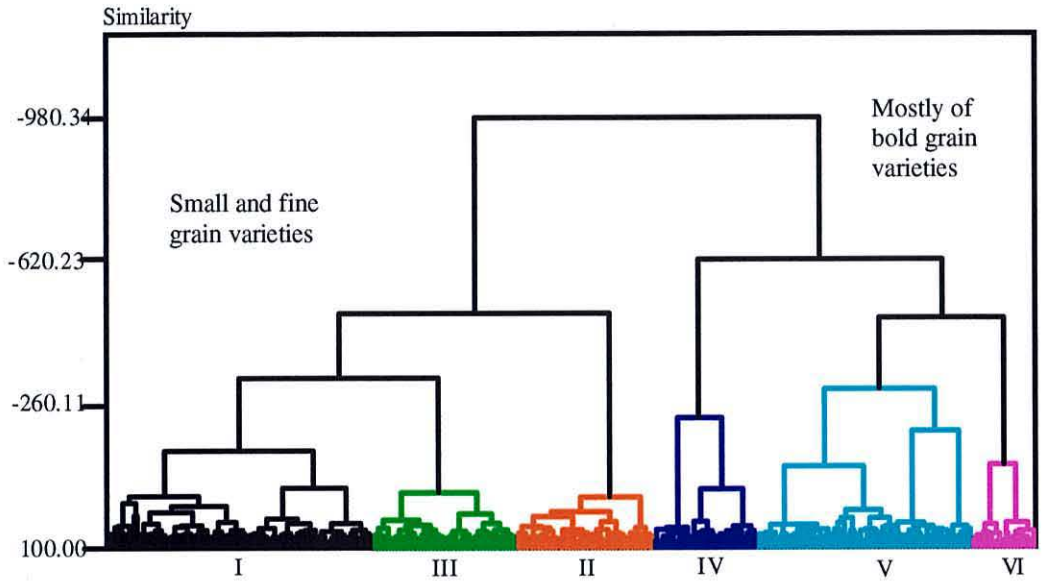


Figure 5.7: Dendrogram showing the clustering pattern of 292 rice accessions from the Kaski ecosite based on qualitative and quantitative traits. (Each colour represents a cluster.)

Table 5.10: Distribution of the Kaski rice accessions by domain in the 6 clusters formed by cluster analysis on 31 qualitative and quantitative traits.

Rice domains	I	II	III	IV	V	VI
1 *Pakho tari* (Upland)	-	-	-	-	5 (7)	19 (95)
2 *Tari* (Rainfed upland)	9 (11)	-	1 (2)	-	1 (2)	-
3 *Kule khet* (Irrigated)	23 (27)	-	1 (2)	-	-	1 (5)
4 *Sim khet* (Marshy land)	11 (13)	10 (23)	14 (31)	-	9 (13)	-
5 *Pakho tari* + *Tari*	2 (2)	13 (30)	-	-	-	-
6 *Tari* + *Kulekhet*	-	-	-	-	2 (3)	-
7 *Tari* + *Kule khet* + *Sim khet*	1 (1)	-	1 (2)	-	-	-
8 *Kule khet* +*Sim khet*	37 (44)	20 (47)	26 (58)	32 (100)	50 (73)	-
9 Not known	1 (1)	-	2 (4)	-	1 (2)	-
Total of clusters	**84**	**43**	**45**	**32**	**68**	**20**

Number in parentheses shows the percentage of accessions based on the total number of accessions of the cluster.

Cluster II was comprised mostly of accessions from the Gurdi group (domain category 8) as well as Naltumme specially adapted to marshy land (domain category 4) (Table 5.11). This cluster mostly included landraces with small and coarse grains. Cluster III was rather mixed and comprised the landraces not specific to particular FUD and not to the agro-ecosystems. All the Anadi group of rice accessions grouped into cluster IV and were characterized by having bold grains with long, coloured sterile lemmas. Likewise cluster V and VI were comprised of bold-grained accessions of Jarneli, Marshi, Madhese, Mansara and Aanga varieties (Table 5.11).

Table 5.11: Distribution in the cluster analysis of the Kaski rice landraces in 6 clusters by FUDs (local names).

Landrace	Domain group	Total accessions	I	II	III	IV	V	VI
Tunde	2	8	100	-	-	-	-	-
Pakhe Tunde	2	1	100	-	-	-	-	-
Panhele	3	10	100	-	-	-	-	-
Jetho Budho	3	15	86	-	7	-	-	7
Tarkaya Jhinuwa	4	1	100	-	-	-	-	-
Juya Jhinuwa	4	1	100	-	-	-	-	-
Madhise Jhinuwa	4	1	100	-	-	-	-	-
Chobo	4	6	63	-	-	-	37	-
Palungtare	4	2	50	-	-	-	50	-
Jyamdi Khole	4	2	50	-	-	-	50	-
Barmeli	4	1	100	-	-	-	-	-
Jadan	4	1	100	-	-	-	-	-
Pakhe Jhinuwa	5	2	100	-	-	-	-	-
Jhinuwa	7	2	50	-	50	-	-	-
Panhelo Jhinuwa	8	3	100	-	-	-	-	-
Tunde Jhinuwa	8	2	100	-	-	-	-	-
Lamcho Jhinuwa	8	1	100	-	-	-	-	-
Seto Jhinuwa	8	1	100	-	-	-	-	-
Masino Jhinuwa	8	1	100	-	-	-	-	-
Andheri Jhinuwa	8	1	100	-	-	-	-	-
Bayarni	8	12	100	-	-	-	-	-
Kalo Bayarni	8	1	100	-	-	-	-	-
Seto Bayarni	8	2	100	-	-	-	-	-
Gajale Bayarni	8	1	100	-	-	-	-	-
Juya Bayarni	8	1	100	-	-	-	-	-
Krishnabhog	8	2	100	-	-	-	-	-
Karna Jira	8	1	100	-	-	-	-	-
Masuli (check)	8	1	100	-	-	-	-	-
Kalo Jhinuwa	8	8	88	12	-	-	-	-
Bhayare	9	2	50	-	-	-	50	-
Masino Lahare Gurdi	4	2	-	100	-	-	-	-
Naltumme	4	8	-	100	-	-	-	-
Biramphool	4	8	-	12	88	-	-	-
Kathe Gurdi	5	13	-	100	-	-	-	-
Lahare Gurdi	8	9	-	100	-	-	-	-
Seto Gurdi	8	3	-	100	-	-	-	-
Sano Gurdi	8	2	-	100	-	-	-	-
Gurdi	8	1	-	100	-	-	-	-
Thulo KaloGurdi	8	1	-	100	-	-	-	-
Thulo Gurdi	8	4	-	50	50	-	-	-
Chhote	2	1	-	-	100	-	-	-
Basmati	4	7	-	-	100	-	-	-
Ramani	8	7	-	-	86	-	14	-
Aanpjhutte	8	2	-	-	50	-	50	-
Gauriya	8	1	-	-	100	-	-	-
Ekle	8	15	-	-	100	-	-	-
Dhabe Gaurama	8	1	-	-	100	-	-	-
Masino	9	1	-	-	100	-	-	-

Tulasi	9	1	-	-	100	-	-	-
Seto Anadi	8	15	-	-	-	100	-	-
Rato Anadi	8	15	-	-	-	100	-	-
Sano Anadi	8	1	-	-	-	100	-	-
Dudhe Anadi	8	1	-	-	-	100	-	-
Mansara	1	14	-	-	-	-	36	64
Pakhe Jarneli	2	8	-	-	-	-	100	-
Manamuri	2	2	-	-	-	-	100	-
Bhatte	2	1	-	-	-	-	100	-
Marshi group	4	4	-	-	-	-	100	-
Makai Khole	4	1	-	-	-	-	100	-
Rate	6	2	-	-	-	-	100	-
Madhise	8	12	-	-	-	-	100	-
Thulo Madhise	8	3	-	-	-	-	100	-
Naulo Madhise	8	2	-	-	-	-	100	-
Sano Madhise	8	1	-	-	-	-	100	-
Jarneli	8	13	-	-	-	-	100	-
Pakhe Ramani	8	1	-	-	-	-	100	-
Sano Aanpjhutte	8	1	-	-	-	-	100	-
Pakhe Gauriya	8	1	-	-	-	-	100	-
Thapachini	8	1	-	-	-	-	100	-
Bale	8	1	-	-	-	-	100	-
Battisara	8	1	-	-	-	-	100	-
Pani Barmeli	8	1	-	-	-	-	100	-
Aanga	1	10	-	-	-	-	-	100
Total of clusters		**292**	**84**	**43**	**45**	**32**	**68**	**20**

† = domain group referring to the table 5.9.

5.3.5.3 Discriminant function analysis

Discriminant function analyses were carried out using the FUDs, agro-morphological clusters derived in cluster analysis based on qualitative and quantitative traits and the rice domains as the group variables. The discrimination based on the qualitative and quantitative traits together was found the most informative for all three group variables used in the analysis. Over 94 % of the rice accessions were correctly classified to their FUDs, over 94 % to agro-morphological clusters and over 77 % to agro-ecosystems (Table 5.12) based on the combined qualitative and quantitative traits. The cross validation reduced their proportion only by a little (Table 5.12).

Table 5.12: Summary of discriminant function analysis of groups of rice accessions from Kaski based on 31 qualitative and quantitative traits.

Group	No of accessions in the group	Prop. correctly discriminanted	Prop. after cross validation
By landrace group:			
Aanga	10	100	100
Anadi group (4 FUDs)	32	100	100
Basmati	7	100	100
Bayarni group (4 FUDs)	17	82	82
Biramphool	8	88	88
Chobo	6	100	100
Ekle	15	93	80
Gurdi group (8 FUDs)	35	94	94
Jarneli group (2 FUDs)	21	100	100
Jethobudho	15	93	93
Jhinuwa group (12 FUDs)	24	79	54
Madhese group (4 FUDs)	18	100	100
Mansara	14	100	93
Marshi (4 FUDs)	4	100	75
Naltumme	8	100	100
Panhele	10	100	100
Ramani group (2 FUDs)	8	88	38
Tunde group (2 FUDs)	9	89	78
Total	**261†**	**94**	**89**
By agro-morphological cluster group:			
Cluster I	84	91	83
C luster II	43	95	93
Cluster III	45	96	93
Cluster IV	32	100	100
Cluster V	68	93	88
Cluster VI	20	100	95
Total	**292**	**94**	**90**
By agroecosystem domains:			
1 Pakho tari	24	79	79
2 Tari	11	67	52
3 Kule khet	25	88	84
4 Sim khet	44	76	53
5 Pakho tari + Tari	15	80	80
6 Tari + Kule khet	2	100	50
7 Tari + Kule khet + Sim khet	2	100	0
8 Kule khet + Sim khet	165	77	67
9 Not known	4	50	0
Total	**292**	**77**	**66**

† = groups represented by only one accession were excluded.

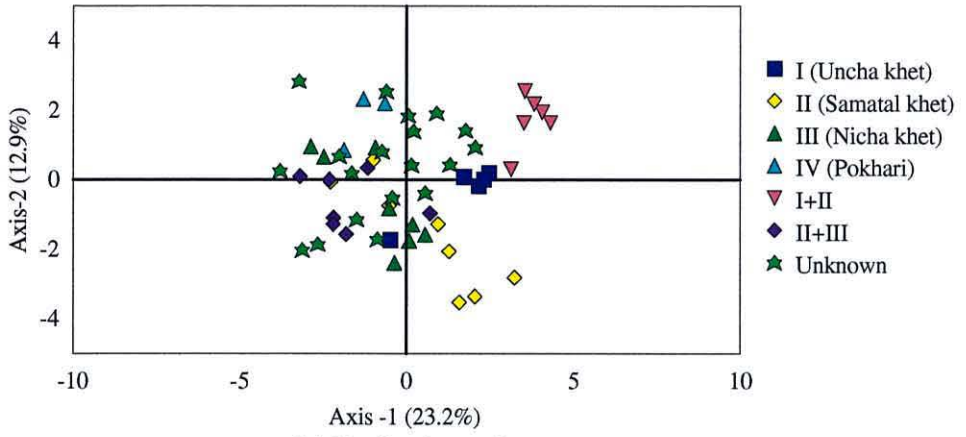
5.3.6 Multivariate analysis for Bara – Terai (lowland) ecosite

5.3.6.1 Principal component analysis

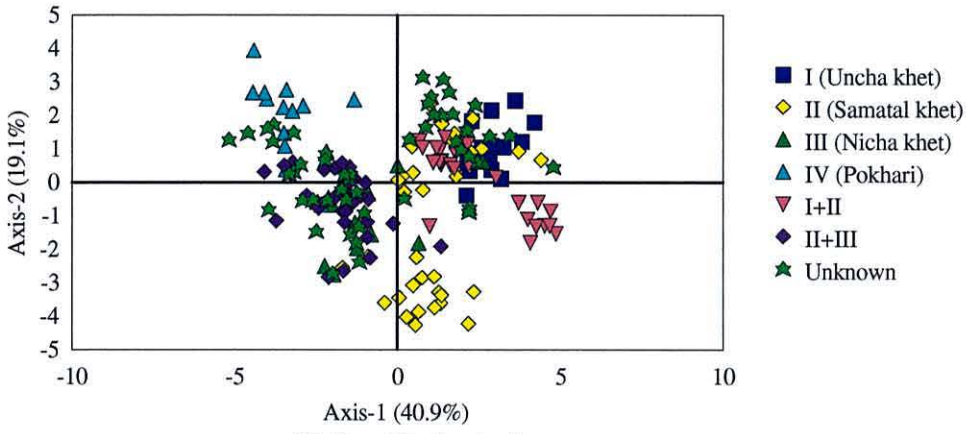
Three principal component analyses were done on the rice accessions from Bara (Figure 5.8). As in the Kaski landraces, the quantitative traits were found to be the most informative and explained 60% of the total variation (Figures 5.8a and 5.8b). Leaf length, ligule length, culm length, phenological traits and yield per plot contributed most to the first principal component and grain characters to the second (Appendix 5.6). The PCA of qualitative traits alone and the qualitative and quantitative traits combined were least informative and explained only about 36% of the total variation. PCA of qualitative traits alone was the least. However, node colour, panicle type, lemma and palea pubescence and grain characters were observed important qualitative traits over these PCAs (Appendix 5.6). The sign and the magnitude of the traits in the PCA over the three sets were also found consistent.

Most the landraces for rainfed conditions (*Uncha* and *Samatal khet*) were discriminated from the rest of the accessions by the first principal component. Furthermore, the accessions adapted to extreme stress conditions, like *Pokhari* and *Uncha khet*, formed two distinct and separate groups. The distribution of the accessions varied little between the PCAs whether the qualitative and quantitative traits were combined or analysed separately.

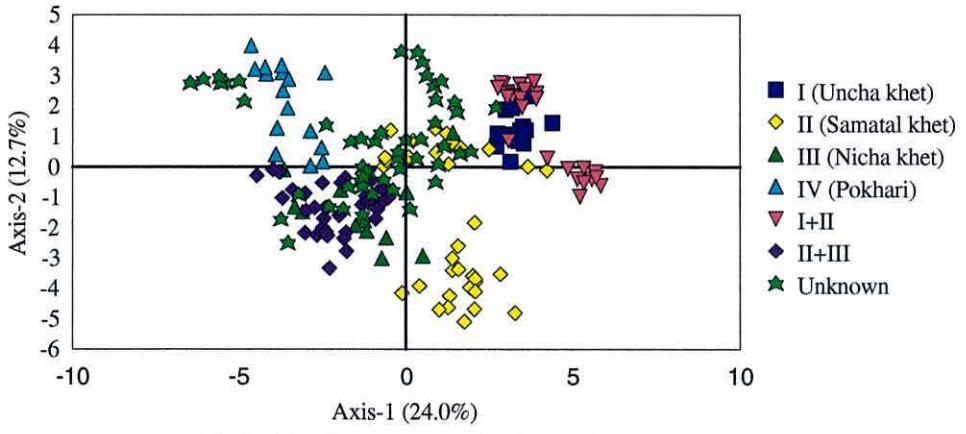
In Bara, the landraces were found distinct to some extent by agro-ecosystem when the accessions with unknown domains were removed from the analysis (Figure 5.8b), but the discrimination by FUDs was not possible.



(a) Qualitative traits



(b) Quantitative traits



(c) Qualitative and quantitative traits

Figure 5.8a: Scatter diagram of 196 rice accessions from the Bara ecosite classified by domains along the axes 1 and 2 of the ordination for (a): 19 qualitative traits; (b): 14 quantitative traits; and (c): 33 qualitative and quantitative traits combined. (Roman numerals denote rice domains.)

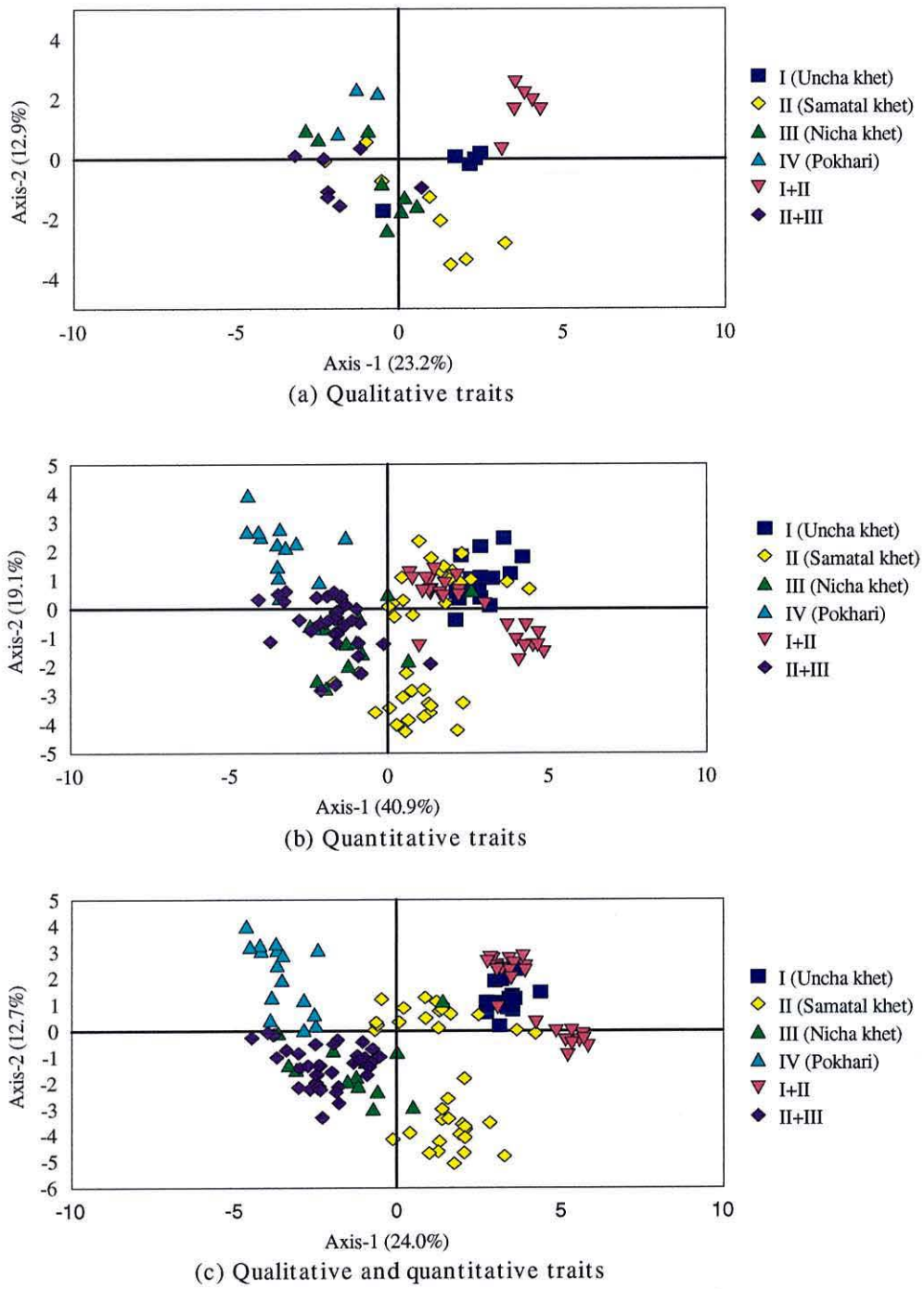


Figure 5.8b: Scatter diagram of 146 rice accessions (excluded 50 accessions without domain information) from the Bara ecosite classified by domains along the axes 1 and 2 of the ordination for (a): 19 qualitative traits; (b): 4 quantitative traits; and (c): 33 qualitative and quantitative traits combined. (Roman numerals denote rice domains.)

5.3.6.2 Cluster analysis

Among cluster analyses, based on all 33 qualitative and quantitative traits together and separately, the combined analysis was most informative and produced 7 clusters. The number of accessions per cluster ranged from 10 in cluster II to 67 in cluster III. Clusters II, V, VI and VII contained all the early maturing varieties and clusters I, III and IV contained the normal to late varieties of rice. There was clear separation of the clusters of two maturity groups, early and normal to late varieties depicted in dendrogram (Figure 5.9). The morphological characteristics of the clusters and the relative distance from each other and from the centroid were presented in Appendix 5.7. The landraces under water stress environment (*Uncha khet*) were all in cluster V (Table 5.13). However, the other varieties having a wide adaptation to water regime environments (*Samatal, Nicha* and *Pokhari*) were distributed across 2-3 clusters.

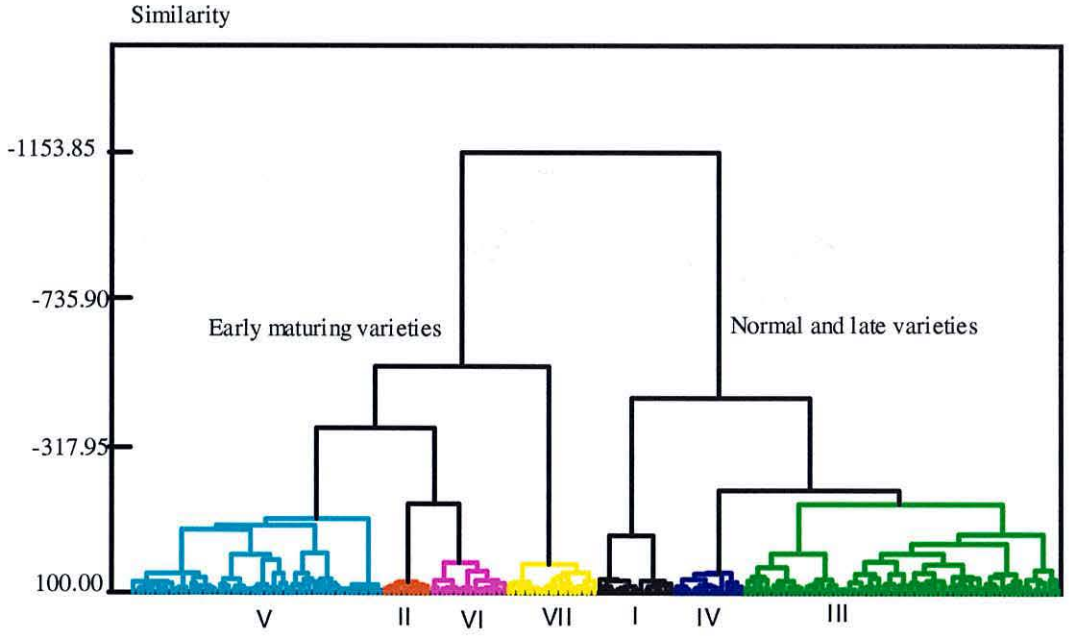


Figure 5.9: Dendrogram showing the clustering pattern of 196 rice accessions from the Bara ecosite based on qualitative and quantitative traits.

Table 5.13: Distribution of the Bara rice accessions by domains in the 7 clusters formed by cluster analysis on 33 qualitative and quantitative traits.

Rice domains	I	II	III	IV	V	VI	VII
1. *Uncha khet* (Upland)	-	-	-	-	23 (43)	-	-
2. *Samatal khet* (Rainfed)	-	-	18 (27)	-	16 (30)	-	19 (100)
3. *Nicha khet* (Irrigated)	7 (44)	-	6 (9)	-	3 (6)	-	-
4. *Pokhari* (Accumulated Water)	9 (56)	-	7 (10)	-	-	-	-
5. Uncha + Samatal	-	10 (100)	-	-	1 (2)	16 (100)	-
6. Samatal + Nicha khet	-	-	23 (34)	14 (93)	-	-	-
7. Not known	-	-	13 (19)	1 (7)	10 (19)	-	-
Total of clusters	**16**	**10**	**67**	**15**	**53**	**16**	**19**

Number in parentheses shows the percentage of accessions based on the total accessions of the group

Accessions with same name were mostly homogeneous and occurred together in the same cluster (Table 5.14). Bhadaiya Basmati, Basmati, and Dipahi were some of the exceptions but they were still distributed in the adjacent clusters (Table 5.14). Clusters I, II, IV and VII were distinct clusters having landraces with specific traits for adaptation, morphology, and quality. For example, cluster I contained landraces grown in *Pokhari* and *Nicha khet* and had all of the 9 landraces named Silhat and all of the 7 named Laltenger. These landraces were late maturing, tall, and thick culmed and had dark green, long and broad leaves, well exserted and pigmented panicles, long awns and erect tillers. All the 10 accessions of Sathi landrace formed the cluster II, these were early maturing and mostly they had the enclosed black panicles, and short and narrow leaves. These landraces have particular religious and cultural values. Cluster IV contained most (88%)

of the Basmati varieties, which have aroma and good grain quality and they are long grained. The 9 Khera and 10 Aanga accessions grouped together and formed cluster VII. These landraces were medium in plant height, had short and narrow leaves, awned panicles and were highly prone to shattering. These landraces were grown in *Uncha khet* and *Samatal khet*. The clusters III (normal rice group) and V (early rice group) contained the largest numbers of rice accessions with various names and also included the landraces adapted to more than one rice domain (Table 5.14).

Table 5.14: Distribution in the cluster analysis of the Bara rice landraces by FUDs (local names).

Landrace	Domain group	Total accessions	I	II	III	IV	V	VI	VII
Laltenger	3	7	100	-	-	-	-	-	-
Silhat	4	9	100	-	-	-	-	-	-
Sathi	5	10	-	100	-	-	-	-	-
B. Basmati	2	4	-	-	25	-	75	-	-
Dipahi	2	2	-	-	50	-	50	-	-
Mansari	2	3	-	-	100	-	-	-	-
U pharam	2	2	-	-	100	-	-	-	-
Dudhraj	2	4	-	-	100	-	-	-	-
Amaghauch	2	2	-	-	100	-	-	-	-
Mansara	3	3	-	-	100	-	-	-	-
Lajhi	3	4	-	-	100	-	-	-	-
Chhataraj	3	2	-	-	100	-	-	-	-
Bhatti	4	7	-	-	100	-	-	-	-
Karma	6	10	-	-	100	-	-	-	-
L pharam	6	8	-	-	100	-	-	-	-
K kamodh	6	1	-	-	100	-	-	-	-
Madhumala	6	2	-	-	100	-	-	-	-
Basmati	6	16	-	-	12	88	-	-	-
Harinkher	7	1	-	-	100	-	-	-	-
Rajala	7	5	-	-	100	-	-	-	-
Pakhad	7	4	-	-	100	-	-	-	-
Balamsar	7	1	-	-	100	-	-	-	-
Budhidayan	7	2	-	-	100	-	-	-	-
Brahmabhusi	7	1	-	-	-	100	-	-	-
Mutmur	1	9	-	-	-	-	100	-	-
Muturi	1	6	-	-	-	-	100	-	-
Rango	1	5	-	-	-	-	100	-	-
Gajargaur	1	2	-	-	-	-	100	-	-
Ghaiya	1	1	-	-	-	-	100	-	-
Nakhisaro	2	9	-	-	-	-	100	-	-
Dudhisaro	2	3	-	-	-	-	100	-	-
Kataush	3	3	-	-	-	-	100	-	-
Sabitri	5	1	-	-	-	-	100	-	-
Sikichan	7	3	-	-	-	-	100	-	-
Hattijhulan	7	5	-	-	-	-	100	-	-
Adalat	7	2	-	-	-	-	100	-	-
Sotawa	5	10	-	-	-	-	-	100	-
Sokan	5	6	-	-	-	-	-	100	-
Khera	2	10	-	-	-	-	-	-	100
Aanga	2	9	-	-	-	-	-	-	100
Total of cluster		**196**	**16**	**10**	**67**	**15**	**53**	**16**	**19**

† = domain group referring to table above

5.3.6.3 Discriminant function analysis.

The clustering results were confirmed by discriminant function analysis using the same set of qualitative and quantitative traits separately and combined together. The FUDs, morphological clusters and the agro-ecological domains were used as group variables. Over 98% of the accessions were correctly classified by name, nearly 100% by morphological clusters and nearly 96% by rice domains. In the cross validation, any reduction in the correct classifications were small (Table 5.15), although, the cross validation of the classification by landrace name was not possible as there were too many groups.

Table 5.15: Summary of discriminant function analysis of groups of rice accessions from Bara based on 33 qualitative and quantitative traits.

Groups	No of accessions per group	Prop. correctly discriminated	Prop. after cross validation
By landrace name:			
Aanga	9	100	NA†
Basmati	16	100	NA
Bhadaiya basmati	4	100	NA
Bhatti	7	100	NA
Dudhraj	4	100	NA
Hattijhullan	5	100	NA
Karma	10	100	NA
Khera group (3 FUDs)	10	100	NA
Lajhi	4	100	NA
Laltenger	7	100	NA
Mansara	5	100	NA
Mutmur	9	100	NA
Muturi	6	100	NA
Nakhisaro	9	100	NA
Pakhad	4	100	NA
Pharam group (2 FUDs)	10	100	NA
Rajala	5	100	NA
Rango	5	100	NA
Sathi	10	100	NA
Silhat	9	100	NA
Sokan	6	100	NA
Sotawa	10	90	NA
Total	**163†**	**98**	
By agro-morphological groups:			
Cluster I	16	100	100
Cluster II	10	100	100
Custer III	67	99	94
Cluster IV	15	100	100
Cluster V	53	100	94
Cluster VI	16	100	100
Cluster VII	19	100	100
Total	**196**	**100**	**98**
By domains:			
1 Uncha khet	23	100	96
2 Samatal khet	53	88	75
3 Nicha khet	16	91	82
4 Pokhari	16	100	100
5 Uncha + samatal khet	27	100	100
6 Samatal + nicha khet	37	100	97
7 Not known	24	100	92
Total	**196**	**96**	**92**

† = The landraces represented by one accession were excluded.

5.4 Discussion

5.4.1 Genetic diversity revealed by different analysis procedures

Classical methods of estimating diversity among the groups of plants have chiefly relied on measuring variation in the morphological traits. Varying levels of phenotypic diversity were observed in the landraces from the three ecosites. The three multivariate procedures: hierarchical clustering, principal component and discriminant function analysis helped to explain the variability among the landrace accessions. In all of these analyses, the FUDs and agroecological domains were used as the group variables to establish relationships and diversity among these landraces within each ecosite. For Kaski and Bara the landrace populations were composed of different genotypes that were grouped into clusters that could be largely explained by FUDs and domains.

These multivariate analyses in diversity estimation have been used and shown to have utility and appropriateness in revealing the germplasm classification and groupings (Furman *et al.*, 1997, and Newbury and Ford-Lloyd, 1997). The principal component analyses used here have shown to some extent the groupings of accessions by agro-ecological domains in Bara and Kaski ecosites. Among the PCAs of agro-morphological data, the analysis of the quantitative data alone was found most informative to explain the highest variation for both of these ecosites and qualitative traits alone for the high-hill ecosite, Jumla. However, the scatter plots showing the distribution of rice landraces based on respective PCAs was not clear enough to differentiate the landraces by FUDs. They were rather dispersed with few exceptional groupings of landraces such as those named Anadi in Kaski and Silhat, Laltenger and Sathi in Bara. The PCA of qualitative and quantitative traits combined together and qualitative traits alone for the high-hill ecosite

(Figure 5.4a and b) revealed a very close relatedness among the accessions under study and the three analyses showed similar groupings to some extent.

Hierarchical clustering and discriminant procedures used in this study were complementary since the information provided by these analyses were to a large extent similar. A number of measures have been proposed and used for the distance or similarity between two individuals or two populations for cluster analysis based on the type of information available. Euclidean distance and taxonomical distance are most frequently used for quantitative variables (Manly 1994). Simple matching coefficient has been commonly used for the variables with multi states and binary presences and absences of traits so that the effects of coding (the particular order in which the states of a trait are coded) have less effect. Ward's linkage and the taxonomical distance measures were found most informative for clustering the landraces in the present study. Both the analyses: clustering and discriminant function grouped the landraces better by FUDs than the PCAs. Furthermore, the clusterings using the qualitative and quantitative data together and qualitative data alone showed a similar distribution and clustering pattern of the accessions across the ecosites. However, the analysis on quantitative traits alone was found most informative in Kaski and Bara in PCAs. Likewise, a clear discrimination of the rice accessions to respective FUDs, clusters and agro-ecological domains was revealed by the discriminant function analysis using the combined qualitative and quantitative traits and better confirmed the clustering pattern than the ordination (PCAs) of the rice accessions within each ecosite. Howeve, the order of coding of the qualitative traits with multiple states could have biased the results to some extent by producing false correlations among traits or failing to detect correlations. Simple matching coefficient as

a distance measure for clustering would be less affected by coding effects, but these analyses have not been done.

5.4.2 Genetic diversity, agro-ecological domains and FUDs

Groupings of the landrace accessions into different clusters and the resulting distances among the cluster centroids (Appendices 5.3, 5.5 and 5.7) suggest that there is high morphological diversity among the landraces grown in different agro-ecological domains within the ecosite. Morphological differentiation was greatest under favourable environment and more limited in stressed environment. Favourable domains (water not limited) comprised a group of highly diverse landraces, which were differentiated into several different clusters in both the mid-hill and lowland ecosites. However, the landraces specific to the water limited environments: *Pakho tari* in Kaski and *Uncha khet* in Bara, had little genetic variation and were hence grouped together into a single cluster. These were Khera and Aanga for the Bara ecosite and Mansara and Aanga for the Kaski ecosite (Table 5.10, 5.11, 5.13, and 5.14). They were best adapted to this specific domain in these regions and were distinct morphological forms with bold and awned grains and were highly shattering. It showed that the morphological variability and the agro-ecological domains are associated with each other.

The traditional farmers' naming and classification of rice landraces (FUDs) indicated the association of names with morphological forms. Most of the accessions with same name grouped together into the same distinct clusters (Tables 5.10 and 5.13; Figures 5.8 and 5.10). For example, in Bara the clusters I (100 % of Silhat and Laltenger), II (100 % of Sathi), IV (88 % of Basmati), VI (100 % of Sotawa and Sokan)

and VII (100 % of Khera and Aanga), and in Kaski, clusters II (94 % of Gurdi groups), IV (100 % of Anadi group) and VI (100 % of Mansara and Aanga) formed distinct morphological clusters having specific morphological traits. It is clear that the ethno-botanic classification of landraces used by the farmers is useful information that helps to describe the variability that is found. In contrast farmers' names for landraces were less important in the out breeding crop pearl millet (*Pennisetum glaucum*) under farmer management in West Africa (Busso *et al.*, 2000).

However, FUDs did not account for all the variation that was found. In Kaski, the Jhinuwa group, Bayarni group, Tunde group, Marshi group, Ramani, Ekle, Mansara, and Jethobudho exhibited within FUD variation with 7 to 62% of the FUDs misclassified. Similarly, 10% of the FUDs Sotawa in Bara were misclassified and 12 to 50% of FUDs named Basmati and Dipahi grouped into two adjacent clusters. It is clear that most of the misclassification was observed in the landrace varieties (Tables 5.10 and 5.11) grown in small area and by few farmers (Tables 3.1b and 3.1c in Chapter III). For example, the high rate of misclassification in Jhinuwa (46% in discriminant and 50% in cluster analysis) suggested that the Jhinuwa was either most diverse for morphological traits or several of the FUDs were misnamed.

5.4.3 Genetic diversity and the distinctive traits

The morphological traits have been used traditionally to distinguish rice varieties. A number of studies have been carried out to assess the germplasm diversity and genetic structure of different crop species using the geographical origin and the agro-morphological data and have pointed out the relative importance of qualitative and

quantitative traits in discriminating the genetic variation. Murphy and Witcombe (1981) on the phenotypic diversity study in a collection of barley accessions from Nepal and Pakistan, pointed out the importance of qualitative and quantitative traits in distribution of genetic variation. They found that high diversity in qualitative traits were related to centres of diversity and diversity in quantitative traits related to diversity in the geographic distribution of the plant germplasm. A similar observation was made by Weltzien (1989) that the quantitatively inherited traits were effective for geographic classification of barley landrace populations from Syria and Jordan.

In this regard, an interesting difference in the perception of local farmers towards the qualitative (multistate) and quantitative characteristics of rice and its classification was observed in the present study. In Jumla, the qualitative traits explained the most variation and were found informative (Figure 5.4), whereas in Kaski and Bara, it was of less importance and the quantitative traits explained the most variation (Figure 5.6 and 5.9). Table 16 shows the important qualitative and quantitative traits by ecosites. In Jumla, there were fewer landraces, and limited variation within the populations. So farmers have selected the qualitative traits such as colour traits and have used them in the ethno-botanic classification of their landraces. For example, Rato Marshi, Kalo Marshi and Seto Marshi are the common landraces in Jumla. Rato Marshi has bold grains with red glumes and red pericarp, Kalo Marshi dark purplish to black glumes, and Seto Marshi white and clear glumes and pericarp. Stigma colour, sterile lemma colour and the awning traits were the important qualitative traits, which resulted in the groupings of the rice accessions into distinct clusters in the PCA and the hierarchical clustering analysis (Appendix 5.2 and Figure 5.4 and 5.5).

Table 5.16: A summary showing the traits important to explain the variation in PCAs in each ecosite.

Traits	Jumla	Kaski	Bara
Qualitative traits:			
Awning	+	-	-
Sigma colour (STIGCOL)	+	-	-
Sterile lemma colour (STLCOL)	+	-	-
Quantitative traits:			
Leaf length (L_LENGTH)	-	+	+
Leaf width (L_WIDTH)	-	+	-
Length width ratio (LBRATIO)	-	+	-
Ligule length (LIGLEN)	-	-	+
Culm length (C_LENGTH)	-	-	+
Effective tillers (TILLERS)	-	-	+
Panicle length (P_LENGTH)	-	+	
Days to heading (DH)	-	-	+
Days to maturity (DM)	-	-	+
1000 grains weight (TGW)	-	+	+
Well-filled grains per panicle (GRAINSP)	-	+	-
Grains weight per panicle (GWP)	-	-	+

Unlike in Jumla, the farmers in Kaski and Bara grew a wide diversity of landraces varying for adaptation, grain quality, straw quality, religious and quality values, and for home consumption. Growth duration (phenological traits) and grain morphological traits were found important traits to cluster the entire accessions of Bara and Kaski in the present study and depicted two big clusters at the root of the respective dendrograms (Figure 5.7 and 5.10). In the Kaski landraces, each cluster was characterized by grain quality and grain morphologies. For example, cluster I was mostly composed of the fine grained and aromatic rice groups; cluster II of the small and coarse grain type while cluster IV had mostly bold, flat grain types with long sterile lemmas (Anadi group) (Table 5.11). Similarly, in Bara, the clusters I, III and IV composed of normal to late maturing accessions; and cluster II, V, VI, and VII the early maturing groups but they differed in plant types, panicle and grain characteristics.

5.4.4 Genetic diversity and the ecosites

In the high-hill ecosite, the diversity was rather limited. There were fewer landraces in this site and they were nearly homogeneous for most morphological traits. There was a low average diversity for qualitative traits and most of the accessions grouped close together into a single cluster (Figure 5.4). In Jumla, the limiting factor is the natural selection. The harsh environment with extreme chilling temperatures for rice restricted the diversity. In the mid-hill region and the lowland plain area, with favourable growing environments, there was high landrace diversity manifested by the wide range of agro-morphological variability among the respective landrace accessions under the study. The existence of heterogeneous environments in the mid-hills and lowlands has contributed to the level of diversity. Both natural selection and conscious selection by farmers to adapt in the diverse growing environments have increased and maintained the landrace diversity. In a morphological diversity study in Sorghum germplasm in Ethiopia, Ayana and Bekele (1999) and Ezeaku *et al.* (1999) found similar observations that selection for the adaptive traits to varied growing environments were the responsible factors in shaping the structure and level of genetic diversity.

5.5 Conclusions

- The multivariate analysis procedures based on agro-morphological traits showed that there were significant differences between the named landraces and they were of different morphological forms within each ecosite. The hierarchical clustering and the discriminant function analysis were found most informative in grouping these landraces by FUDs, clusters and agro-ecological domains.

- The structure and level of morphological variation in rice landraces is influenced by the eco-geographic variation and the growing environment of the region. In the high-hills, the qualitative (multi-state) traits were more important. In the landraces from the mid-hills and lowland (terai) region, quantitative traits were more important for the discrimination of rice landraces.
- Assigning a landrace 'name' in the traditional classification based on morphological traits, adaptation traits, use and quality traits reflected the genetic identity in the Kaski and Bara ecosites. The identically named accessions collected from different farmers were clustered together. However this was not true for the high-hill ecosite where there was low diversity and the landraces names did not exactly relate to agro-morphology.
- The landraces grown under marginal (stressed) environments: Marshi in Jumla with chilling temperature and landraces that were grown in *Pakho tari* in Kaski, and *Uncha khet*, *Pokhari* in Bara are genotypes with specific adaptive traits for these environments. These landraces require conservation as they can be used to breed varieties for those marginal environments.
- The rice diversity decreased from the high-hill to mid-hill and lowland. The harsh environment with chilling temperature restricted the diversity with low adaptation in the high hills. In mid-hill and lowland (*Terai*) regions, the diverse growing environments helped maintain the landrace diversity in farmers' fields.

CHAPTER VI

MOLECULAR MARKER DIVERSITY IN RICE LANDRACES

6.1 Introduction

Agro-morphological characteristics have long been used to classify plant genotypes and distinguish among them. The evaluation of Nepalese rice accessions based on these agro-morphological traits showed that there was a high amount of morphological differentiation among the landraces within each ecosite. Landraces with different names (FUDs) tended to be morphologically distinct and adapted to their growing environments. These landraces therefore constitute an important genetic resource for crop improvement. However, very little is known about the genetic structure (the allelic composition) and relationships of these Nepalese rice landraces.

The genetic diversity of a population depends upon the number and frequency of alleles (allelic composition). Molecular markers have been found useful technique to measure the extent of diversity within and between plant populations (Rongwen, *et al.*, 1995). Of the molecular markers, microsatellite markers are one of the most powerful new genetic markers. They are based on the number of simple sequence DNA repeats (SSRs) and have become the markers of choice for a wide spectrum of genetic and population studies due to their high polymorphism, codominance, abundance throughout the genome and relative ease of scoring (Gao *et al.*, 2002). Microsatellite (SSR) markers have extensively been used in molecular characterization, genetic diversity studies and management of genetic resources in seed genebanks. Hundreds of microsatellites from rice have been developed, and mapped (Wu and Tanksley, 1993; Paunad *et al.*, 1996; Chen *et al.*, 1997; McCouch *et al.*, 1997; Cho *et al.*, 2000; and Temnykh *et al.*, 2000).

The present study on molecular marker diversity in rice landraces was carried out in the newly established laboratory in Nepal with the objective to measure and assess the genetic relationships and variability in rice landraces from the three Nepalese ecosites by the use of DNA markers.

6.2 Materials and Methods

To support this study, the first functional molecular laboratory in Nepal was established in the Agriculture Botany Division, NARC, Nepal. IPGRI-APO (IPGRI regional office for Asia, Paciffic and Oceania), Malaysia; National Institute of Agrobiological Resources (NIAR), Japan; the DFID-UK Plant Science Research Programme; and NARC, Nepal supported this effort technically, and financially.

A sample of 21 rice accessions from 147 in the Jumla ecosite, 24 from 291 in the Kaski ecosite and 25 from 194 in the Bara ecosite was randomly sampled within strata of economic importance (average area under cultivation and the number of HHs growing a named landrace). In each ecosite, the sampled accessions represented the 10 differently named farmers' unit of diversity (FUDs) and represented diverse rice growing environments (Appendices 4.2a, 4.2b and 4.3 in Chapter IV). Details of the accessions under study are shown in Table 6.1. Three modern varieties: Kalinga III, IR64 and IR36 were included as reference cultivars.

Thirty nine rice microsatellite (SSR) primers synthesized by Research Genetics, USA, representing different regions of the rice genome were used for the amplification of rice landraces. These microsatellite (SSR) primers were selected based on the level of published polymorphic information content and their distribution in the rice genome

(Appendix 3.4 in Chapter III). The present study was based on the bulk DNA of 10 individual seedlings of each accession to provide a better representation than using a single plant (Virk *et al.*, 1995).

Molecular techniques, band scoring and statistical analysis were described in Chapter III (3.3 for SSR marker diversity analysis; 3.4.1.6 for measures of distances; and 3.4.2 for molecular marker analysis). All the polymorphic microsatellite loci observed in the studied landrace accessions from each of the ecosites were subjected to cluster analysis (except for Jumla where both polymorphic and monomorphic loci were used) both with and without the check varieties: Kalinga III, IR64 and IR36 using the NTSYS-pc program (Rohlf, 1993). Dendrograms were created based on Nei's genetic distance using the polymorphic microsatellite DNA data for each set of rice accessions both with and without check varieties using an Unweighted Pair Group Method with Arithmetical Averages (UPGMA). The resulting dendrograms from Bara and Kaski ecosites were further compared with results using Jaccard's similarity coefficients with UPGMA.

Table 6.1: Rice genotypes used in the study on DNA variation by microsatellite (SSR) markers

S.No.	Jumla ecosite		Kaski ecosite		Bara ecosite	
	Landrace	**Accession number**	**Landrace**	**Accession number**	**Landrace**	**Accession number**
1	Darime	1052 – 1	Aanga	219	Basmati	052
2	Darime	2006 – 1	Aanga	224	Basmati	044
3	Dhan	1001 – 1	Gurdi	058	Basmati	320
4	Dhan	1013 – 1	Jetho Budho	203	Dudhraj	215
5	Jumli R. Marshi	2036 – 1	Jetho Budho	205	Karma	036
6	Kalo Marshi	1023 – 1	Jetho Budho	207	Karma	033
7	Kalo Marshi	1057 – 1	Jetho Budho	214	Karma	029
8	Kalo Marshi	1084 – 1	Jhinuwa	012	Lajhi	101
9	Mehele	2051 – 1	Jhinuwa	013	Lajhi	104
10	Mehele	2068 – 1	Kalo Bayarni	070	Lajhi	097
11	Palte Dhan	2020 – 1	Kalo Bayarni	072	Laltenger	304
12	Rato Dhan	1020 – 1	Kathe Gurdi	025	Laltenger	125
13	Rato Marshi	1009 – 1	Kathe Gurdi	028	Mansara	146
14	Rato Marshi	1015 – 1	Kathe Gurdi	029	Mansara	150
15	Rato Marshi	2003 – 1	Kathe gurdi	033	Mansara	148
16	Rato Marshi	2035 - 1	Ramani	158	Mutmur	065
17	Seto Dhan	2033 – 1	Ramani	161	Mutmur	066
18	Seto Marshi	1022 – 1	Rato Anadi	092	Mutmur	301
19	Seto Marshi	1064 – 1	Rato Anadi	094	Nakhisaro	118
20	Seto Marshi	1066 – 1	Rato Anadi	104	Nakhisaro	117
21	Seto Marshi	1024 – 1	Seto Gurdi	052	Nakhisaro	308
22	Kalinga III	Check variety	Seto Gurdi	053	Sathi	012
23	IR36	Check variety	Tunde Jhinuwa	014	Sathi	018
24	IR64	Check variety	Tunde Jhinuwa	015	Sokan	183
25			Kalinga III	Check variety	Sokan	307
26			IR36	Check variety	Kalinga III	Check variety
27			IR64	Check variety	IR36	Check variety
28					IR64	Check variety

6.3 Results

In order to optimize and examine the reliability of the microsatellite protocol on rice landraces, preliminary PCR amplifications were carried out using template bulk DNA of 10 individuals of each of two round-grained medium tall progenies of a cross Kalinga III x IR64 (# MT2-3 and # 128-3) from a participatory plant breeding programme in Nepal (Witcombe *et al.*, 2000) and Kalinga III, IR64 and IR36. Twelve primer pairs were used and most of them (10 markers) yielded good PCR products. Based on these results, the appropriate protocol was determined for successive amplification of PCR products in assessing the variation in rice landrace genotypes. In each PCR amplification run, three check varieties were included to verify the reproducibility of bands. Only the markers with doubtful banding pattern and runs where there was no amplification were repeated. The primers and the accessions where many amplification products were missing from the gels (failure of the PCR reaction) were excluded from the analysis. Most of the SSR markers demonstrated single locus variation and only a few accessions had two alleles.

6.3.1 Jumla: the high-hill ecosite

6.3.1.1 Genetic diversity of microsatellite (SSR) markers

Out of 40 markers (primers) tested, 39 markers yielded the amplification products in all accessions under study, except 1052-1 named Darime, that was therefore excluded from the diversity analysis. The results of PCR amplification of SSR loci in 23 rice accessions for 39 rice microsatellite (SSR) primer pairs are summarized in Table 6.2. A total of 81 alleles (bands) at 39 loci were detected with an overall average of 2.1 alleles per locus in the whole set (20 FUDs and 3 MVs) of rice accessions under study and 40 alleles in 20

landrace accessions from Jumla. Thirty two of the 39 primer pairs (82 %) detected DNA polymorphism among the 23 rice accessions (PIC 0.08 to 0.23) (Table 6.2). However RM226 (3 %) was the only polymorphic molecular marker in the landrace accessions (1013-1) with two alleles in one accession only (PIC 0.05 for landraces alone) (Plate 6.1). Seven primer pairs (RM48, RM203, RM22, RM148, RM349, RM215, and RM120) produced monomorphic bands across all 23 rice accessions (landraces and checks) (PIC = 0). Polymorphisms between the modern check varieties and landrace accessions were observed at each of 32 loci out of the 39 primer pairs. Hence, landraces from Jumla, although they were called by different FUDs, were found to be genetically homogenous with a common allele for all 39 markers. Only one extra allele along with common allele at the locus (RM226) was detected occurring in accession 1013-1 (Table 6.2). This result with double band in accession 1013-1 was found repeatable and confirmed by repeated amplification and gel analyses for three times (Plate 6.1).

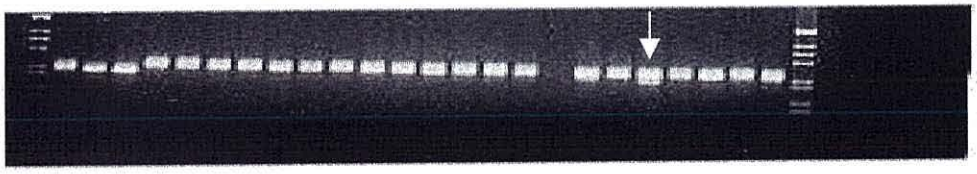


Plate 6.1: Microsatellite DNA profiles of rice landraces from Jumla detected by primer pair RM226 showing single banding pattern except in sample 20 (1013-1) with double bands. The amplified microsatellite DNAs were separated in 3 % agarose gel and detected by ethidium bromide.

(From left to right: Ladder, Kalinga III, IR64, IR36, 1022-1, 1064-1, 1066-1, 1024-1, 1009-1, 1015-1, 2003-1, 2035-1, 1023-1, 1057-1, 1084-1, 2051-1, 2068-1, 1052-1, 2006-1, 1001-1, 1013-1, 1020-1, 2033-1, 2020-1, 2036-1, and ladder).

Table 6.2: Number of alleles and polymorphic information content (PIC) of each of SSR locus in rice accessions from the Jumla ecosite.

S.No	Primer	Chromosome	Allele			PIC	
			Total sample	Landrace accessions	Check varieties	Total sample	Landrace accessions
1	RM246	1	3	1	3	0.16	0
2	RM5	1	3	1	2	0.23	0
3	RM122	2	2	1	1	0.23	0
4	RM211	2	2	1	2	0.08	0
5	RM213	2	2	1	2	0.08	0
6	RM48	2	1	1	1	0	0
7	RM203	2	1	1	1	0	0
8	RM60	3	2	1	1	0.23	0
9	RM232	3	2	1	2	0.08	0
10	RM22	3	1	1	1	0	0
11	RM148	3	1	1	1	0	0
12	RM226	4	3	2	1	0.26	0.05
13	RM201	4	2	1	1	0.23	0
14	RM349	4	1	1	1	0	0
15	RM164	5	3	1	2	0.23	0
16	RM26	5	2	1	1	0.23	0
17	RM3	6	2	1	1	0.23	0
18	RM234	7	3	1	2	0.23	0
19	RM248	7	3	1	2	0.23	0
20	RM11	7	2	1	2	0.16	0
21	RM351	7	2	1	1	0.23	0
22	RM350	8	2	1	1	0.23	0
23	RM223	8	2	1	1	0.23	0
24	RM215	9	1	1	1	0	0
25	RM257	9	3	1	2	0.23	0
26	RM242	9	3	1	2	0.23	0
27	RM228	10	2	1	1	0.23	0
28	RM222	10	2	1	1	0.23	0
29	RM244	10	2	1	2	0.16	0
30	RM167	11	2	1	1	0.23	0
31	RM206	11	2	1	2	0.08	0
32	RM229	11	2	1	1	0.23	0
33	RM224	11	3	1	2	0.23	0
34	RM17	11	2	1	2	0.16	0
35	RM21	11	3	1	2	0.23	0
36	RM120	11	1	1	1	0	0
37	RM202	11	2	1	2	0.16	0
38	RM247	12	2	1	1	0.23	0
39	RM20	12	2	1	1	0.23	0
Total number of alleles			81	40	57		
Total number of accessions			24	20	3		
Average alleles per locus (A)			2.1	1.0	1.5		
Number of polymorphic loci			32	1	17		
Number of polymorphic alleles			74	1	35		
Av. alleles per poly. Locus (Ap)			2.3	1.0	2.1		
PPL†			82.1	2.6	43.6		
PPA‡			91.3	2.5	61.5		
Average gene diversity (PIC)						0.17	0

PPL† = percentage of polymorphic loci; PPA‡ = percentage of polymorphic alleles

6.3.1.2 Clustering and principal component analysis of SSR variations

The observed 81 alleles were subjected to multivariate analyses.The dendrogram structures and the order of the accessions was not affected by the use of Nei's genetic distance and Jaccard's pairwise similarity coefficient. All the landrace accessions clustered together and formed a distinct group from the check modern varieties. The landraces from Jumla, all but with different traditional names, were genetically homogenous with the sole exception of 1013-1. A similar observation was encountered in the principal component analysis that differentiated the landrace accessions and the check varieties accounting for 93 % of the total variation (Figure 6.3).

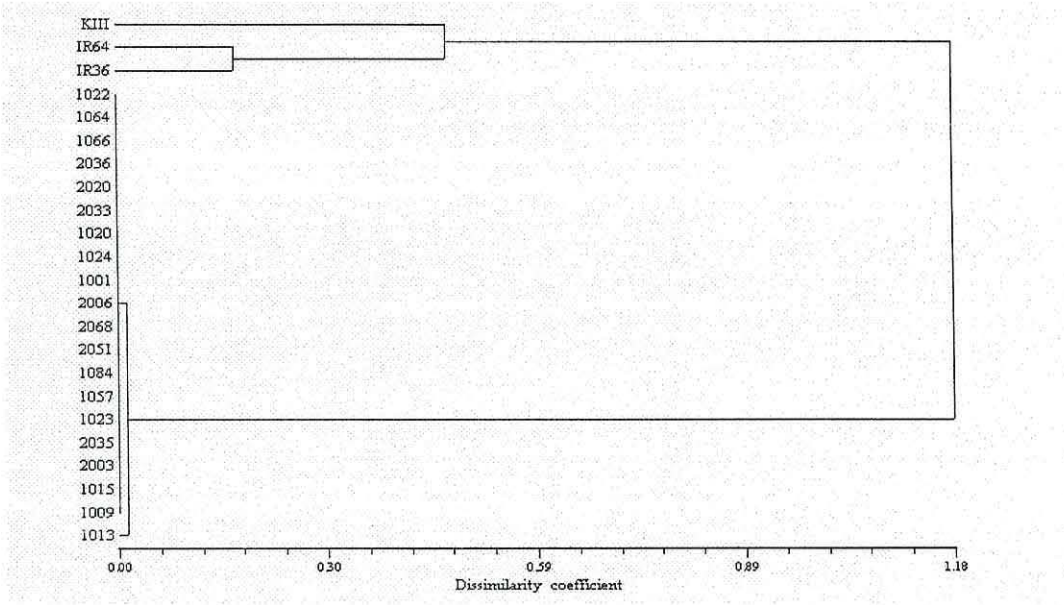


Figure 6.1: Dendrogram of 20 landrace accessions and 3 check varieties using Neis genetic distance and UPGMA clustering based on all 81 alleles generated by 39 SSR markers in Jumla landraces.

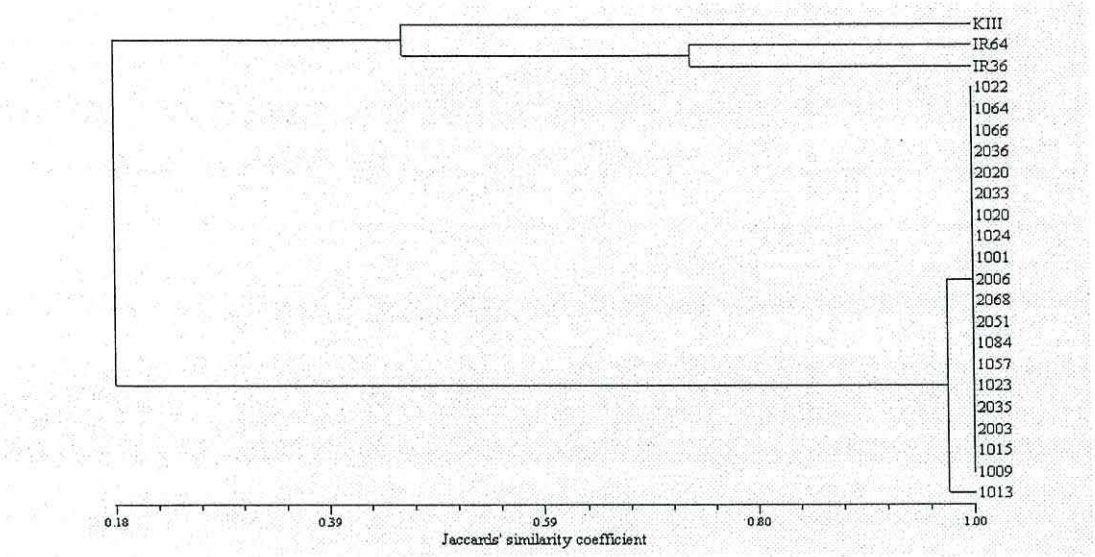


Figure 6.2: Dendrogram of 20 landrace accessions and 3 check varieties using Jaccards' similarity coefficient and UPGMA clustering based on all 81 alleles generated by 39 SSR markers in Jumla landraces.

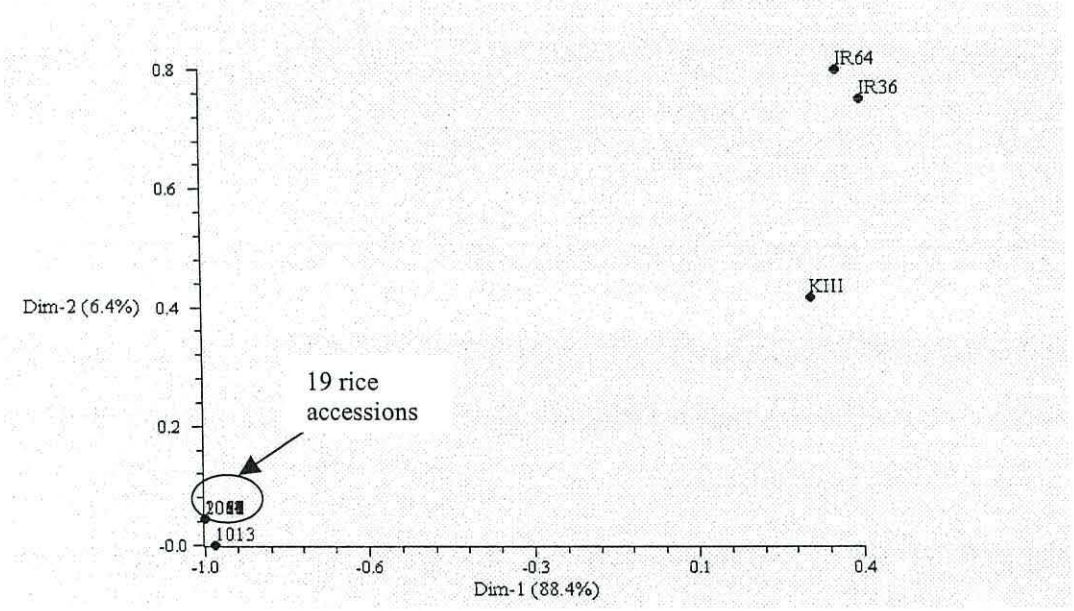


Figure 6.3: Scatter plot of rice accessions including 20 landraces from Jumla and 3 check varieties based on principal component analysis of 81 alleles generated by 39 microsatellite markers.

6.3.2 Kaski: the mid-hill ecosite

6.3.2.1 Genetic diversity of microsatellite (SSR) markers

SSR RM201, which gave no amplified product in most landrace accessions, was excluded from the diversity analysis. With the remaining 38 primer pairs 105 alleles were scored in the total set of 27 rice landraces and checks and 91 alleles in the 24 landraces (Table 6.3). Out of 38 primer pairs, 95% detected polymorphisms in the 27 landraces and checks and 90% in 24 landraces. The number of alleles observed at each locus ranged from 1 (RM60 and RM120) to 6 (RM247) with an average of 2.7 alleles per locus for total set, 2.4 for landrace set and 1.5 for check modern varieties. The average number of alleles per polymorphic locus and percentage of polymorphic alleles (PPA) varied among rice accessions under study (Table 6.3). The maximum number of allelic variants in the landraces was 5 alleles with RM247. Four alleles were resolved with markers RM246, RM164 and RM20 while the rest of the markers either resolved 2-3 alleles, or were monomorphic with the case of RM213 and RM244 (Plate 6.2a-e). There were several alleles unique to Kaski landraces at a lower frequency (less than 0.05) (Figure 6.4). These were at loci RM164, RM211, RM215, RM232, RM206, RM20, RM48, and RM21.

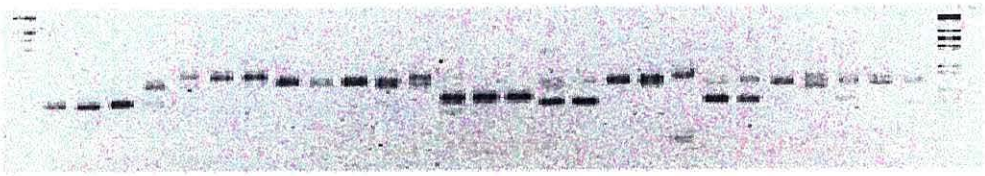
The PIC values were estimated over 27 accessions (the whole set with 3 check varieties) and 24 accessions (without check varieties). The highest values of 0.77 and 0.72 were recorded for RM247 on chromosome 12 (Table 6.3).

Table 6.3: Number of alleles and polymorphic information content (PIC) of each of SSR locus in various rice accessions from the Kaski ecosite.

S.No	Primer	Chromosome	Allele			PIC	
			Total sample	Landrace accessions	Check varieties	Total sample	Landrace accessions
1	RM246	1	4	4	3	0.54	0.44
2	RM5	1	3	3	2	0.66	0.66
3	RM122	2	3	2	1	0.2	0.22
4	RM211	2	3	3	2	0.29	0.26
5	RM213	2	2	1	2	0.07	0
6	RM48	2	3	3	1	0.5	0.51
7	RM203	2	2	2	1	0.4	0.30
8	RM60	3	1	1	1	0	0.
9	RM232	3	3	3	2	0.46	0.44
10	RM22	3	2	2	1	0.35	0.22
11	RM148	3	2	2	1	0.05	0.12
12	RM226	4	3	3	1	0.5	0.43
13	RM349	4	2	2	1	0.48	0.5
14	RM164	5	4	4	2	0.63	0.56
15	RM26	5	2	2	1	0.49	0.5
16	RM3	6	3	2	1	0.49	0.38
17	RM234	7	3	2	2	0.50	0.43
18	RM248	7	4	3	2	0.49	0.38
19	RM11	7	2	2	2	0.38	0.33
20	RM351	7	2	2	1	0.30	0.15
21	RM350	8	2	2	1	0.10	0.12
22	RM223	8	2	2	1	0.50	0.5
23	RM215	9	2	2	1	0.07	0.08
24	RM257	9	3	2	2	0.38	0.25
25	RM242	9	3	2	2	0.44	0.33
26	RM228	10	4	3	1	0.54	0.43
27	RM222	10	2	2	1	0.50	0.5
28	RM244	10	2	1	2	0.07	0
29	RM167	11	2	2	1	0.20	0.22
30	RM206	11	4	2	2	0.52	0.4
31	RM229	11	2	2	1	0.38	0.28
32	RM224	11	3	3	2	0.43	0.42
33	RM17	11	3	3	2	0.62	0.57
34	RM21	11	4	3	2	0.39	0.24
35	RM120	11	1	1	1	0	0
36	RM202	11	3	2	2	0.42	0.39
37	RM247	12	6	5	1	0.77	0.72
38	RM20	12	4	4	1	0.53	0.44
Total number of alleles			105	91	56		
Total number of accessions			27	24	3		
Average alleles per locus (A)			2.7	2.4	1.5		
Number of polymorphic loci			36	34	16		
Number of polymorphic alleles			103	87	34		
Av. alleles per poly. Locus (Ap)			2.8	2.6	2.1		
PPL†			94.7	89.5	42.1		
PPA‡			98.1	95.6	60.7		
Average gene diversity (PIC)						0.39	0.33

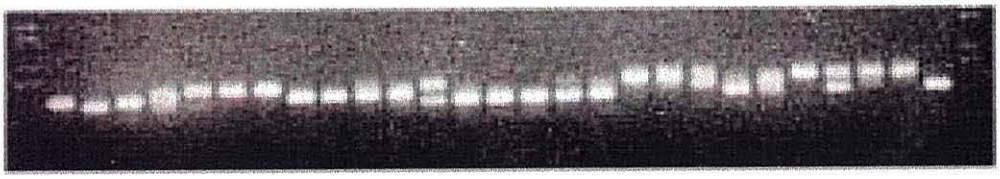
PPL† = percentage of polymorphic loci; PPA‡ = percentage of polymorphic alleles

(a) RM247 with 5 alleles in landraces (negative photograph)

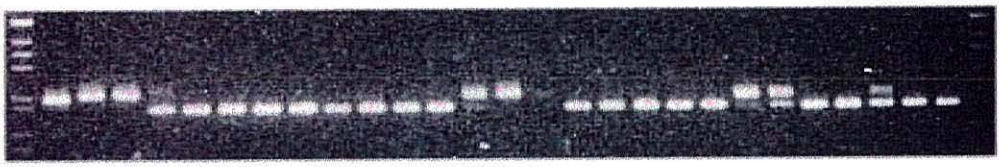


(b) RM164 with 4 alleles in landraces

(c) RM5 with 3 alleles in landraces



(d) RM242 with 2 alleles in landraces



(e) RM213 with monomorphic allele in landraces

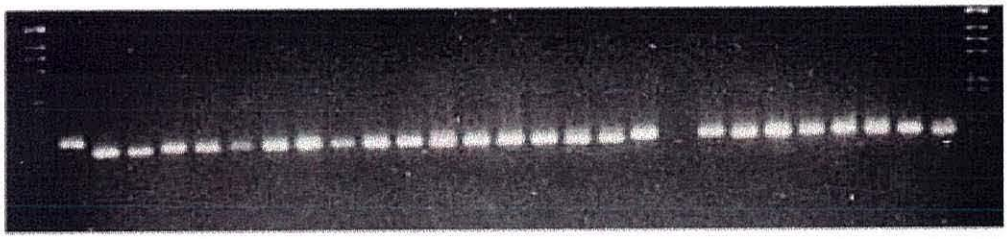


Plate 6.2a-e: Microsatellite DNA polymorphism in rice landraces from the Kaski ecosite detected by 5 different primer pairs showing banding patterns. The amplified microsatellite DNAs were separated in 3 % agarose gel and detected by ethidium bromide.

(From left to right: Ladder, Kalinga III, IR64, IR36, Kathe Gurdi: 025, 028, 029, and 033; Jetho Budho: 203, 205, 207, and 214; Gurdi: 058; Rato Anadi: 092, 094, and 104; Aanga: 219, and 224; Ramani: 158, and 161; Seto Gurdi: 052, and 053; Jhinuwa: 012, and 013; Tunde Jhinuwa: 014, and 015; Kalo Bayarni: 070, and 072 and ladder).

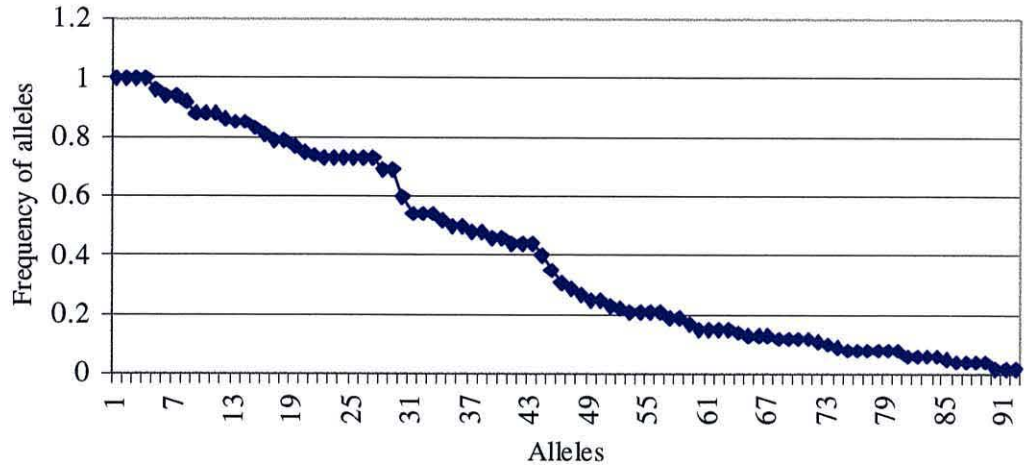


Figure 6.4: Frequency distribution of alleles in 24 Kaski landraces produced by 38 markers.

6.3.2.2 Clustering and principal component analysis

The clustering of rice landraces from Kaski ecosite based on 103 polymorphic alleles showed that landraces had a considerable level of genetic diversity within and between different FUDs. With the genetic dissimilarity coefficient < 0.24, the 27 rice accessions could be divided into 7 clusters by UPGMA (Figure 6.5a) that were related to the agro-morphological and ethno-botanical classifications. For example, landraces with the names Rato Anadi, Jetho Budho, Ramani, Anga and Gurdi landraces formed distinct clusters. However, Jhinuwa, Tunde Jhinuwa, Seto Gurdi and Kalo Bayarni did not cluster in groups showing variation among landraces with the same name. The cluster analysis of these SSR products without the check varieties gave a similar division of the landraces into the groups (Figure 6.5b). The use of the Jaccard's similarity coefficient produced similar result to Nei's genetic distance (Figure 6.5c). Among 24 landrace accessions, two accessions of Kathe Gurdi (29KG and 33KG) and two accessions of Jetho Budho (203JB

and 205JB) were observed genetically identical populations. Aanga (219A and 224A) and Rato Anadi (92RA, 94RA and 104RA) formed the distinct clusters.

The PCAs, whether done with or without the check varieties (Figures 6.6 and 6.7) were relatively less informative than the dendrograms, and failed in most cases to identify landrace groupings. The only exceptions were Rato Anadi that were clearly separated whether or not the check varieties were included. However, Kathe Gurdi, Gurdi, Seto Gurdi, Jhinuwa, Tunde Jhinuwa, Ramani, and Kalo Bayarni formed a tight and mixed grouping.

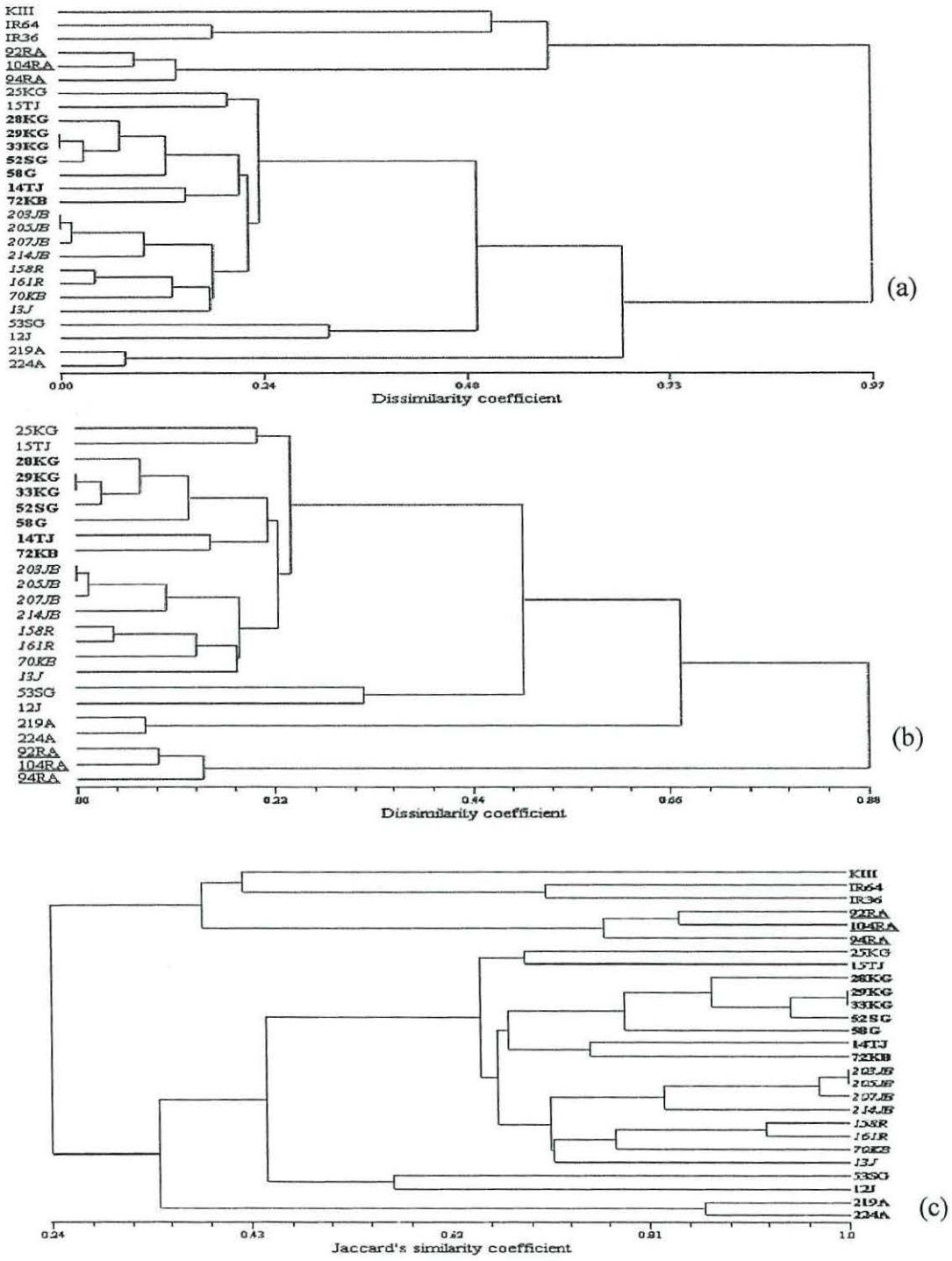


Figure 6.5: Dendrograms of rice landraces from Kaski obtained UPGMA clustering based on polymorphic alleles (a) with Neis' genetic distance including landraces and check varieties; (b) with Neis genetic distance excluding the check varieties; and (c) with Jaccards' similarity coefficient including landraces and check varieties. (RA=Rato Anadi; KG=Kathegurdi; G=Gurdi; SG=seto Gurdi; TJ=Tunde Jhinuwa; KB=Kalo Bayarni; JB=Jetho Budho; J=Jhinuwa and A=Aanga)

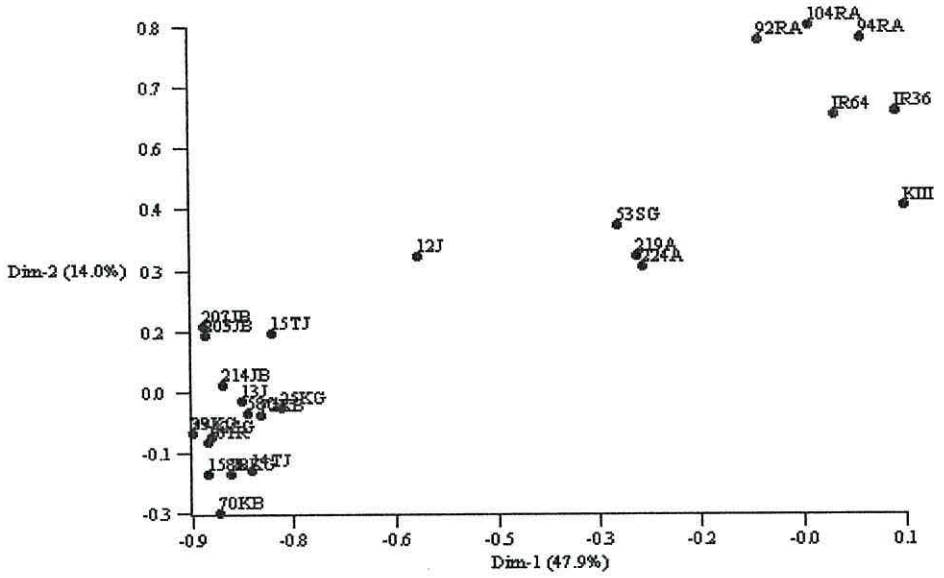


Figure 6.6: Scatter plot of rice accessions including landraces and 3 check varieties based on principal component analysis of alleles generated by polymorphic SSR markers (the Kaski ecosite).

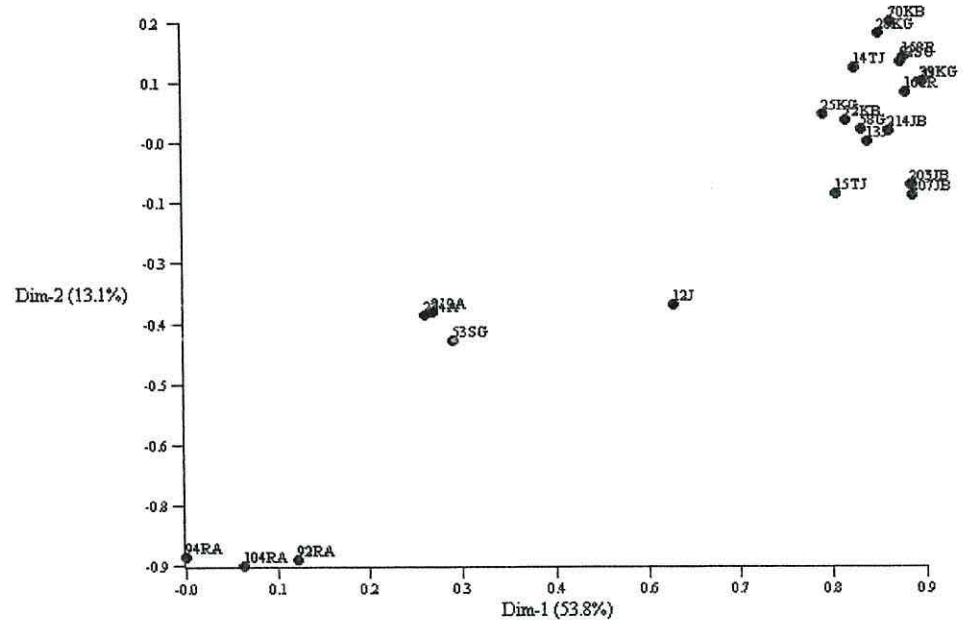


Figure 6.7: Scatter plot of landrace accessions only excluding check varieties based on principal component analysis of alleles generated by polymorphic SSR markers (the Kaski ecosite).

6.3.3 Bara: the lowland ecosite

6.3.3.1 Genetic diversity of microsatellite (SSR) markers

All the microsatellite primer pairs except RM201 produced prominent PCR products in all 27 rice accessions. With 38 primer pairs, 102 alleles were detected in the landraces and checks; 95 alleles in the 25 landraces; and 56 in the 3 check rice varieties (Table 6.4). Number of alleles at each of locus ranged from 1 to 4 with an average of 2.7 for landraces and checks; 2.5 for the landraces; and only 1.5 for the 3 check varieties. A maximum of four alleles were resolved at each of 7 loci (RM48, RM232, RM226, RM248, RM257, RM167, RM206 and RM21). Two to three alleles were identified for most of the loci except for RM122, RM22, and RM244, which were monomorphic for all the rice accessions from Bara (Plate 6.3a-d) and RM60 and RM120 that monomorphic across the landraces and checks. Several allelic variants unique to Bara landraces were revealed occurring at a lower frequency (less than 0.05). These were at loci RM226, RM246, RM232, RM213, RM215, RM257, RM167, and RM26 (Figure 6.8).

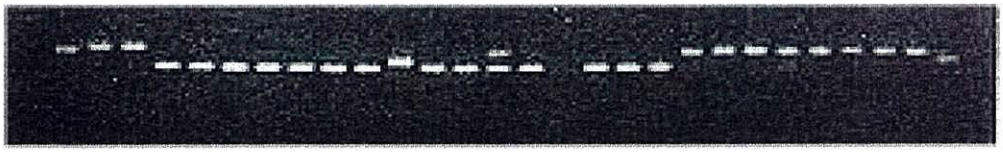
The PIC values were relatively high for most loci, averaging 0.35 (landraces alone) to 0.40 (landraces and checks) and the highest values of 0.66 and 0.68 were recorded for RM21 on chromosome 11(Table 6.4).

Table 6.4: Number of alleles and polymorphic information content (PIC) of each of SSR locus in various rice accessions from the Bara ecosite.

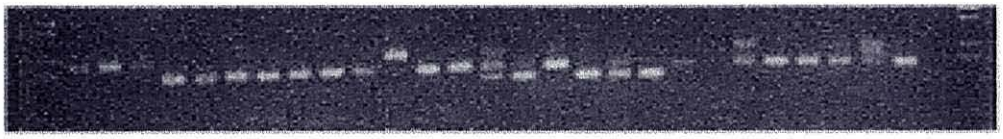
S.No	Primer	Chromosome	Allele			PIC	
			Total sample	Landrace accessions	Check varieties	Total sample	Landrace accessions
1	RM246	1	3	3	3	0.33	0.25
2	RM5	1	2	2	2	0.46	0.44
3	RM122	2	2	1	1	0.19	0
4	RM211	2	2	2	2	0.16	0.11
5	RM213	2	2	2	2	0.13	0.08
6	RM48	2	4	4	1	0.48	0.55
7	RM203	2	2	2	1	0.29	0.32
8	RM60	3	1	1	1	0	0
9	RM232	3	4	4	2	0.66	0.66
10	RM22	3	2	1	1	0.19	0
11	RM148	3	2	2	1	0.37	0.38
12	RM226	4	4	4	1	0.36	0.36
13	RM349	4	2	2	1	0.13	0.15
14	RM164	5	3	3	2	0.59	0.58
15	RM26	5	2	2	1	0.07	0.08
16	RM3	6	3	3	1	0.52	0.44
17	RM234	7	3	3	2	0.64	0.63
18	RM248	7	4	4	2	0.62	0.56
19	RM11	7	2	2	2	0.49	0.5
20	RM351	7	2	2	1	0.5	0.5
21	RM350	8	2	2	1	0.13	0.15
22	RM223	8	2	2	1	0.32	0.34
23	RM215	9	2	2	1	0.07	0.08
24	RM257	9	4	4	2	0.6	0.56
25	RM242	9	3	2	2	0.53	0.49
26	RM228	10	3	2	1	0.5	0.38
27	RM222	10	2	2	1	0.39	0.42
28	RM244	10	2	1	2	0.13	0
29	RM167	11	4	4	1	0.39	0.42
30	RM206	11	4	4	2	0.64	0.6
31	RM229	11	3	3	1	0.23	0.22
32	RM224	11	3	3	2	0.59	0.6
33	RM17	11	3	2	2	0.56	0.5
34	RM21	11	4	4	2	0.68	0.66
35	RM120	11	1	1	1	0	0
36	RM202	11	3	2	2	0.26	0.15
37	RM247	12	3	3	1	0.47	0.51
38	RM20	12	3	3	1	0.58	0.6
Total number of alleles			102	95	56		
Total number of accessions			28	25	3		
Average alleles per locus (A)			2.7	2.5	1.5		
Number of polymorphic loci			36	33	16		
Number of polymorphic alleles			100	90	34		
Av. alleles per poly. Locus (Ap)			2.8	2.7	2.1		
PPL†			94.7	86.8	42.1		
PPA‡			98.0	94.7	60.7		
Average gene diversity (PIC)						0.4	0.35

PPL† = percentage of polymorphic loci; PPA‡ = percentage of polymorphic alleles

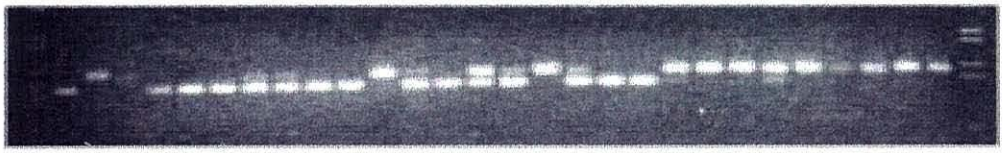
(a) RM257 with 4 alleles in landraces



(b) RM3 with 3 alleles in landraces



(c) RM11 with 2 alleles in landraces



(d) RM244 with monomorphic allele in landraces

Plate 6. 3a-d: Microsatellite DNA polymorphism in rice landraces from the Bara ecosite detected by 4 different primer pairs. The amplified microsatellite DNAs were separated in 3 % agarose gel and detected by ethidium bromide.

(From left to right: Ladder, Kalinga III, IR64, IR36, Mutmur: 065, 066, and 301; Nakhisaro: 118, 117, and 308; Sathi: 012, and 018; Sokan: 183, and 307; Mansara: 146, 150 and 148; Basmati: 052, 044, and 320; Karma: 033, 036 and 029; Lajhi: 101, 104, and 097; Dudhraj: 215; Laltenger: 304, and 125 and ladder).

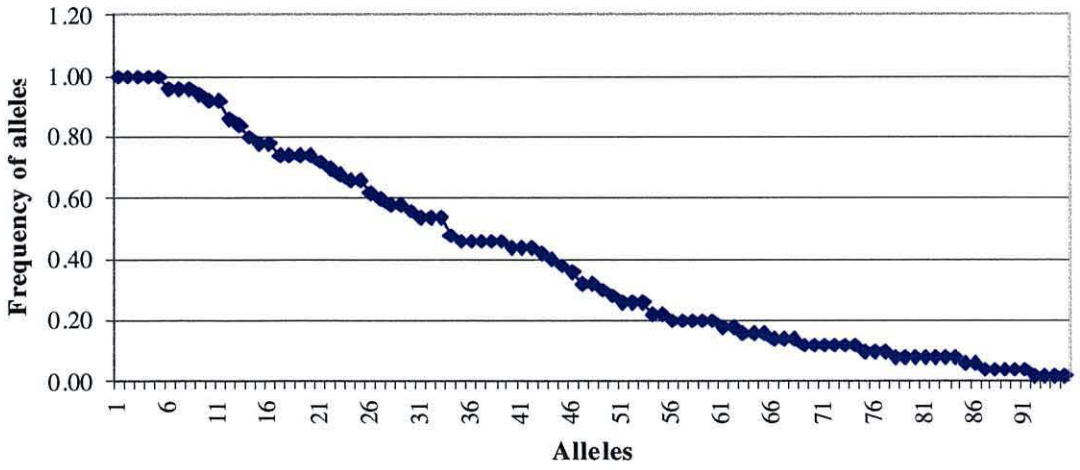


Figure 6.8: Frequency distribution of alleles in 25 Bara landraces produced by 38 markers.

6.3.3.2 Clustering and principal component analysis

In the dendrogram based on 100 polymorphic alleles the rice accessions from Bara were included in 5 major clusters at the dissimilarity level < 0.28 (Figure 6.9a). The clustering clearly discriminated the early varieties from remainder of the rice varieties (both normal and late). These were in clusters:

- ten accessions of normal & late rice varieties: Lajhi, Dudhraj, Karma, Mansara and Laltenger
- five accessions of normal & late rice varieties: Basmati and Mansara
- nine accessions of early maturing varieties: Mutmur, Sokan, Nakhisaro, and Sathi

The early varieties could be further differentiated into three sub-clusters: cluster for Mutmur accessions; cluster for Nakhisaro accessions; and cluster for Sathi (12S) and Sokan (307So).

An accession of Sokan (183) and Sathi (18S) were two early rice varieties which separated clearly from other accessions and did not fall in any of the cluster. Similar grouping of the rice landraces were obtained without the check varieties (Figure 6.9b) as was also the case with Jaccard's similarity coefficient (Figure 6.9c). Two accessions of Mutmur (65M and 66M) and two accessions of Nakhisaro (117N and 118N) were identical to each other.

The PCA of Bara landraces showed similar groupings of accessions as observed in dendrograms, so both the analyses were equally informative (Figures 6.10 and 6.11). Few exceptions of Mansara (146Ma and 150Ma) grouped in the cluster of early rice varieties. Basmati (44B, 52B and 320B) and Laltanger (125La and 304La) were clearly separated from rest accessions in both the PCA analyses with and without check varieties.

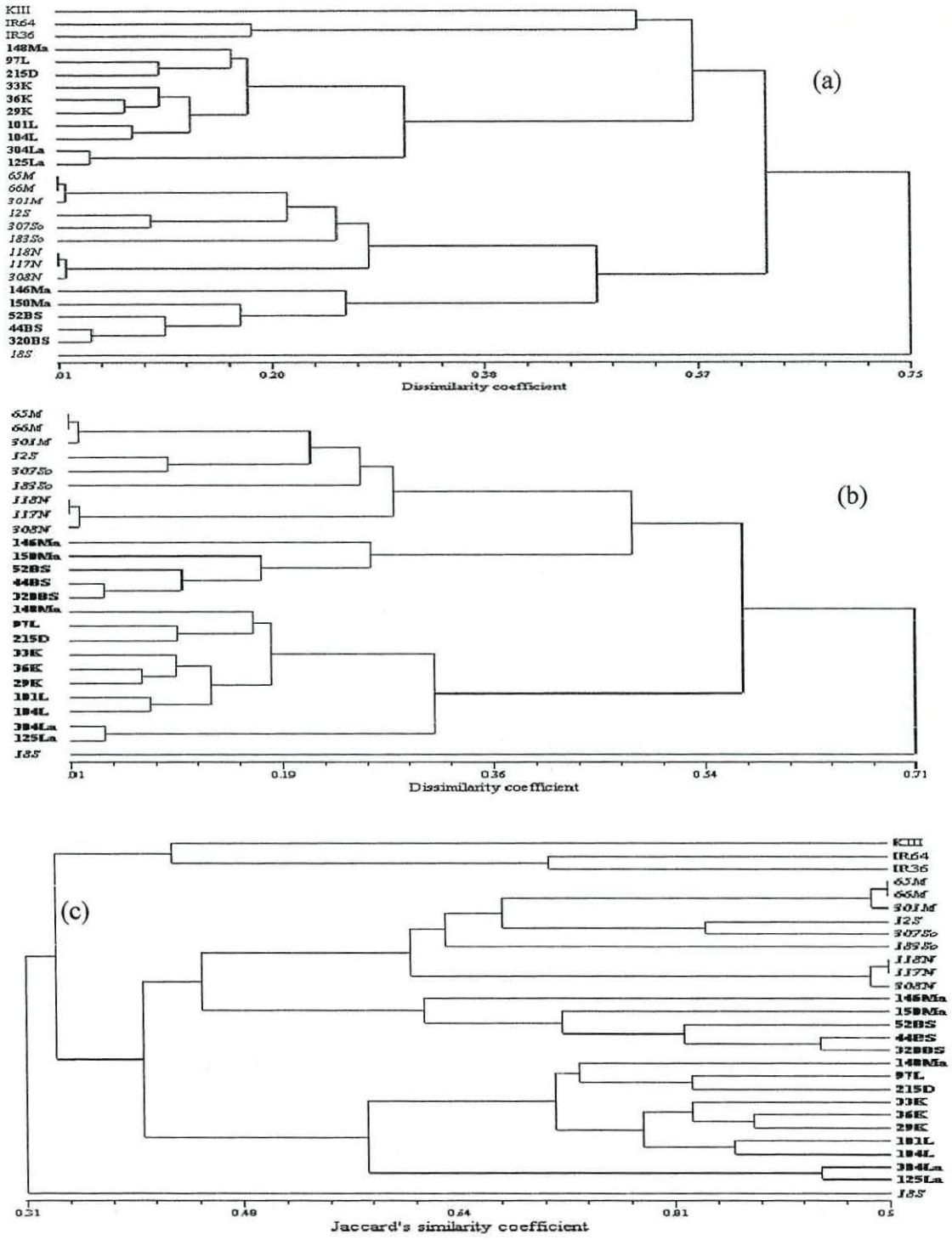


Figure 6.9: Dendrograms of rice landraces from Bara obtained UPGMA clustering based on polymorphic alleles (a) with Neis' genetic distance including landraces and check varieties; (b) with Neis genetic distance excluding the check varieties; and (c) with Jaccards' similarity coefficient including landraces and check varieties. (M=Mutmur; So=Sokan; S=Sathi; N=Nakhisaro; Ma=Mansara; BS=Basmati; L=Lajhi; D=Dudhraj; K=Karma; and La=Laltenger).

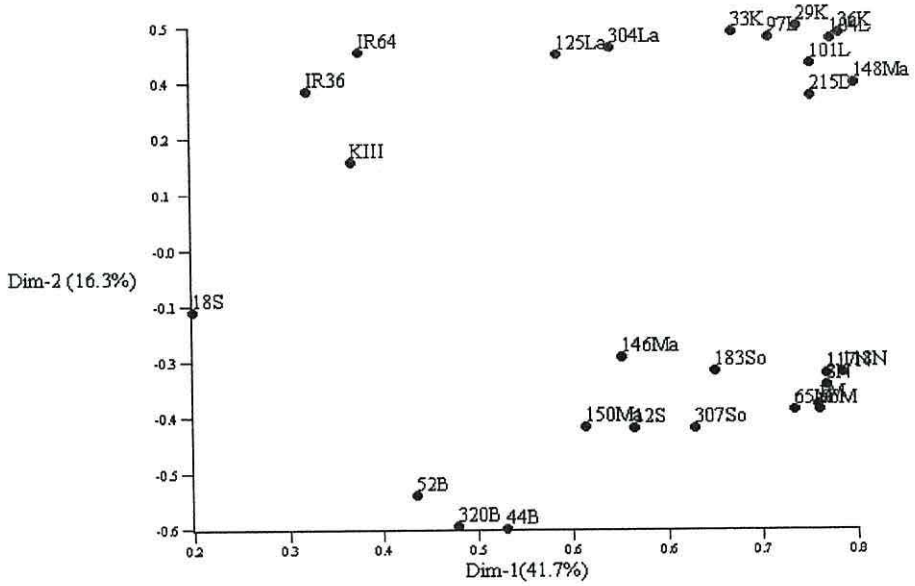


Figure 6.10: Scatter plot of landrace accessions including check varieties based on principal component analysis of alleles generated by polymorphic SSR markers (the Bara ecosite).

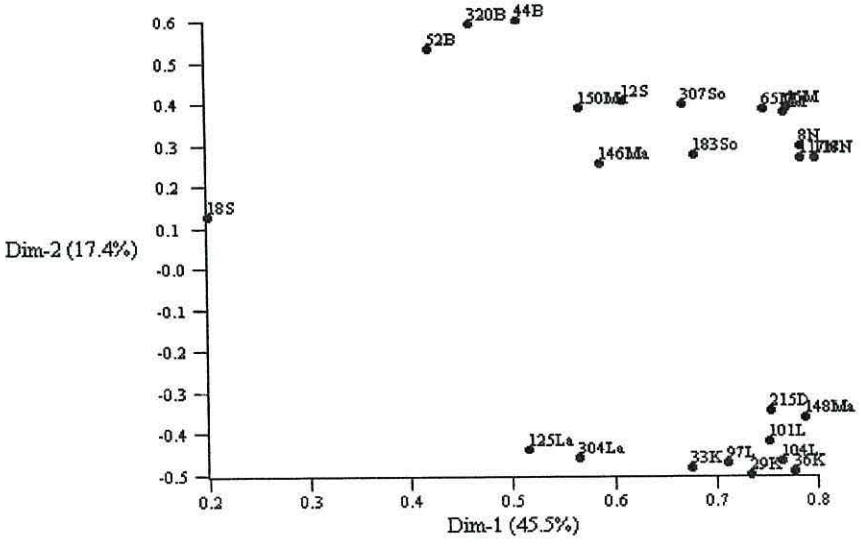


Figure 6.11: Scatter plot of landrace accessions excluding check varieties based on principal component analysis of alleles generated by polymorphic SSR markers (the Bara ecosite).

6.3.4 Comparison of genetic diversity between rice landraces from three ecosites

A comparison of the genetic diversity of rice accessions was performed among 3 sets of landraces from three different agro-ecosystems of the country. The set of landraces from the mid-hill ecosite (Kaski) and the lowland ecosite (Bara) were more diverse, producing a significantly greater average number of alleles ($P < 0.0001$) and higher PIC values ($P < 0.0001$) than the landraces from the Jumla ecosite (Table 6.5). The average gene diversity (PIC) values, a reflection of allelic diversity and their frequency in landraces was found significantly correlated with the number of alleles across the sites as expected (0.78, Jumla; 0.69, Kaski; and 0.73, Bara; $P < 0.01$). The average genetic dissimilarity coefficients (Nei's genetic distance) showed that the Kaski and Bara landraces are equally diverse and much more so than the Jumla landraces (Table 6.5).

Table 6.5: Summary of microsatellite polymorphism in rice landraces and their comparison between three ecosites.

Diversity parameters	Jumla	Kaski	Bara
Total size of accessions	20	24	25
Total number of primers	39	38	38
Total number of polymorphic primers detected	1	34	33
Total number of bands amplified	40	91	95
Average number of bands per primer (A)	1.0	2.4	2.5
Number of polymorphic bands identified	1	87	90
Percentage of polymorphic bands (PPA)	2.5	95.6	94.7
Average number of bands per poly. Primer (Ap)	1.0	2.6	2.7
Average genetic dissimilarity	0.12	0.45	0.45
Average gene diversity (PIC)			
(a) all loci	0.00	0.33	0.35
(b) Polymorphic loci	0.05	0.37	0.40

The PIC values for the microsatellite (SSR) markers were compared across chromosomes for 1 to 3 sets of landraces both with and without check varieties (Figures 6.12a and 6.12b). The level of diversity differed between Kaski and Bara rice landraces

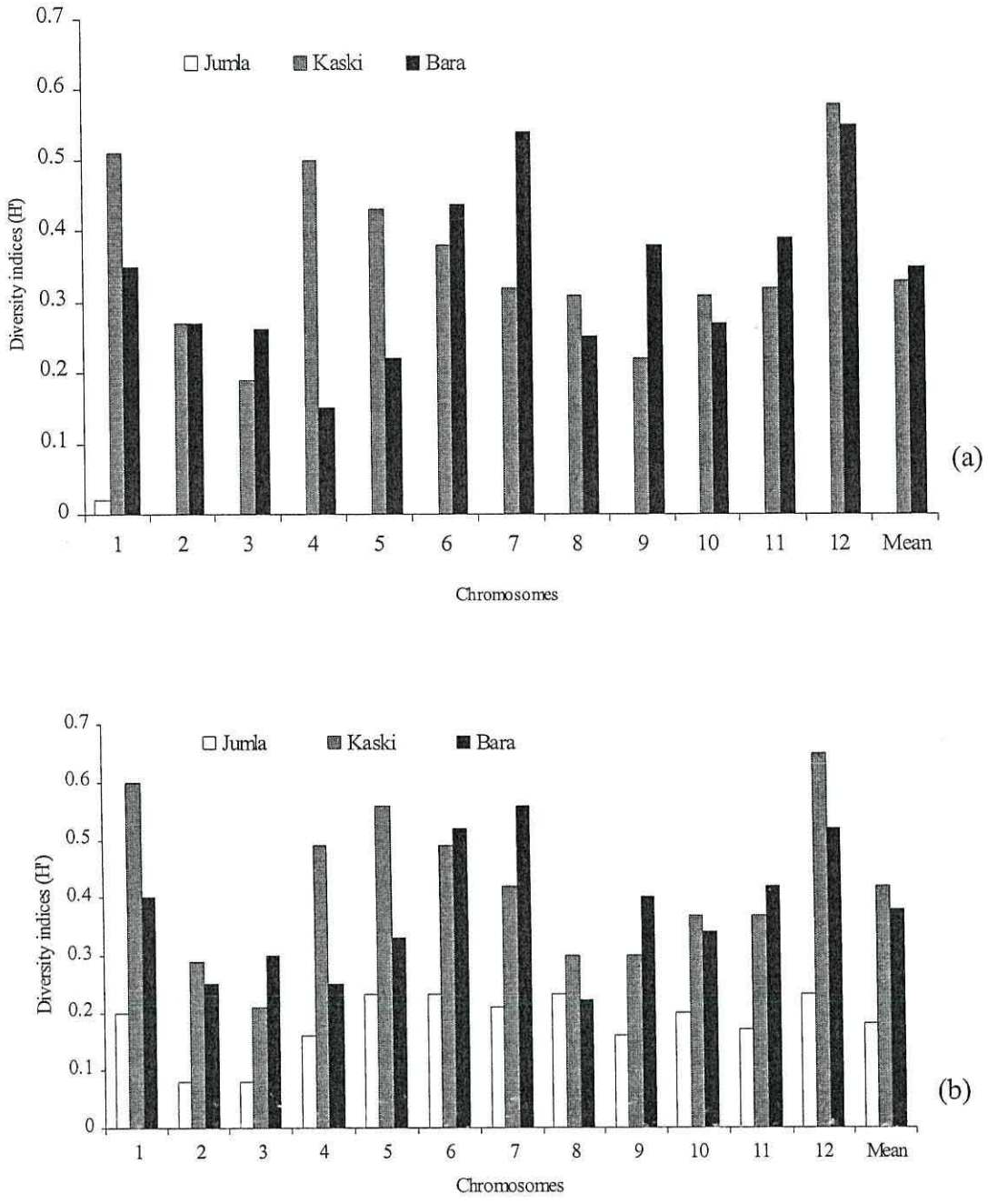


Figure 6.12: Comparison of polymorphic information content of 38 SSR markers distributed on different rice chromosomes within the landraces from three ecosites: (a) average PIC over landrace accessions only, and (b) average PIC over total set of rice accessions.

on all the chromosomes based on the average PIC values except the chromosome 2 (Figure 6.12a). However there was no significant differences in the level of diversity among landraces of two sites. The high level of genetic variation over the 12th chromosome was observed both in Kaski and Bara landrace genotypes and it was not affected by the check varieties.

6.4 Discussion

6.4.1 Diversity analysis

Each of the microsatellite primer pairs detected variation at a single locus with alleles varying from 1-6 per locus. When only the landraces were considered, out of the 39 primer pairs studied, 34 showed polymorphism in at least one of the ecosites. When the check varieties were also included then 36 primer pairs showed the polymorphism. Hence set of microsatellites (SSRs) used in the present study have detected diversity within and among the landraces of the ecosites.

The patterns of genetic diversity varied among the study sites. Landraces from Bara and Kaski exhibited high genetic variation (0.33 - 0.35 PIC) and most of the markers were polymorphic (87 - 90%). In contrast, very little variation was encountered within the landraces from Jumla. Only a single marker (3%) was polymorphic and that diversity was displayed by only one landrace. The rice diversity in Jumla was therefore much lower than in Kaski and Bara (Table 6.12). Indeed, the landraces in Jumla were essentially monomorphic and genetically homogenous for the SSR studied. In contrast, the PIC was recorded as high as 0.72 for landraces from Kaski with RM247 (Table 6.3). The value of gene diversity (PIC) mostly increased with the number of alleles at a locus.

Therefore, the number of alleles and the gene diversity were important parameters for the evaluation of the genetic diversity. However, in comparison with the results from previous studies using microsatellites and other types of markers including isozyme on the genetic diversity of rice cultivars (*indica* and *japonica* groups), landraces and wild species, these diversity levels are much lower (Second, 1982; Glaszman, 1987; Oka, 1988; Wang and Tanksley, 1989; Zhang *et al.*, 1992; Xiao *et.al*, 1996; Yang *et al.*, 1994; Davierwala *et al.*, 2000; Qian *et.al.*, 2001 and Blair *et al.*, 2002). Yang *et. al.,* (1994) detected up to 25 diverse alleles for a single microsatellite marker RM163 in 140 rice landraces from various parts of China, Japan and India. However, it was observed very low in the present study and not comparable as the diversity depends on the number and type of SSRs chosen and the rice materials included in the study. Thanh *et.al.*, (1999) also showed significant variation in microsatellite DNA polymorphisms among 31 upland rice accessions of Vietnam. Likewise series of studies on the genetic diversity of natural populations of cultivated rice and common wild rice have reported the diversity at different morphological, isozyme and DNA levels (Zhang *et. al.*, 1994, Bustos *et.al.*, 1998, Ge *et.al.*, 1999, Gao *et.al.*, 1999, and Sun *et.al.*, 2001,).

6.4.2 Diversity observed in clustering and principal component analyses

The landraces were, as expected, clearly distinct from the check modern varieties. The landraces in Jumla clustered into a single group irrespective of the FUDs, so the information from microsatellites does not agree with the identification of varieties by farmers. Sebastian *et al.* (1998) in an assessment of diversity and identity of farmers' rice varieties of Philippines using microsatellite markers and Busso *et al.* (2000) in pearl

millet found a similar trend. Variations in landraces were not related to the names but was more related to the farmers and the traditional practices of seed management.

However, the cluster analyses of the landraces from Kaski and Bara (Figures 6.5 and 6.9) supported much more the traditional classification of landraces as FUDs and their differentiation that was found on an agro-morphological basis in Chapter V. It was apparent among landraces from the Kaski ecosite where landraces named Rato Anadi, Aanga, Jetho Budho, Ramani, and Gurdi formed distinct clusters. Similarly, landraces called Basmati, and Laltenger from Bara formed distinct sub-clusters of normal duration rice. In Bara, the clustering was more pronounced by the growth duration of the landraces. However, the grouping of some accessions with the same name was not consistent in both Kaski and Bara. Accessions named Jhinuwa, Tunde Jhinuwa, Kalo Bayarni and Kathe Gurdi from Kaski and Sathi and Mansara from Bara either belonged to different clusters or were not clustered at all. These tended to be the landraces grown in small area by few HHs; this could be as a result of intra-varietal variations, or simply from misnaming by farmers, or the results of sampling effect during sample collection or molecular works.

A very similar grouping of landraces was observed in principal component analyses across the ecosites. However, PCA with and without check varieties appeared less informative than the cluster analysis as the diversity concerned with the FUDs.

6.4 Conclusions

- Microsatellites were useful markers. They detected a high level of allelic variation and gave an understanding of the genetic relationships and diversity among landrace accessions from the three different agro-ecozones of the country.
- Rice landraces from the high-hill (Jumla) ecosite had a narrow genetic base and showed homogeneity in their allelic composition with common allele for all the 39 microsatellite (SSR) markers evaluated.
- This pattern of DNA variation is probably due to the founder effect of a single population of Jumli Marshi from which, by continuous selection by farmers for major gene characteristics, a morphological range of landraces (FUDs) has evolved. Molecular diversity will only be found in those regions where major gene changes have occurred. This molecular uniformity is even more remarkable considering the variation found even with supposedly pure lines (Chapter VIII).
- A much large genetic diversity was detected in the landraces from Kaski and Bara that was also related to the morphological differences and traditional classifications of these landraces within each ecosite.
- The most abundantly grown genotypes had little variation within a named group (they clustered together) but the genotypes grown in small area and by only few households, such as Sathi, Sokan from Bara and Jhinuwa, Seto Gurdi and Kalo Bayarni from Kaski exhibited more variation (they did not cluster together).

CHAPTER VII

COMPARISION OF AGRO-MORPHOLOGICAL AND SSR MARKER DIVERSITIES IN RICE LANDRACES

7.1 Introduction

Both agro-morphological and molecular markers have been widely used to assess diversity in rice and have provided valuable information on genetic relationships between rice varieties at various levels (Davierwala *et al.,* 2000). Agro-morphological characters, though commonly used, can vary with the environment unlike molecular markers, which detect variation at the DNA level, are not affected by the environment. However, the ability to resolve genetic variation at the molecular marker level is directly related to how well the markers detect polymorphism (Yee *et al.*, 1999). In the present study, both agro-morphological traits and microsatellite (SSR) markers were used to detect the genetic relationships and diversity in a collection of rice landraces from three ecosites in Nepal. In this chapter, the diversity revealed by each of these two methods is compared.

7.2 Materials and methods

The set of landraces under microsatellite (SSR) marker diversity study in Chapter VI were used for the comparison made in this chapter (Table 6.1 in Chapter VI). For the comparison of the agro-morphological variability and molecular marker diversity among the landraces within the ecosite and among the ecosites, the data from all 3 ecosites was pooled for both agro-morphological traits and molecular traits. For both types of data, a new data matrix with 68 landrace accessions from three ecosites was formed. Two accessions: 1052-1 from Jumla (no amplification for many markers) and 304 from Bara

(poor germination in the field trial) were excluded from this analysis. Thirty three agro-morphological traits that were evaluated at all three sites were pooled and used to calculate average taxonomic distance matrices.

All statistical analyses were performed by NTSYS-pc version 1.8 (Rohlf, 1993). Dissimilarity matrices at the phenotypic level, using average taxonomic distance, (Sneath and Sokal, 1973) were calculated for qualitative and quantitative traits together, for qualitative traits and quantitative traits alone. This was done separately for each ecosite and also for the combined data of the three ecosites. Jaccard's similarity coefficient and Neis' genetic distance matrices at the DNA level for landraces from individual ecosites and across ecosites were constructed for the molecular data (Jaccard, 1908; and Nei, 1972). Principal component analysis was carried out on pooled agro-morphological and allelic data separately for the landraces across the ecosites (Sneath and Sokal, 1973).

Nei's genetic distance matrix and Jaccard's similarity coefficient matrix based on molecular data were compared by the Mantel test statistic with the taxonomic distance matrix based on agro-morphological data (Mantel, 1967) using the MXCOMP function of the NTSYS-pc (Rohlf, 1993; Autrique *et al.*, 1996; and Martinello *et al.*, 2001). In this test a correlation greater than 0.5 for two independent matrices (similarity or dissimilarity matrices) is significant at the 1% level (Lapointe and Legendre, 1992).

7.3 Results

7.3.1 Comparison of agro-morphological and SSR marker diversity measures

The matrices of average taxonomic distances for each site based on qualitative and quantitative traits together and one each for qualitative and quantitative traits alone were

obtained and were compared with the matrices of Nei's genetic distance obtained for SSR markers (Table 7.1) and Jaccard's similarity coefficients (Table 7.2). There were non-significant correlations between the matrices when compared individually for each site. However, when across site data were compared the correlations were all significant, and that with combined qualitative and quantitative traits was highest (Table 7.1 and 7.2).

Table 7.1: Comparison of the matrices of Nei's genetic distance for molecular marker data and taxonomic distance estimates for agro-morphological data made in rice landraces using Mantel test in NTSYS pc (Rohlf, 1993).

Variables	**Jumla (n=20)**	**Kaski (n=24)**	**Bara (n=24)**	**All three sets combined (n=68)**
MM† with Q+Q‡	- 0.11	+ 0.03	- 0.03	+ 0.55
MM with Ql.¶	- 0.06	+ 0.06	+ 0.05	+ 0.50
MM with Qn.§	- 0.16	- 0.02	- 0.08	+ 0.50

† = Average Neis genetic distance matrix based on molecular marker traits
‡ = Average taxonomic distance matrix of qualitative and quantitative traits
¶ = Average taxonomic distance matrix of qualitative traits
§ = Average taxonomic distance matrix of quantitative traits

Table 7.2: Comparison of matrices between Jaccard's similarity coefficients for molecular marker data and taxonomic distance estimates for agro-morphological data made in rice landraces using Mantel test in NTSYS pc (Rohlf, 1993).

Variables	**Jumla (n=20)**	**Kaski (n=24)**	**Bara (n=24)**	**All three sets combined (n=68)**
MM† with Q+Q‡	+ 0.11	- 0.06	- 0.01	- 0.69
MM with Ql.¶	+ 0.06	- 0.08	- 0.10	- 0.63
MM with Qn.§	+ 0.16	- 0.00	+ 0.07	- 0.62

† = Jaccard's similarity matrix based on molecular marker traits
‡ = Average taxonomic distance matrix of qualitative and quantitative traits
¶ = Average taxonomic distance matrix of qualitative traits
§ = Average taxonomic distance matrix of quantitative traits

7.3.2 Combined PCA analysis of agro-morphological traits and SSR markers and genetic relationships among landraces from three ecosites

In the principal component analysis (PCA) plots based on agro-morphological traits, a clear separation of the landrace genotypes could be seen according to the agro-ecosystems (except two accessions of Aanga from Kaski). Nearly 53% of the total variation was explained and the accessions were grouped into 3 distinct clusters (Figure 7.1). However, in the PCA based on the SSR markers, the accessions were grouped into 4 clusters and 60% of the total variation was explained (Figure 7.2). The Jumla landraces made a highly distinct single cluster but the distinction between the Bara and Kaski landraces was less than for the agro-morphological traits (Figure 7.2).

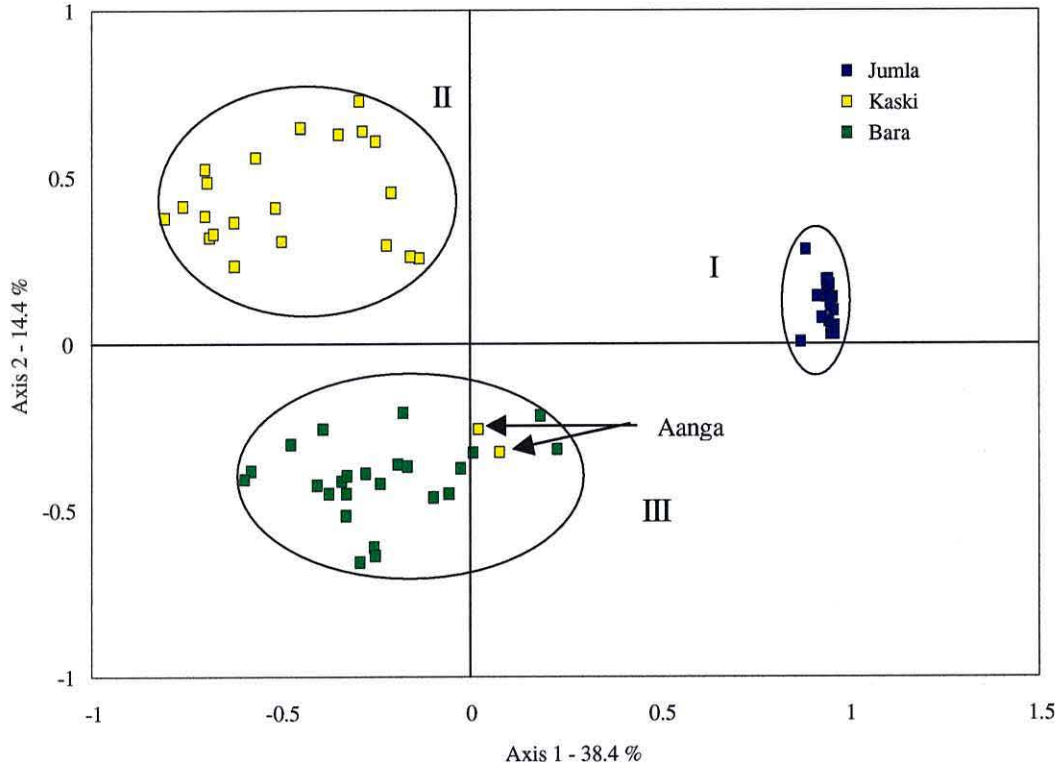


Figure 7.1: Scatter plot of 68 rice landraces based on PCA of 33 significant qualitative and quantitative traits.

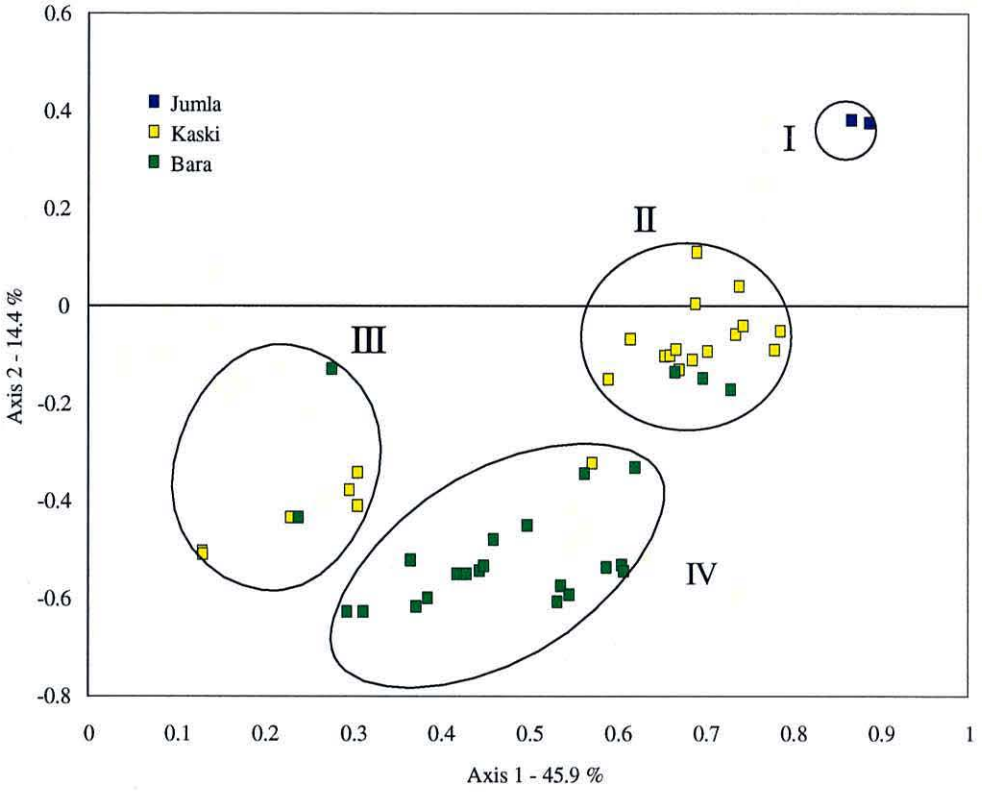


Figure 7.2: Scatter plot of 68 rice landraces based on PCA of 117 microsatellite alleles.

The Kaski landraces based on the microsatellite markers however clustered into two groups II and III (Figure 7.2). Cluster II was comprised of the small and fine grained, aromatic landraces and cluster III was comprised of bold and glutinous grained varieties (Figure 7.2) grown in small area. The aromatic rice landraces from Bara (Basmati: 44B, 52B and 320B accessions) were clustered with group II of the Kaski landraces. Moreover, a bold-grained, late-maturing and pigmented accession (125L) from Bara was also clustered with group III of the Kaski landraces that also have bold grains (Figure 7.2).

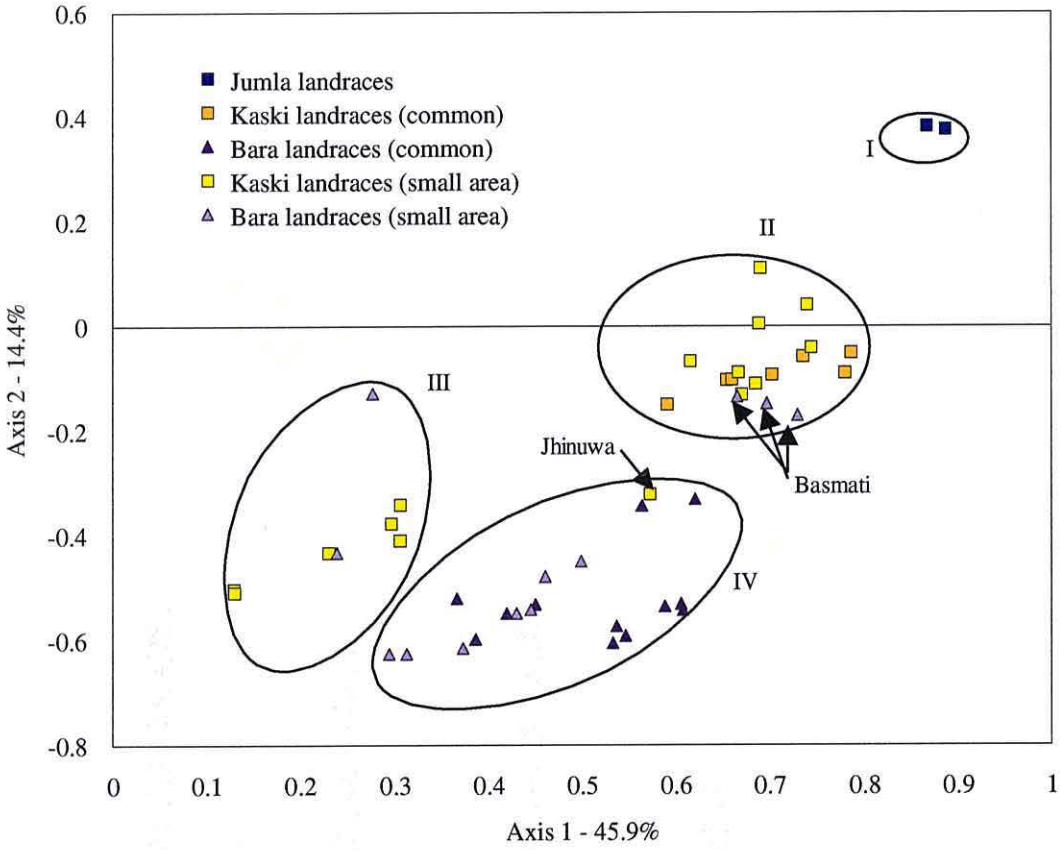


Figure 7.3: Scatter plot of 68 rice landraces based on PCA of 117 microsatellite alleles showing the dispersion of less common and specialized rice varieties from Kaski and Bara.

The agro-morphological and microsatellite markers were equally able to detect a certain level of genetic diversity among the landraces within each ecosite. The microsatellite markers were more powerful than the morphological traits as they distinguished between fine grained aromatic and the glutinous and coarse grained landraces of Kaski. The less common rice varieties were more diverse than the common ones: the landraces from both the ecosites belonging to this category were distributed in three adjacent clusters whereas the common landraces were within a single cluster for

each ecosite (Figure 7.3). Such clustering was not detected in the PCA based on the agro-morphological traits.

For each ecosite, the average taxonomic distance obtained from the qualitative and quantitative traits together and alone, using UPGMA linkage analysis, was higher than the average Nei's genetic distance and Jaccard's dissimilarity coefficients for the molecular marker data (Figure 7.4).

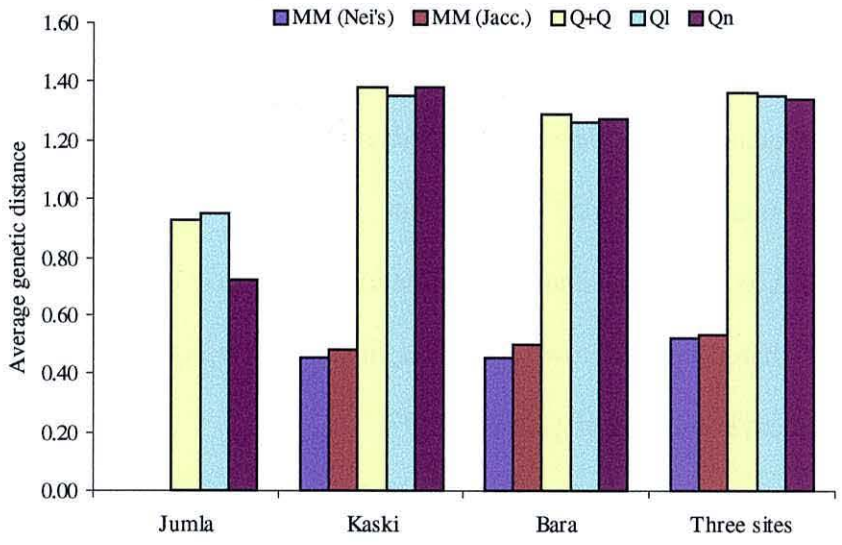


Figure 7.4: Average genetic distance obtained using microsatellites, and agro-morphological traits. (MM = molecular markers with Nei's and Jaccard's parameters; Q+Q = qualitative and quantitative traits together; Ql = qualitative traits alone; and Qn = quantitative traits alone).

7.4 Discussion

It has been demonstrated in rice that microsatellite (SSR) markers are an efficient marker system for measuring the diversity among closely related cultivars with a narrow genetic base (Yang *et al.*, 1994; Panaud *et al.*, 1996; Akagi *et al.*, 1997; and Olufowote *et al.*,

1997). However, there was a poor agreement or correlation between the matrices generated by agro-morphological traits and molecular markers data on a within site basis. Low correlations between the similarity or dissimilarity matrices based on different marker systems have been obtained for rice (Parsons *et al*, 1997; and Gao *et al.*, 2002) and many other crops: wheat (Plaschke *et al.*, 1995; and Autrique *et al.*, 1996); barley (Graner *et al*, 1994; and Schut *et al.*, 1997); oat (Beer *et al.*, 1993); and okra (Martinello *et al.*, 2001). The poor correlation or the disagreement could be due to fundamental differences in the concepts underlying these measures of genetic diversity (Davierwala *et al.*, 2000). The measure of genetic diversity by morphological markers is an indirect measure, which quantifies the degree of relatedness or differences of two genotypes for a phenotypic trait under a provided environment. In contrast, molecular markers directly sample the DNA composition and reflect relationships based on the proportion of the alleles (bands) shared between two genotypes in the DNA sequences across the entire genome with no environmental effect (Nei, 1987).

The high and significant correlations between agro-morphological and marker data across the three may be artificial. The genetic distance between sites for the agro-morphologiocal traits is confounded by environmental effects that have increased the diversity between sites. Each set of landraces were evaluated in a different site according to their origin (Chapter III). The PCA on molecular marker data provided truer estimates of genetic diversity and relationships among the landrace genotypes from three ecosites. It showed that the landrace accessions from Jumla (the high-hill ecosite) had a very narrow genetic base and were genetically distinguishable from the landraces from Kaski

and Bara. But the level of landrace diversity in Bara and Kaski was almost the same and the landraces from these two sites were not clearly separated (Figure 7.2).

The specialized rices (aromatic, glutinous or bold-grained) where conscious selection is primarily for quality traits were less geographically differentiated than the less specialized rices where conscious selection pressure is for yield and hence local adoption. Moreover, specialized landraces are grown on only a small area and hence are more susceptible to loss. They are more frequently replaced by farmers perhaps, from sources outside of the village. This would slow the genetic differentiation between ecosites. It is obvious in this molecular comparison analysis that three accessions of Basmati (Bara) and one accession of Jhinuwa (Kaski)-fine grained aromatic rices were out-clustered irrespective to the landraces of respective ecosites. The commonly grown rice varieties where conscious selection pressure is for yield were locally adapted and clustered together.

7.5 Conclusions

- The diversity estimates on agro-morphological traits of the rice accessions recorded was influenced by both genotype and environment. The microsatellite markers, on the other hand, provided estimates of genetic similarities and differences at the genome level.
- Genetic distances based on agro-morphological traits and microsatellite (SSR) markers showed a very poor correlation within sites. However, the correlations were significant when the data of all three sets of landrace accessions were combined.

However, this was biased by environmental differences between sites in the agro-morphological evaluations.

- The landraces from Jumla (the high-hill) had an extremely narrow genetic base, particularly when assessed by molecular markers, and were morphologically and genetically distinct from those of Kaski and Bara.
- The Kaski and Bara rice landraces were not very distinct on the basis of molecular markers analysed in the study. Differences may become clearer if more SSR markers are used.
- Some phenotypic traits of grain of rice varieties, such as aroma and stickiness, were important and appeared to affect the structure of genetic diversity as assessed by molecular markers.

CHAPTER VIII

COMPARISION OF LANDRACE DIVERSITY AND PPB DIVERSITY ASSESSED BY MOLECULAR MARKERS

8.1 Introduction

Rice is the most intensively evaluated and one of most polymorphic cereal crop species with the largest *ex situ* germplasm in the world (Virk *et al.,* 1996). The genus has extensive intra-specific variation, differentiation into sub-species (*sativa* and *japonica*) and further differentiation into different cultivar groups as landraces, breeding lines, and modern varieties. Landraces are most often heterogeneous with a blend of different individual plants maintained by farmers in a local environment and constitute a significant portion of cultivated rice gene pool in Asia (Yang *et al.,* 1994).

Participatory plant breeding (PPB) and participatory variety selection (PVS) are two participatory crop improvement approaches that aim to use, enhance and improve these landraces according to the needs and interest of farmers. These have helped to increase the genetic diversity in farmers' fields (Witcombe *et al.*, 1996; Sthapit *et al.*, 1996; and Witcombe *et al.*, 2001). Both landraces and varieties produced from PPB are therefore important means of maintaining genetic variation in farmers' fields.

For landraces, genetic variability is maintained not only between but also within accessions which is normally obvious. Molecular markers, along with morphological traits, have made it possible to evaluate genetic diversity contained within and between cultivars and have helped in identifying duplicate accessions in genebanks (Virk, *et al.*, 1995; and Zhu, 1996). However, there is comparatively little information available on intra-accession (cultivar) variation in landraces compared to variation between them.

Olufowote *et al.*, (1997) evaluated the within-cultivar variation in 71 phenotypically purified rice varieties using a combination of RFLP and microsatellite (SSR) markers and revealed diversity even in phenotypically homogenous populations. Luce *et al.* (2001) used microsatellites in detecting inter- and intra- varietial diversity in a collection of 419 rice accessions from gene banks in European countries and compared the diversity with a collection of 57 Asian varieties. Most cultivars were identified with distinct genotypes and 82 cultivars displayed an intra varietal polymorphism with two alleles on one or more loci. In the present study the extent of inter and intra diversity in landraces and PPB varieties were compared at the molecular level.

8.2 Materials and Methods

The study was carried in two parts:

(a) within-cultivar variation based on DNA, extracted from individual plants, and

(b) between cultivars (landrace and PPB) variation based on DNA, bulked from 10 individuals.

It was not technically possible to compare all the genotypes at the same time in a single gel, so the second experiment (b) was separately designed and included additional PPB bulk populations, to those in experiment (a). The genetic materials in these experiments are shown in Tables 8.1 and 8.2.

Twelve accessions (landraces, varieties and PPB bulks), represented by 108 individual plant DNA extracts, were included in the within-cultivar variation study (Table 8.1). Twenty accessions (landraces, varieties and bulks) were included in the experiment

(b) using bulk DNA of (Table 8.2). All the landraces included in this study were morphologically distinct FUDs in farmers' fields.

The methods used for extraction of DNA for both experiments and the methods for molecular analyses are described in Section 3.3 of Chapter III.

Table 8.1: Rice varieties of different origin assessed for within-cultivar variation by the use of 25 microsatellite (SSR) markers.

Variety	Lines / origin	DNA sample	Remarks
Jetho Budho (207)†	Landrace, Kaski	9 individuals	Aromatic
Kathe Gurdi (029)	Landrace, Kaski	10 individuals	Common and widely grown
Rato Anadi (094)	Landrace, Kaski	9 individuals	Bold grained
Basmati (320)	Landrace, Bara	6 individuals	Aromatic rice
Nakhisaro (308)	Landrace, Bara	8 individuals	Early rice
Laltenger (104)	Landrace, Bara	10 individuals	Late maturing
Machhapuchhre-3	PPB variety, Nepal	10 individuals	Chhommrong x Fuji 102
Machhapuchhre-9	PPB variety, Nepal	10 individuals	Chhommrong x Fuji 102
Chhommrong	Pure line, Nepal	10 individuals	Pure line selection
ET	PPB bulk, Nepal	8 individuals	Kalinga III x IR64 (early tall)
Ashoka 200F	PPB variety, India	9 individuals	Kalinga III x IR64, selection, India
Kalinga III	Pure line, India	9 individuals	Pure line selection, India

† = Number in parenthesis indicates the accession number of landrace

Twenty five microsatellite (SSR) markers distributed in 12 chromosomes of rice were used for experiment (a) within-cultivar variation and 36 markers for experiment (b) between-cultivar variation which have been found polymorphic (Chapter VI). These SSRs are listed in Appendix 3.3 and 8.1. The bands (alleles) detected in amplification at each of SSR loci were scored based on their relative mobility as present (1), absent (0), or heterozygous (0.5) for each of the two alleles observed per locus. The frequency of bands was calculated for each polymorphic locus for each variety (Appendix 8.1) and Shannon-Weaver diversity index (H') was used to quantify the within-cultivar diversity with multiple individuals per accession (Puccher *et al.*, 1996).

Table 8.2: Rice varieties assessed for between-cultivar variation by the use of 36 microsatellite (SSR) markers.

Variety	Lines / origin	DNA sample	Remarks
Seto Marshi	Landrace, Jumla	Bulk of 10	Cold tolerant, widely grown
Kalo Marshi	Landrace, Jumla	Bulk of 10	Cold tolerant, widely grown
Jetho Budho (207)†	Landrace, Kaski	Bulk of 10	Aromatic
Kathe Gurdi (029)	Landrace, Kaski	Bulk of 10	Common and widely grown
Rato Anadi (094)	Landrace, Kaski	Bulk of 10	Bold grained
Basmati (320)	Landrace, Bara	Bulk of 10	Aromatic rice
Nakhisaro (308)	Landrace, Bara	Bulk of 10	Early rice
Laltenger (104)	Landrace, Bara	Bulk of 10	Late maturing
Machhapuchhre-3	PPB variety, Nepal	Bulk of 10	Chhommrong x Fuji 102
Machhapuchhre-9	PPB variety, Nepal	Bulk of 10	Chhommrong x Fuji 102
Chhommrong	Pure line, Nepal	Bulk of 10	High altitude variety
ET	PPB bulk, Nepal	Bulk of 10	Kalinga III x IR64 (early tall)
MD	PPB bulk, Nepal	Bulk of 10	Kalinga III x IR64 (medium dwarf)
MT1	PPB bulk, Nepal	Bulk of 10	Kalinga III x IR64 (medium tall 1)
MT2	PPB bulk, Nepal	Bulk of 10	Kalinga III x IR64 (medium tall 2)
MT3	PPB bulk, Nepal	Bulk of 10	Kalinga III x IR64 (medium tall 3)
Ashoka 200F	PPB variety, India	Bulk of 10	Kalinga III x IR64 (selection, India)
Kalinga III	Pure line variety, India	Bulk of 10	Upland variety
IR64	IRRI variety	Bulk of 10	Irrigated
IR36	IRRI variety	Bulk of 10	Irrigated

† = Number in parenthesis indicates the accession number of landrace

Besides, other genetic variation parameters: total number of alleles, alleles per locus, percentage of polymorphic loci (PPL), percentage of polymorphic alleles (PPA), number of heterozygotes and the maximum number of alleles were also calculated and measured the dispersions of individuals of each variety. Number of individuals under test varied from 6 to 10 (Table 8.1).

The individual allelic data (108) were subjected to a principal component analysis to determine relationships among and between the varieties. A principal component analysis was also performed for experiment (b).

8.3 Results

8.3.1 Experiment (a)

8.3.1.1 Within-cultivar variation (intra-varietal heterogeneity):

Although the SSR variability was tested in 10 individual plants of each variety, the presumed off-type individuals in Basmati and the individuals in other varieties with no amplification for many markers were excluded in the analysis. The number of individual plants analysed per variety was shown in Table 8.1. Out of 25 markers tested, 22 detected the within-cultivar variation in 12 varieties (Figure 8.1, and Table 8.3). RM229, RM234 and RM257 were found non-informative for detecting within-cultivar variation in the varieties in the present study, but were observed polymorphic between the varieties. All varieties showed within-cultivar heterogeneity at a minimum of 1 and a maximum of 12 markers out of the 22 polymorphic markers with 4-48% of PPL, 8-66% of PPA and 0.15-2.64 of H'. The maximum values for the diversity parameters: alleles per locus, PPL, PPA, and Shannon-Weaver diversity index (H') was found in Rato Anadi (3) and early tall (11). Four of the six landraces and the PPB bulk (early tall) were more diverse than the PPB varieties, and their parental pure line varieties.

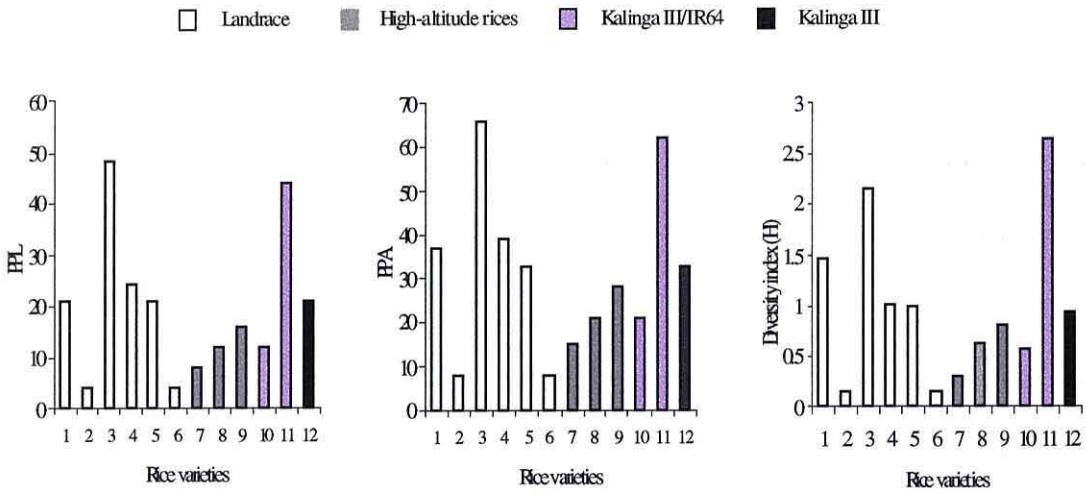


Figure 8.1: Within cultivar allelic variation among 12 rice varieties based on the variant individuals produced by 25 microsatellite (SSR) markers. (1. Jetho Budho; 2. Kathe Gurdi; 3. Rato Anadi; 4. Nakhisaro; 5. Basmati; 6. Laltenger; 7. Machhapuchhre 3; 8.Machhapuchhre 9; 9. Chhommrong; 10. Ashoka 200 F; 11.Early tall bulk; and 12. Kalinga III). (PPL = percentage of polymorphic loci; PPA = percentage of polymorphic alleles; H'= Shannon-Weaver Diversity Index).

The PPB bulk (early tall, 11) was as diverse as the most diverse landrace and had 44% polymorphic loci (PPL); 62% polymorphic alleles (PPA) and 2.64 diversity index (Figure 8.1, Table 8.3). Although the PPB varieties: Machhapuchhre 3 (7); Machhapuchhre 9 (8) and Ashoka 200F were relatively homogeneous with low diversity values (Figure 8.1) they were more diverse than the least diverse landraces. The supposedly pure line varieties Chhommrong (9) and Kalinga III (12) were more diverse than PPB varieties (Figure 8.1, Table 8.3).

Although rice is predominantly inbreeding, out crossing is known to occur. There was a maximum of 3 heterozygote individuals in Rato Anadi and Nakhisaro landraces for 1-2 loci, and early tall (ET) and Kalinga III each had one to two heterozygotes at one locus (Table 8.3).

Table 8.3: Allelic variation observed in 12 rice varieties in SSR analysis

Varieties	Total alleles	Allele per locus	Polymorphic allele	Maximum no. of alleles detected	Maximum heterozygote individuals
Jetho Budho	31	1.2	11	3	-
Kathe Gurdi	26	1.0	2	2	-
Rato Anadi	38	1.5	25	3	3 (2Ls)§
Nakhisaro	31	1.2	12	2	3 (1L)
Basmati	30	1.2	10	2	-
Laltenger	26	1.0	2	2	-
Machhapuchhre 3	27	1.1	4	2	-
Machhapuchhre 9	28	1.1	6	2	-
Chhommrong	29	1.2	8	2	-
Ashoka 200F	28	1.1	6	2	-
Early tall	37	1.5	23	3	2 (1L)
Kalinga III	30	1.2	10	2	1 (1L)
Average		1.2			-

§ = Loci

8.3.1.2 Microsatellite (SSR) polymorphism

Six of the 25 SSR detected a single banding pattern with no within-cultivar variation in all individuals of the three landraces from Kaski, but 5 of them produced multiple bands for remaining 9 varieties (Appendix 8.1). Likewise 14 SSRs detected a single banding pattern in three landraces from Bara. A different set of seven markers detected single banding pattern with no heterogeneity within accessions of the remaining 6 rice varieties (2 high-altitude rice varieties and 2 PPB varieties and 2 cross parent varieties) under study.

The number of alleles detected varied among the SSR markers and among accessions (Appendix 8.1). The maximum number 3 of alleles were detected in Jetho Budho for RM226; Rato Anadi for RM17 and early tall for RM21 (Plate 1). However, across 120 individuals a maximum of six alleles was resolved with RM206. RM48, RM213 and RM226 were the most informative and detected within-cultivar variation in maximum of 42% of the rice varieties under study. Moreover, RM213 and RM48

(chromosome 2) detected inter- and intra-variation among the individuals of PPB varieties for high-altitude each with 2 alleles but RM213 was non-informative for the landraces (Plate 8.2).

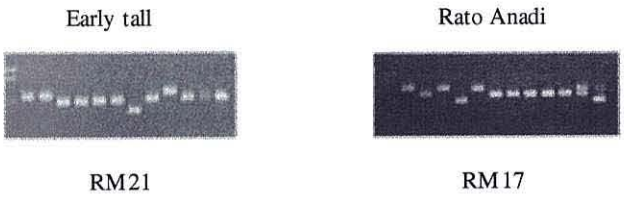


Plate 8.1: Microsatellite (SSR) markers and the rice varieties observed with a maximum of three alleles in individual DNA analysis.

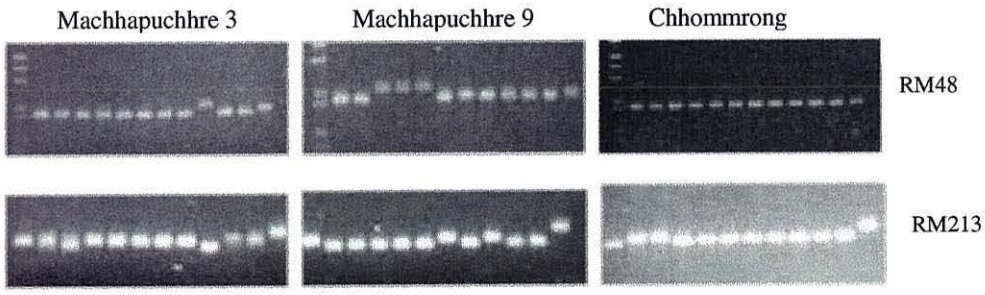


Plate 8.2: Microsatellite (SSR) markers with polymorphism in high altitude rice varieties (Chhommrong was monomorphic for RM48).

8.3.1.3 Genetic relationships by principal component analysis

The principal component analysis, using data from individual plants, explained 53% of the total variation over the first two axes and grouped the individual plants of the accessions into distinct clusters. The individuals of a named variety clustered closely showing relatively little within-cultivar variation compared to between-cultivar variation (Figure 8.2). For example, all the high altitude rice varieties of Nepal (individuals of Machhapuchhre 3, Machhapuchhre 9 and Chhommrong) grouped together. While individuals of Kalinga III, Ashoka 200F, and early tall (ET) (materials from cross

between Kalinag III and IR64) clustered together into a separate group. The individuals of landrace accessions each formed distinct clusters with somewhat varying amounts of within-cultivar variation. Individuals of Laltenger (LAL) and Basmati (BAS) from Bara and Kathe Gurdi (KG) from Kaski showed little within variation (Figure 8.2). However, landraces Rato Anadi (RA), Jetho Budho (JB) from Kaski; Nakhisaro (NS) from Bara and PPB bulk early tall (ET) varieties showed relatively high within-cultivar (accession) variation and the individuals were scattered within the cluster. Among the high-altitude varieties, Chhommrong (CHH) and Machhapuchhre 9 (M9) were more diverse than Machhapuchhre 3 (M3). In general, the variability in the PCA (Figure 8.2) agrees with the diversity measures shown in Figure 8.1.

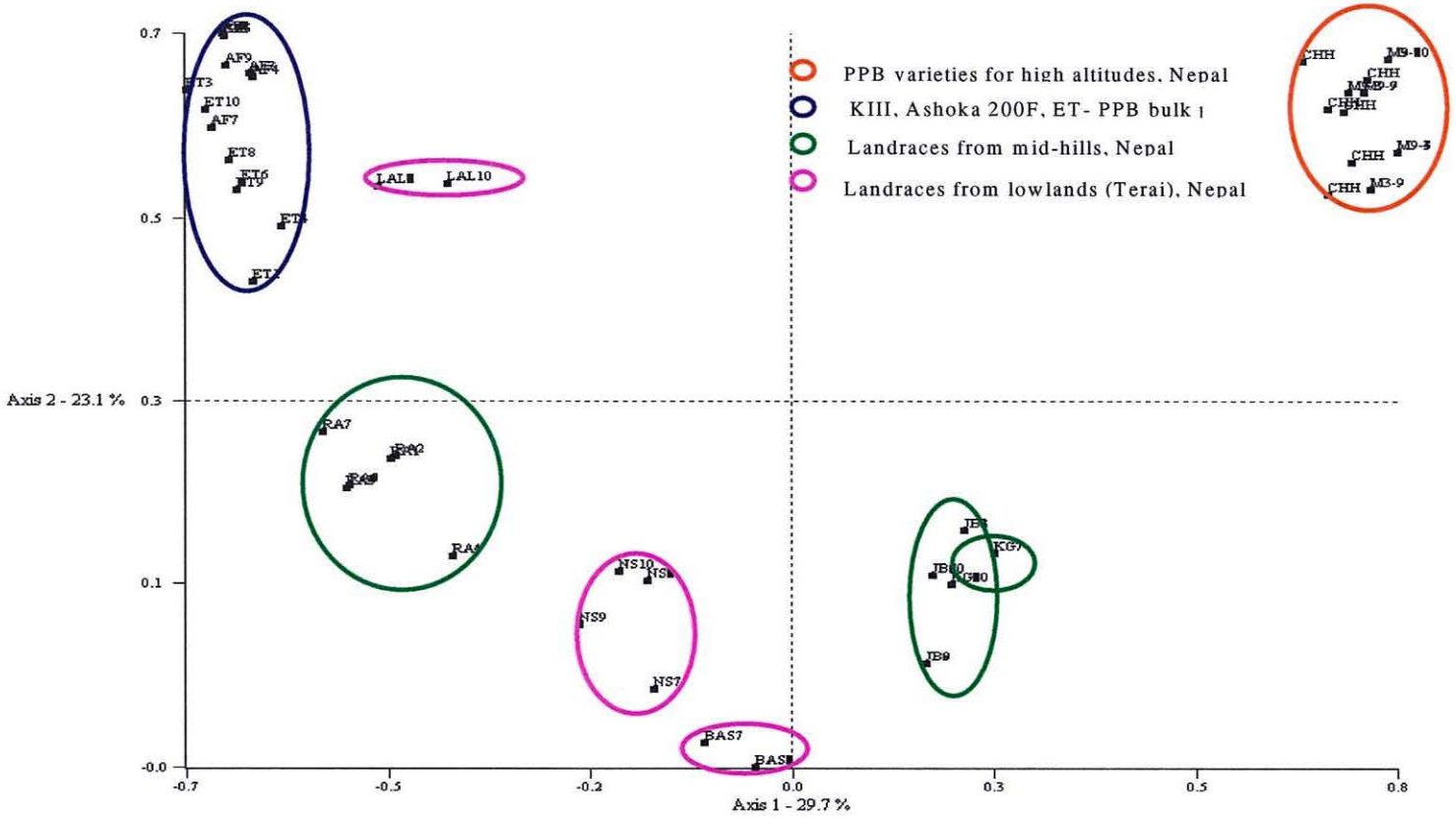


Figure 8.2: Scatter plot of 108 individuals of 12 varieties based PCA analysis on allelic data generated by 23 microsatellites (SSR)

8.3.2 Experiment (b)

8.3.2.1 Between-cultivar variation on bulk DNA samples

The 34 polymorphic microsatellite (SSR) markers (two were monomorphic) detected 109 alleles among the 20 rice accessions that represented 5 cultivar groups: checks (IRRI-bred, pure-line varieties for irrigated environments and upland variety); PPB bulk population and high-altitude PPB varieties and one of their parents; and, Nepalese landraces. The alleles per locus significantly varied among the 20 varieties ($F = 5.67$; $P<0.01$). The diversity values for these cultivar groups under study varied considerably and landraces from the Kaski and Bara ecosites and the PPB bulk populations were the most diverse (Table 8.4). The check varieties, consisting of two irrigated and one upland variety were, not unexpectedly diverse. Jumla landraces were the least diverse and were homogeneous for all of the markers tested. The next least diverse group were also from high altitudes (Chhommrong and two high-altitude PPB varieties).

Table 8.4: Comparative allelic diversity among different cultivar groups of rice consisting 20 different accessions.

Diversity parameters	Check varieties (2)	PPB bulk pop^n (5)	PPB varieties (3)	Landraces			Total accessions (20)
				Jumla (2)	Kaski (3)	Bara (3)	
Rice varieties	IR36 IR64 KIII	ET MD MT1 MT2 MT A200F	M3 M9 CHH	SM KM	JB KG RA	NS BAS LAL	20 varieties
No. of markers	34	34	34	34	34	34	34
Total alleles	52	57	36	34	57	60	109
Alleles per locus	1.5	1.7	1.1	1.0	1.7	1.8	3.2
Polymorphic markers	17	19	2	0	23	22	34
Alleles per polymorphic marker	2.1	2.2	2.0	0	2.0	2.2	3.2
PPA†	67	74	11	0	81	80	100
PPL‡	50	56	6	0	68	65	100

PPA† = Percentage of polymorphic alleles; PPL‡ = Percentage of polymorphic loci. (For the rice varieties see table 8.3 above).

8.3.2.2 Genetic Relationships analysed by principal component analysis (between cultivar diversity)

The principal component analysis on 20 rice varieties, showed similar groupings and relationships among the varieties as with the individuals of the 12 varieties. The first two axes explained 64% of total variation (Figure 8.3). The Nepalese landraces including three high altitude varieties Machhapuchhre 3, Machhapuchhre 9 and Chhommrong were clearly separated from the check varieties, and the PPB variety (Ashoka 200F) and bulks of Kalinga III x IR64 by the second axis with the exceptions of two landrace accessions (RA and LAL). It indicated the diverse origins of the rice materials under the study.

As expected, the landraces from Bara and Kaski displayed a higher inter-cultivar diversity compared with high-altitude materials. Landraces from Bara and Kaski were scattered far apart in the plot. However, Jetho Budho and Kathe Gurdi showed some over lapping. Rato Anadi (RA), an upland, late maturing variety with glutinous rice from Kaski and Laltenger (LAL), an irrigated, late maturing variety from Bara, however, were observed close to the check varieties. The scatter of the varieties shows that there is inter-cultivar variation among these varieties.

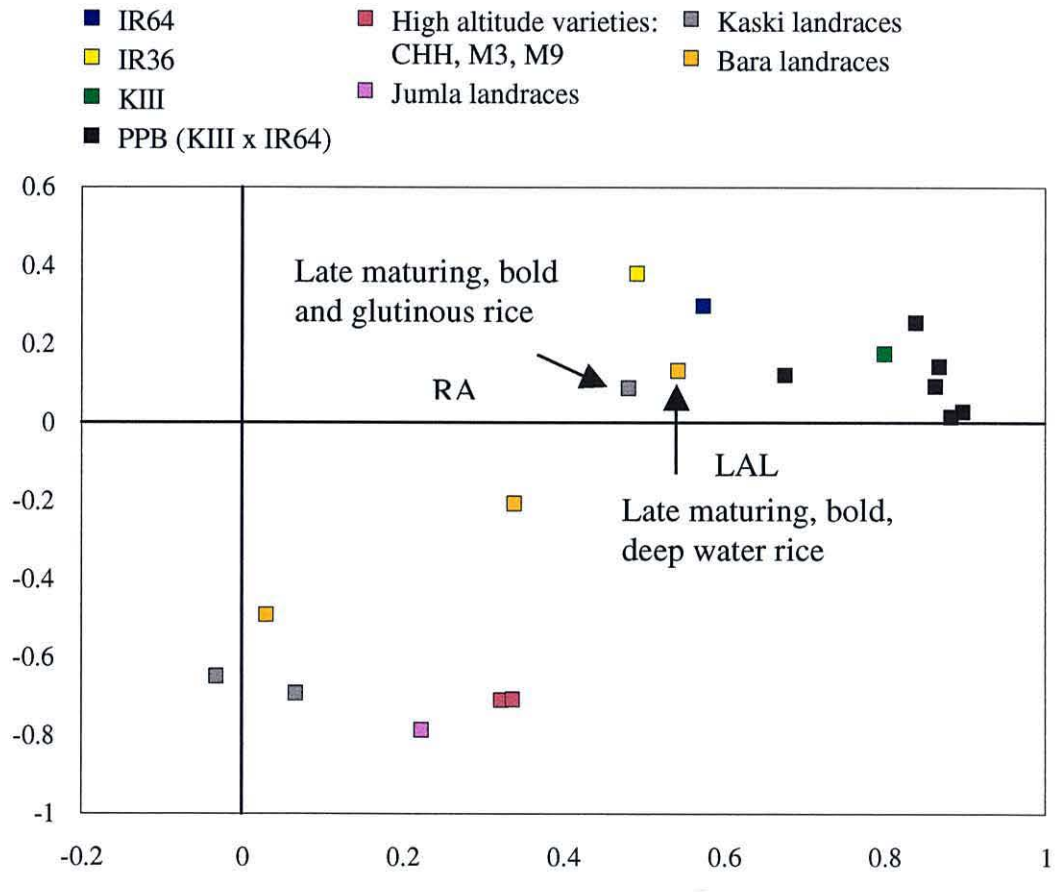


Figure 8.3: Scatter plot of 20 rice varieties based on 109 alleles generated by 34 microsatellite markers.

8.4 Discussion

Various molecular markers have been used to examine diversity in rice at the inter- and intra-varietal and species levels (McCouch *et al.*, 1988; Wang and Tanksley 1989; Yang *et al.*, 1994; Virk *et al.* 1995; Zhu, 1996 and Qian *et al.*, 2001)). These studies have shown that number of alleles per locus reduces from landraces to cultivars, *indica* to *japonica*, and wild to cultivated rice and no single technique is more effective than others in discriminating among the studied groups of rice. Olufowote *et al.* (1997) in a study on within-cultivar variation in rice observed the importance of within-cultivar variation in landraces and found the SSR as an efficient assay for detection of within-cultivar variation in rice.

The rice materials examined here represented diverse varieties: landraces from different agro-ecosystems ranging from lowland to high-altitudes, pure-line varieties and varieties from PPB from both India and Nepal. All the SSR markers that were polymorphic in analyses in Chapter VI were chosen for this inter- and intra-varietal heterogeneity and represented the whole genome of rice. The results of this study indicate that the SSRs provided information on the structure of a landrace population and supported the genetic relationships and differences among cultivars of different groups.

The landraces were the most heterogeneous group in the study and had the high levels of polymorphism with the highest average allelic variation for both within- and between-cultivar diversity. However, the landraces from Jumla alone display little diversity. Yang *et al.*, (1994) and Olufowote, *et al.*, (1997) indicated that landraces had a higher number of alleles than modern cultivars. However, this was observed in a collection of landraces and modern cultivars that represented a wide array of rice

germplasm of the world including South-East Asia. The lower number of alleles per locus detected in the present study was in a smaller set of rice landraces from only three ecozones of Nepal. Diversity depends upon the number, choice and discriminating ability of markers, the sample size, and on how diverse is the origin of the samples.

Can PPB effectively maintain within-cultivar diversity? The PPB bulk ET (early tall) from the Kalinga III / IR64 cross was as diverse as the most diverse landrace variety Rato Andi (Figure 8.1 and Table 8.3). The breeding history of ET was such (bulk population breeding with only moderate selection pressure) that relatively high within-cultivar diversity is not unexpected. When selection pressure in the same cross (Kalinga III x IR64) was more intense, as was the case for Ashoka 200F, – even though bulk population breeding was still the selection method – then within-cultivar diversity declined dramatically (Figure 8.2) and was less than that of the four of the six landraces and one of its parents, Kalinga III. When PPB was applied in another situation, high altitude rice in Nepal the PPB varieties Machhapuchhre 3 and 9 (M3 and M9) were less variable than Chhommrong. However, the PPB had generated more diversity for this high altitude site than was found among the landraces at high-altitude Jumla.

However, the genetic variation between bulks produced from a single cross using modified bulk population breeding (Witcombe *et al.*, 2001) was almost as high, when assessed by SSR markers, as the variability found among the landraces from Kaski and Bara. More over, the Kaski and Bara landraces were adapted to a much greater range of physical and socio-economic environments than the six PPB bulks, five of which were adapted to the same upland environment.

8.5 Conclusions

- Landraces represented a genetically diverse and heterogeneous group and had multiple alleles at many of the SSR loci. Landraces from Kaski and Bara had the largest within- and between- cultivar variations than others.
- A PPB bulk to which a low selection pressure had been applied conserved within-cultivar diversity better than conventional pure-line breeding, and also preserved as much within-cultivar diversity as found in many landraces.
- A PPB programme for high-altitude rice conserved more genetic variation than was found among landraces in the high-altitude Jumla ecosite.
- PPB bulks derived from a single cross can contain almost as much diversity as, the variability from a single ecosite, such as Kaski or Bara where many diverse landraces are found.

CHAPTER IX

GENERAL DISCUSSION

Knowledge of the traits and information on genetic variation in germplasm and its relationships is crucial for deciding conservation strategies. Agro-morphological characteristics and the PCR based markers have provided valuable information about the genetic diversity of Nepalese rice landraces. In this thesis, agro-morphological characteristics and molecular markers were employed to estimate genetic variation and the relationships among the rice landraces from three contrasting agro-ecosystems of the country (high-hill, mid-hill and lowland). The rice materials included in the study represented the major landraces under cultivation in these ecosites. Key findings on genetic diversity of rice landraces that were drawn from the present study and their relevance and implications in relation to the management of rice biodiversity are discussed in this concluding chapter. This is the first use of the molecular markers (SSRs) in characterizing and measuring the diversity of rice landraces of Nepal that was carried out in newly established molecular marker laboratory in Nepal.

The research began with a recognition of the importance of on-farm diversity of rice landraces in the three study ecosites (Chapter IV). The base-line survey showed that farmers in these sites have adopted a broad range of varieties adapted to diverse agro-ecosystems and production systems, and which satisfy various culinary, nutritional, social and ritual needs of the farming households. Across the ecosites, the physically and socio-economically favourable environments (favourable rice domains and resource-rich farming households) conserved the greatest diversity of rice. The marginal environments (stress-prone domains) had fewer landraces. Farmers use a set of traits and give names to

characterize and describe these units of crop diversity. This diversity is therefore reflected in the diverse names of landraces (FUDs) that farmers assign to their varieties. Landrace names are, therefore, the initial indicator of genetic diversity on-farm. Rao *et al.* (2002) in evaluating the rice landraces of Laos also pointed out the significance and utility of traditional names for use in crop improvement.

Several approaches have been used to characterize the extent of genetic diversity of rice including landraces (Second, 1982; Olufowote *et al.*, 1997; Sebastian *et al.*, 1998; Li and Rutger, 2000; Virk *et al.*, 2000). The results of the present study on the agro-morphological and molecular marker have revealed that there is an enormous amount of genetic variation among the landrace accessions from three diverse agro-ecological regions of Nepal but the extent of variation was much less in the high-hill ecosite.

The multivariate analyses (hierarchical clustering, principal component and discriminant function) of agro-morphological traits showed the taxonomic (morphological) relationships among the landraces within each ecosite. Such techniques have been widely used in divergence studies in a range of crop species including rice to reveal complex genetic relationships (Bhatta, 1970; Tolbert *et al.*, 1979; Sinha *et al.*, 1991; Newbury and Ford-Lloyd, 1997; and Alemayechu and Becker, 2002). The principal component analysis showed that quantitative traits were the most informative of the agro-morphological traits and explained the highest variation for Kaski and Bara landraces. However, in Jumla (the high-hill ecosite) qualiative traits were important. Hierarchical clustering and discriminant procedures were complementary. A clear discrimination of the rice accessions by FUDs, agro-morphological clusters (cluster analysis) and agro-ecological domains was revealed using the combined qualitative and

quantitative traits in discriminant function analysis and better confirmed the clustering pattern than the PCAs.

The results presented in Chapter V indicate that the diversity patterns were determined by the growing environments within each ecosite. In Bara (the lowland ecosite), the diversity was largely determined by the growth duration, plant height, grain characteristics and the adaptive traits of landraces while in Kaski (the mid-hill ecosite), it was for the leaf and grain characteristics. In contrast, in Jumla (the high-hill ecosite), the landraces exhibited differences for only a few qualitative traits. Colour of the stigma, colour of the sterile lemma and the presence or absence of awning were the important traits in describing the diversity in this ecosite. The groupings that landraces made in clustering analysis agreed with the ethno-botanical classification that farmers traditionally employ to identify and describe their rice landraces (Kaski and Bara). The exceptions were accessions of quality rice: Jetho Budho, Jhinuwa, Ramani from Kaski, and Basmati and Sotawa from Bara that showed intra-landrace variation for morphological traits and clustered in adjacent clusters.

The relative importance of qualitative and quantitative traits in estimating the extent and geographic distribution of the genetic variation varied across the ecosites. In the high-hills, the genetically inherited qualitative traits were more important. In the landraces from the mid-hills and lowlands, the quantitative traits were more important for the discrimination of rice landraces. The study also found that the rice diversity decreased from high-hill to mid-hill and lowland. The smaller diversity in Jumla (the high-hill ecosite) is possibly restricted due to natural selection for extreme environment with chilling temperature (Sthapit, 1994). There was a marked predominant cultivation of

Marshi varieties (bold grain), classified as Rato Marshi, Seto Marshi and Kalo Marshi by the farmers. In contrast, diverse growing environments, a wide range of altitudes in the mid-hill ecosite and homogenous but favourable environments in the lowland ecosite and both the natural and conscious selection by the farmers for diverse values have helped the maintenance of high diversity in these two ecosites.

The molecular marker study in the present research showed that SSR markers were useful and efficient and the level of polymorphism was high enough for characterizing and estimating the extent and distribution of genetic variation in selected rice landraces of Nepal. It is evident from the results in Chapters VI and VIII that a wide range of genetic diversity, as revealed by SSR polymorphism, exists among these landraces with intra-and inter-cultivar variation. Landraces from Bara and Kaski were characterized by high estimates for the number of alleles per locus, gene diversity and percentage of polymorphic loci and provided a genetic basis to the high morphological diversity observed in these landraces. The level of diversity in Kaski and Bara was almost the same and both had a considerable level of heterogeneity within and between landraces. Landraces in Jumla were essentially monomorphic and genetically homogenous for all the SSRs studied. The molecular marker analysis revealed that most landraces with same FUDs had similar microsatellite allelic composition. On the contrary, a different observation was encountered in the molecular analysis of traditional rice varieties of Philippines (Sebastian *et al.*, 1998) and in pearl millet (Busso *et al.*, 2000). Samples with same names had different allelic compositions but the samples with different names collected from different farmers had similar allelic composition.

Rice varieties developed by participatory plant breeding were almost as diverse as landraces in the present molecular study, indicating that the PPB approach can be useful in broadening the genetic base.

The diversity of rice landraces based on either agro-morphological traits and microsatellite (SSR) markers were compared. A very poor correlation was encountered within individual ecosites but a considerable correlation level was found when the comparison was on the combined data of three sites (Chapter VII). However, morphological differences between sites were confounded by environmental differences. The levels of genetic diversity in the landraces from the three ecosites were generally similar for both agro-morphological traits and SSR markers. The landraces from Jumla (the high-hill ecosite) which had low level of morphological variability, displayed the lowest genetic diversity at SSR loci as well and the diversity in Jumla was distinct from that in Kaski and Bara. In Kaski and Bara there is a large extent of diversity distributed in a range of environments and differentiated between the specialized and common rice landraces of the sites. So the use of greater number of SSR markers may provide still more information on genetic diversity in these two sites.

The results on agro-morphological and molecular variability in the present study generally established that a tremendous diversity has been maintained by the farmers as landraces. Amount of variation in landraces from high altitude is very low based on the present information. Both the marker systems are complementary to each other and have provided and assisted identifying the level of diversity. But with the increase in the number of markers survey, it will be possible to detect the diversity preciously and provide the greater number of information on diversity of these landraces.

REFERENCES

Adhikary NP, Palikhe MM, Devkota RN (1995) Status of rice, maize and wheat genetic resources in Nepal. pp 74-77 In Plant genetic resources: Nepalese perspectives, Proceedings of the national workshop on plant genetic resources conservation, use and management organized by NARC at Kathmandu, 28 Nov-1 Dec IPGRI/ NARC

Agriculture Perspective Plan (1995) A main document prepared for National Planning Commission, HMG of Nepal and Asian Development Bank. Agriculture Projects Service Centre, Kathmandu and John Mellor Associates, Inc. Washington, D.C.

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

Akagi H, Yokozeki Y, Inagaki, A, Fijimura, T (1997) Highly polymorphic microsatellites of rice consists of AT repeats, and a classification of closely related cultivars with these microsatellite loci. Theor Appl Genet 94: 61-67

Akkaya MS, Bhagawat AA, Cregon PB (1992) Length polymorphism of simple sequence repeats in soybean. Genetics 132: 1131-1139

Akkaya MS, Bhagwat AA, Cregan PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci 35: 1439-1445

Alemayechu N, Becker H (2002) Genotypic diversity and patterns of variation in a germplasm material of Ethiopian mustard (*Brassica carinata* A. Braun). Genetic Resources and Crop Evolution 49: 573-582

Almgard G, Landegren U (1974) Isoenzymatic variation used for the identification of barley cultivars. Z Pflanzenzuchtung 72: 63-73

Aman RA (1997) A comparative assessment of molecular techniques employed in genetic diversity studies and their suitability in resource limited settings. pp 119-127 In WG Ayad, T Hodgkin, A Jaradat, and VR Rao (eds) Molecular genetic techniques for plant genetic resources, Report of IPGRI workshop, 9-11 October, 1995, Rome, Italy, International Plant Genetic Resources Institute, Rome, Italy

Amin AY (2002) Markers for quantitative traits analysis and bulk segregant analysis of salt tolerance in wheat (*Triticum aestivum* L.). A thesis submitted to the University of East Anglia for the degree of philosophae. UK

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181-186

Arumuganathan, K, Earle ED (1995) Nuclic DNA content of some important plant species. Plant Mol Biol Rep 9: 208-218

Autrique E, Nachit MM, Monnenveux P, Tanksley SD, Sorrells ME (1996) Genetic diversity in durum wheat based on RFLPs, morphological traits and coefficient of parentage. Crop Sci 36: 735-742

Ayala EJ, Kiger JA (1984) Modern Genetics. 2nd (edn) Renjamin-Cumings, Menlo Park, USA.

Ayana A, Bekele E (1999) Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. Genetic Resources and Crop Evolution 46: 273-284

Becker J, Heun M (1995) Barley microsatellites: Allele variation and mapping. Plant Mol Biol 27: 835-845

Becker J, Vos P, Kuiper M, Salamini F, Heum M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 249: 65-73

Beer SC, Goffreda J, Phillips TD, Murphy JP, Sorrells ME (1993) Assessment of genetic variation in *Avena sterilis* using morphological traits, isozymes, and RFLPs. Crop Sci 33: 1386-1393

Bell CJ, and Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19: 137-144

Bellon, MR (1991) The ethnoecology of maize variety management: a case study from Mexico. Human Ecology 19:86-96

Bellon MR, Pham JL, M.T. Jackson (1997) Genetic conservation: a role for rice farmers. Pp 263-290 In N Maxted, BV Ford-Lloyd, JG Hawkes (eds) Plant Genetic Conservation: An *in situ* approach. Chapmann and Hall London, UK

Beuningen LT, Busch RH (1997) Genetic diversity among North American spring wheat cultivars: III cluster analysis based on quantitative morphological traits. Crop Sci 37:981-988

Bhat KV (1996) Molecular genetic techniques for characterization of plant genetic resources. Pp 21-25 In Biochemical and molecular techniques for characterization and classification of plant genetic resources USAID-PGR project and ICAR, Pusa Campus, New Delhi, India

Bhatta GM (1970) Multivariate analysis approach to selection of parents for hybridization aiming at yield improvement in self pollinated crops. Aust J agric Res 21: 1-7

Blair MW, Hedetale V, McCouch SR (2002) Fluorescent-labelled microsatellite panels useful for detecting allelic diversity in cultivated rice (*Oryza sativa* L.). Theor Appl Genet, on line published 23 May, 2002

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Gene 32: 314-331

Braun P, Tanksley SD (1996) Characterization and genetic mapping of simple repeat sequences in tomato genome. Mol Gen Genet 250: 39-49

Brown AHD (1978) Isozymes, plant population genetic structure and genetic conservation. Theor Appl Genet 52:145-157

Brown AHD (1999) The genetic structure of crop landraces and the challenges to conserve them *in situ* on farm. pp 29-50 In SB Brush (ed) Genes in the Field: On Farm Conservation of Crop Diversity. Lewis Publishers, Boca Raton/IDRC, Ottawa/IPGRI, Rome, Italy

Brown AHD, Clegg MT (1983) Isoenzyme assessment of plant genetic resources. Isozymes: Current Topics in Biological and Medical research, Medical and other applications 11: 285-295

Brunel D (1994) A microsatellite marker in *Helianthus annus* L. Plant Mol Biol 24: 397-400

Brush SB (1991) A farmer-based approach to conserving crop germplasm. Economic Botany 45:153-165

Brush SB (1995) *In situ* conservation of crop landraces in centres of crop diversity. Crop Sci 35:346-354

Buell CR (2002) Obtaining the sequence of the rice genome and lessons learned along the way. Trends in Plant Science 7(12):

Busso CS, Devos KM, Ross G, Mortimore M, Adams WM, Ambrose MJ, Alldrick S, Gale MD (2000) Genetic diversity within and among landraces of pearl millet (*Pennisetum glaucum*) under farmer management in West Africa. Genetic Resources and Crop Evolution 100:1-8

Bustos AD, Casamova C, Soler C, Jouve N (1998) RAPD variation in wild populations of four species of the genus *Hordeum* (Poaceae). Theor Appl Genet 96: 101-111

Byerlee D, Moya P (1993) Impacts of international wheat breeding research in the developing world 1966-90. CIMMYT, Mexico, D.F

Caldo RA, Sebastian LS, Hernandez JE (1997) Diversity of Philippine improved rice varieties and their progenitors using morphological and molecular markers. Philippine Journal of Crop Science (Philippines) 22: 11-26

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE, Second G, McCouch SR, Tanksley SD (1994) Saturated molecular map of the rice genome based on an interspecific backcross populations. Genetics 138: 1251-1274

Centre Bureo of Statistics (CBS) (2000) Statistical information on Nepalese agriculture 1999/2000. HMG/MOA/ Agri-Business Promotion and Statistic Division, Kathmandu, Nepal

Chakraborty R, Kimmel M, Strivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. Proc Natl Acad Sci USA. 94: 1041-1046

Chambers R (1998) Institutions and practical change. Reversals. institutions and change. In R Chambers, A Paccy and IA Thrupp (eds). Farmer First 181-195 Intermediate Technology Publications, London

Chang TT (1976) The origin, evolution, cultivation, dissemination and diversification of Asian and African rices. Euphytica 25: 425-441

Chang TT (1985) Principle of genetic conservation. IOWA State journal of research, 59:325-348

Chang TT, Bardens EA (1965) The morphology and varietal characteristics of the rice plant. IRRI Tech Bull 4: 40pp

Chang, TT, Oka HI (1976) Genetic various ness in the climatic adaptation of rice cultivars. pp 87-111 In Proceedings of the symposium on climate and rice, IRRI, Los Banos, Philippines

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

Cheng HH, Crittenden LB (1994) Microsatellite markers for genetic mapping in the chickens. Poult Sci 73: 539-546

Cho YG, Blair MW, Panaud O, McCouch SR (1996) Clonning and mapping of variety specific rice genomic sequences: amplified fragment length polymorphisms (AFLP) from silver stained polyarylamide gels. Genome 39: 373-378

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

Cho YG, Kim YW, Kang MR, Seok SJ, Eun MY (1998) Construction of a molecular frame map using MG recombinant inbred populations and RFLP markers mapped in rice. Korean J Plant breeding

Cho YG, McCouch SR, Kuiper SR, Pot J, Kang MR, Groenen JTM, Eun MY (1997) Integrated map of AFLP, SSLP, and RFLP markers using a recombinant inbred populations of rice (*Oryza sativa* L.) Theor Appl Genet 97: 370-380

Chowdhury AK, Srinives P, Tongpamnak P, Saksoong P (2001) Genetic diversity based on morphology and RAPD analysis in vegetable soybean. Korean J Crop Sci 46: 112-120

Conklin HC (1957) Hanunoo agriculture: a report on an integral system of shifting cultivation in the Philippines. FAO Forestry Development Paper No. 12. FAO, Rome

Cox TS, Wood D (1999) The nature and role of crop biodiversity. pp 35-58 In D Wood and JM Lenne (eds) Agrobiodiversity: Characterisation, Utilization and Management, CABI Publishing, CAB International, Wallingford, Oxon, UK

Crossa J, Delacy LH, Taba S (1995) The use of multivariate methods in developing a core collection. In T Hodgkin, AHD Brown, Th JL van Hintum, EAV Morales (eds). Core Collections of Plant Genetic Resources John Wiley Chichester England

Cruzan MB (1998) Genetic markers in plant evolutionary ecology. Ecology. 79: 400-412. cultivation. Human Ecology 11: 16-34

Davierwala AP, Chowdari KV, Kumar S, Reddy APK, Ranjekar PK, Gupta VS (2000) Use of different marker systems to estimate genetic diversity of Indian elitr rice varieties. Genetica 108: 269-284

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Mirissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5264 microsatellites. Nature 380: 152-154

Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor TJ, Evans CA, DeAngelis MM, Levinson DM, Kruglyak L, Goodman N, Copelang NG, Jenkins NA, Hawkins TL, Stein L, Page DC, Lander ES (1996) A comprehensive genetic map of the mouse genome. Nature 380: 149-152

Ellis RP, McNicole JW, Baird L, Booth A, Lawerence P, Thomas B, Powell W (1997) The use of AFLP to examine genetic relatedness in barley. Mol Breed 3;359-369

Estoup A, Pressa P, Krrieg F, Vaiman D, Guyomard R (1993) (CT)-n and (GT)-n microsatellites: a new class of genetic markers for *Salmo triata* L.(brown trout). Heredity 71: 488-496

Eyzaguirre P, Iwanga M (1995) Farmers contribution to maintaining genetic diversity in crops, and its role within the total genetic resources systems. pp 9-18 Proceedings of a workshop on participatory plant breeding, 26-29 July, 1995 Wageningen, the Netherlands. IPGRI, Rome

Ezeaku, L, Gupta SC, Prabhakar VR (1999) Classification of sorghum germplasm accessions using multivariate methods. African crop Science Journal 7: 97-108

Fahima T, Sun GL, Beharev A, Krugman T, Beiles A, Nevo E (1999) RAPD polymorphism in wild emmer wheat populations, *Triticum diococcoides*, in Isarel. Theor Appl Genet 98: 434-447

Falconer DS (1989) Introduction to Quantitative Genetics. Third Edition, Longman Scientific and Technical, England 438pp

FAO (Food and Agriculture Organization of the United State) (1996) Global plan of action for the conservation and sustainable utilization of plant genetic resources for food and agriculture. Division of Plant Production and Protection, FAO, Rome Italy

Fehr WR (1987) Principles of Cultivar Development: Theory and Techniques. Mcgraw Hill.Inc., New York

Ford-Lloyd BV (1990) The conservation of horticultural plant genetic resources. Advances in Horticultural Sciences 4: 31-38

Ford-Lloyd BV, Jackson MT (1986) Plant Genetic Resources: An Introduction to Their Conservation and Use. Edward Arnold, London

Ford-Lloyd BV, Jackson MT, Newbury HJ (1997) Molecular markers and the management of genetic resources in seed genebanks: a case study of rice. pp 103-105 In JA Callow, BV Ford-Lloyd and HJ Newbury (eds). Biotechnology and Plant Genetic Resources: Conservation and Use. CAB INTERNATIONAL, Wallingford, UK

Ford-Lloyd BV, Newbury HJ (1994) New methods for detecting genetic variation in plant germplasm. Anadolu 4: 94-98

Francisco-Ortega J, Jackson MT, Catty JP, Ford-Lloyd BV (1992) Genetic diversity in the *Chamaecytisus proliferos* complex (Fabaceae: Genisteae) in the Canary Islands in relation to *in situ* conservation. Genetic Resources and Crop Evolution 39: 149-158

Franco J, Crossa J, Ribaut JM, Betran J, Warburton ML, Khairallah M (2001) A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. Theor Appl Genet 103: 944-952

Frankel OH, Brown AHD, Burdon JJ (1995) The conservation of plant biodiversity. Cambridge University Press, Cambridge, UK

Furman BJ, Qualset CO, Skovmand B, Heaton JH, Corke H, Wesenberg DM (1997) Characterisation and analysis of north American triticale genetic resources. Crop Sci 37: 1951-1959

Gao LZ, Ge S, Hong DY, Zhang JW, Luo QY, Tao GD, Xu ZF (1999) A study on the population genetic structure of *Oryza granulata* (Zolf. Et. Mor. Ex. Steud) Baill. from Yunnan and its *in situ* conservation significance. Science in China Series C0 42: 102-108

Gao LZ, Schaal BA, Zhang CH, Jia JZ, Dong YS (2002) Assessment of population genetic structure in common wild rice *Oryza rufipogon* Griff. using microsatellite and allozyme markers.http://link.Springer.de/link/servi.../paper/s00122-002-0917-1ch100.htm

Garland SH, Lewin L, Abedinia M, Henry R, Blakency A (1999) The use of microsatellite polymorphisms for identification of Australian breeding lines of rice (*Oryza sativa* L.). Euphytica 108: 53-63

Ge S, Oliveira GC, Schaal BA, Gao LZ, Hong DY (1999) RAPD variation within and between natural populations of the wild rice *Oryza rufipogon* from China and Brazil. Heredity 82: 638-644

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. Theor Appl. Genet 74: 21-30

Gomez KA, Gomez AA (1976) Statistical procedures for agricultural research with emphasis on rice. The International Rice Research Institute, Los Banos, Laguna, Philippines

Goodman MM (1972) Distance analysis in biology. Syst Zool 21:174-186

Goodwin ID, Aitkin EAB, Smith LW (1997) Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18: 1524-1528

Graner A, Ludwig WF, Melchinger AE (1994) Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. Crop Sci 34: 1199-1205

Guldagar P (1975) Ex situ conservation stands in the tropics. In L Roche (ed) Methodology of Conservation of Forest Genetic Resources, FAO, Rome

Gurung JB, Vaidya AK (1998) Cultural context of agricultural biodiversity and sustainable agriculture in the Nepalese hills and mountains. pp 55-60 In T Pratap and BR Sthapit (eds). Managing Agrobiodiversity: Farmers' changing perspectives and institutional responses in Hindu Kush Himalayan (HKH) Region. International Centre for Integrated Mountain Developments/International Plant Genetic Resources Institute, ICIMOD, Nepal

Hamada H, Kakunaga T (1982) Potential Z-DNA forming sequences are highly dispersed in the human genome. Nature 298: 396-398

Hamrick JL, Godt MJW, Murawski DA, Loveless MD (1991) Correlations between species traits and allozyme diversity: implications for conservation biology. In Donald Folk and Kent E. Holsinger (eds). Genetics and Conservation of Rare Plants Oxford University Press, Oxford, UK

Harlan JR (1975) Our vanishing genetic resources. Science 188: 618-621

Hodgkin T, Roviglioni R, De Vicente MC, Dudnik N (2001) Molecular methods in the conservation of plant genetic resources. Int Symp on Molecular Markers. Acta Hort 546: 107-118

Illig T (2002) High-throughput SNP genotyping on a MALDI-TOF ms system, using Abgene thermo-start DNA polymerase. Insight Abgene. 15: 8-9

Innan H, Terauchi R, Miyashita T (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146: 1441-1452

IPGRI (1993) Diversity for development: The strategy of the International Plant Genetic Resources Institute. International Plant Genetic Resources Instirute, Rome , Italy

IRRI (1993) IRRI rice Almanc 1993-1995. International Rice Research Institute, Manila, The Philippines

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

Ishii T, McCouch SR (2000) Microsatellite and microsynteny in the chloroplast genomes of *Oryza* and eight other gramineae species. Theor Appl Genet 100: 1257-1266

Jaccard P (1908) Nouvelle researches sur la distribution florale. Bull. Soc. Vaud Sci Nat 44: 223-270

Jackson MT (1997) Conservation of rice genetic resources: the role of the International Rice Genebank at IRRI. Plant Molecular Biology 35: 61-67

Jackson MT, Pham JL, Newbury HJ, Ford-Lloyd BV, Virk PS (1999) A core collection for rice – needs, opportunities and constraints. pp 18-27 In RC Johnson, T. Hodgkin (eds). Core collection for today and tomorrow IPGRI, Rome, Italy

Jarvis DI, Hodgkin T (1998) Strengthening the scientific basis of *in situ* conservation of agricultural biodiversity on-farm: option for data collecting and analysis. IPGRI, Rome, Italy

Jarvis DI, Myer L, Klemick H, Guarino L, Smale M, Brown AHD, Sadiki M, Sthapit B, Hodgkin T (2000) A training guide for *in situ* conservation on farm. Version 1. International Plant Genetic Resources Institute, Rome, Italy

Jeffreys AJ, Wilson V, Thein SL (1985) Nature (London) 314: 67-73

Jena KK, Khush GS, Kochert G (1992) RFLP analysis of rice (*Oryza sativa* L.) introgression lines. Theor Appl Genet 84: 608-616

Joshi A, Witcombe JR (1996) Farmer participatory crop improvement II. Participatory varietal selection, a case study in India. Experimental Agriculture 32: 461-477

Joshi KD, Rana RB, Subedi M, Kadayat KB, Sthapit BR (1997) Addressing diversity through farmer participatory variety testing and dissemination approach: a case study of Chaite rice in the Western hills of Nepal. pp 158-176 In L Sperling, ML Loevinsohn (eds) Using Diversity, Proceedings of Conference on Using Diversity and Maintaining Genetic Resources on Farm, New Delhi, June 1995, International Development Research Centre, New Delhi

Juliano AB, Naredo EB, Jackson MT (1998) Taxonomic status of *Oryza glumaepatula* Steud. I. Comparative morphological studies of New world diploids and Asian AA genome species. Genetic Resources and Crop Evolution 45: 197-203

Kantety RV, Zeng X, Bennentzen JL, Zehr B (1995) Assessment of genetic diversity in dent and popcorn (*Zea mays* L.) inbred lines using inter-simple sequence repeats (ISSR) amplification. Mol Breed 1: 365-373

Karihaloo, JL (1996) Origin and evolution of eggplant. pp 43-47 In Biochemical and Molecular techniques for characterization and classification of plant genetic resources, USAID-PGR project and ICAR, Pusa campus, New Delhi, India

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

Kato S, Kosaka h, Hara S (1928) On the affinity of rice varieties as shown by fertility of hybrid plants. Bull Sci Fac Agric Kyushu Univ Fukuoka, Japan 3: 132-147

Khush GS (1997) Origin, dispersal, cultivation and variation of rice. Plant Molecular Biology 35: 25-34

Kumar R, Singh DN, Prasad SC, Gangawar JS, Virk DS, Witcombe JR (2001) Participatory plant breeding in rice in Eastern India. pp 233-236 An exchange of experiences from South and South east Asia: proceedings of the international symposium on participatory plant breeding and plant genetic resource enhancement, Pokhara, Nepal, 1-5 May 2000. Cali, Colombia: Participatory Research and Gender Analysis Program, Coordination Office; International Center for Tropical agriculture

Kurata N, Nagamura Y, Yamanoto K, Harushima Y, Suc N, Wu J, Antonio BA, Shomura A, Shimizu T, Lin SY, Inouc T, Fukuda A, Shimano T, Kuboki Y, Toyama T, Miyamoto Y, Kirihaya T, Hayasaka K, Miyao A, Monna L, Zhong HS, Tamura Y, Wang ZX, Momma T, Umchara Y, Yano M, Sasaki T, Minobe Y (1994) A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genet.8: 365-372

Lambert DH (1985) Swamp rice farming: the indigenous Pahang Malay agricultural system. West view Press, Boulder

Lanaud C, Lebot V (1997) Molecular techniques for increased use of genetic resources. In WG Ayad, T Hodgkin, A Jaradat, VR Rao (eds). Molecular genetic techniques for plant genetic resources, Report of IPGRI workshop, 9-11 October, 1995, International Plant Genetic Resources Institute, Rome, Italy

Lando RP, Mak S (1994) Combodian farmers decision-making in choice of traditional rainfed lowland rice varieties. IRRI Research Paper Series 154, International Rice Research Institute, Manila, Philippines

Lapointe FJ, Legendre P (1992) Statistical significance of the matrix correlation coefficients for comparing phylogenetic trees. Systematic Biology 41: 378-384

Lee M (1995) DNA markers and plant breeding programs. Adv Agron 55:265-344

Li Z, Rutger JN (2000) Geographic distribution and multilocus organization of isozyme variation of rice (*Oryza sativa* L.) Theor Appl Genet 10: 379-387

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics 44: 397-401

Loutte D, Charrier A, Berthaud, J (1997) *In situ* conservation of Maize in Mexico: genetic diversity and maize seed management in a traditional community. Econ Bot 5(1):20-38

Love JM, Knight AM, Mealeer MA, Todd JA (1990) Towards construction of a high resolution map of the mouse genome using PCR-analyzed microsatellites. Nucl Acis Res 18 14: 4123-4130

Lu JI, Chang TT (1980) Rice in its temporal and spatial persepectives. pp 1-74 In BS.Luh (ed) Rice: Production and Utilization. AVI Publishing Co. Inc. Westport, CT

Luce C, Noyer JL, Tharreau D, Ahmadi N, Feyt H (2001) The use of microsatellite markers to examine the genetic resources of rice (*Oryza sativa* L.) adapted to European conditions. Proc Int Symp On Molecular markers Acta Hort. 546: 221-234

Mackill DJ, Zhang Z, Redona ED, Colowit PM (1996) Level of polymorphism and genetic mapping of AFLP markers in rice. Genome 39:969-977

Magurran A (1991) Ecological diversity and its Management. Princeton: Prnceton university Press, UK

Maheswaran M, Subudhi PK, Nandi S, Xu JC, Parco A, Yang DC, Huang N (1997) Polymorphism, distribution, and segregation of AFLP markers in double haploid rice populations. Theor Appl Genet 94: 39-45

Manly BFJ (1994) Multivariate statistical methods – A primer. Chapman and Hall/CRC, London

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Research. 27: 209-220

Martin C, Juliano A, Newbury HJ, Lu BR, Jackson MT, Ford-Lloyd BV (1997) The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. Genetic Resources and Crop Evolution 44: 175-183

Martinello GE, Leal NR, Amaral Junior AT, Pereira MG, Daher RF (2001) Comparison of morphological characteristics and RAPD for estimating genetic diversity in *Abelmoschus* spp. Intl. Symp. On Molecular Markers. Dore, Dosba and Baril. Acta Hort. 546: 101-104

Maughan PJ, Saghai Maroof MA, Bush GR, Huestis GM (1996) Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritence and near-isogenic line analysis. Theor Appl Genet 93:392-401

Maurya DM, Bottrall A, Farrington J (1988) Improved livelihood, genetic diversity and farmers' participation: a strategy for rice-breeding in rainfed areas of India. Experimental Agriculture 24:311-320

Maxted N, Ford-Lloyd BV, Hawkes JG (1997) Complementary conservation strategies. pp 15-19 In N Maxted, BV Ford-Lloyd and JG Hawkes (eds). Plant Genetic Conservation- The *In Situ* Approach. The University of Burmingham, Chapman & Hall, London, UK

Maxted N, Guarino L, Myer L, Chiwona EA (2002) Towards a methodology for on-farm conservation of plant genetic resources. Genetic Resources and Crop Evolution 49: 31-46

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

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76; 815-829

Moncada P, Martinez CP, Borreru J, Chatel M, Gauch Jr.H, Guimaraes E, Tohme J, McCouch SR. (2001) Quantitative trait loci for yield and yield components in an *Oryza sativa* x *Oryza rufipogon* BC_2F_2 population evaluated in an upland environment. Theor Appl Genet 102: 41-52

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J. 1: 175-182

Moringa T (1954) Classification of rice varieties on the basis of affinity. pp 1-14 In Reports for 5th Meeting of International Rice Commission a Working Party on Rice breeding, Ministry of Agriculture and Forestry, Tokyo

Morishima H, Oka HI (1960) The patterns of interspecific variations in the genus *Oryza*: Its quantitative representation by statistical methods. Evolution 14: 153-165

Murphy PJ, Witcombe JR (1981) Variation in Himalayan barley and the concept of centers of diversity. pp 26-36 In Barley Genetics IV. Proceedings of the Fourth International Barley Genetics Symposium, Edinburgh

Nakagahra M, Akihama T, Iwata N (1975) Genetic variation and geographic lines of esterase isozyme in native rice varieties. Jpn J genet 50: 373-382

Nandi HD (1936) The chromosome morphology, secondary associations and origin of cultivated rice. Journal of Genetics 33: 315-336

Nei, M (1972) Genetic distances between populations. American Naturalist 106: 283-292

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc. Natl. Acad.Sci. USA 70:3321-3323

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Newbury HJ, Ford-Lloyd BV (1993) The use of RAPD for assessing variation in plants. Plant growth Regulation, 12: 43-51

Newbury HJ, Ford-Lloyd BV (1997) Estimation of genetic diversity. pp 192-206 In N Maxted, BV Ford-Lloyd, JG Hawkes (eds). Plant Genetic Conservation the *In situ* Approach Chapman and Hall, London, UK

Ni J, Colowit PM, Mackill DJ (2002) Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci. 42: 601-607

Nielson G (1985) The use of isozymes as probes to identify and label plant varieties and cultivars. pp 1-32 In MC Ruttazi, JG Scandalios, GS Whitt (eds) Isozymes: Current topics in biological and medical research 12 Alan R Liss, New York, NY

Oka HI (1958) Inter-varietal variation and classification of cultivated rice. Indian J Genet Plant Breed 18: 79-89

Oka HI (1988) Origin of cultivated rice. Japan Scientific Societies Press, Tokyo pp 25

Olufowote JO, Xu Y, Chen X, Park WD, Beachell HM, Dilday RH, Goto M, McCouch SR (1997) Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLp markers. Genome 40: 370-378

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

Parsons BJ, Newbury JH, Jackson MT, Ford-Lloyd BV (1997) Contrasting genetic diversity relationships are revealed in rice (*Oryza sativa* L.) using different marker types. Mol Breed.3: 115-125

Paudel CL, Tiwari PR, Neupane, JD, Devkota DP, (1998) Strengthening the Scientific Basis for *In situ* Conservation of Agrobiodiversity: Findings of *In situ* selection exercise in Jumla. NP Working paper No. 3/98 NARC/ LIBIRD/IPGRI

Paudel KP (1999) Genetic diversity of Pummelo in Nepal and improvement of propagation methods. A thesis submitted to the Unversity of South Hampton for the degree of doctrine in philosophae, UK

Perlas N, Vellve R (1997) *Oryza* Nirvana? A SEARICE Publication, Manila, Philippines

Pham JL, Morin SR, Sebastian LS, Abrigo GA, Calibo MA, Quilloy SM, Hipolitto L, Jackson MT (2001) Rice, farmers and genebanks: a case study in the Cagayan valley, Philippines. Pp 149-160 In JMM.Engels, VR Rao, AHD Brown, Jackson MT (eds). Managing Plant Genetic Diversity IPGRI, CAB International. Wallingford, UK

Plaschke J, Ganal MW, Roder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91: 1001-1007

Poehlman JM (1979) Breeding Field crops. 2nd edn AVI Publishing Co, New York

Porceddu E, Ceoloni C, Lafiandra D, Tanzarella OA, Scarascia Mugnozza GT (1988) Genetic resources and plant breeding: problems and prospects. In TE Miller, RMD Koebner, (eds) Proc Int Wheat Genetic Symp, 7th, Cambridge, England, 13-19 July 1988. Inst Plant Sci Res, Cambridge Lab., Cambridge, England

Pot J, Kuiper M, Vos P, Bastiaana E, Vanden Boschl N, Weller J (1996) A high density AFLPTM map of *Arabidoposis thaliana*. Plant genome. IV Abstracts. Poster 76 (http;//probe.nalusda.gov.8000/otherdocs/pg/pg4/PG-IV.html)

Poulsen GB, Kahl G, Weising K (1993) Abundance and polymorphism of simple repetitive DNA sequences in *Brassica napus* L. Theor Appl Genet 85: 994-1000

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S Rafalski, A (1996) The comparison of RFLP, RAPD, AFLP and SSR (Microsatellite) markers for germplasm analysis. Mol. Breeding 2: 225-238

Prabhu RR, Webb D, Jessen H, Luk S, Smith S, Greshoff PM (1997) Genetic relatedness among soybean genotypes using DNA amplification fingerprinting (DAF), RFLP and pedigree. Crop Sci 37: 1590-1595

Prance GT (1997) The conservation of botanical diversity. pp 3-15 In N Maxted, BV Ford-Lloyd, JG Hawkes (eds) Plant genetic conservation: the *in situ* approach. Chapmann and Hall, London, UK

Price AH, Steele KA, Moore BJ, Barraclough PB, Clark, LJ (2000) A combined RFLP and AFLP linkage map of upland rice (*Oryza sativa* L.) used to identify QTLs for root-penetration ability. Theor Appl Genet 100: 49-56

Puccher DI, Ibanez MA, Ranzo MA (1996) Classification and diversity values of seventeen cultivars of *Eragrostis curvula*. Seed Sci and Technol 24: 139-149

Qiagen, (1999) DNeasy plant mini handbook for DNA isolation from plant tissue. 1011408, 03/99 West Sussex, RH10 2AX

Qian W, Ge S, Hong DY (2001) Genetic variation within and among populations of a wild rice *Oryza granulata* from China detected by RAPD and ISSR markers. Theor Appl Genet 102: 440-449

Qualset, CO, Damania AB, Zanatta ACA, Brush SB (1997) Locally based crop plant conservation. pp 160-175 In N Maxted, BV Ford-Lloyd, JG Hawkes (eds) Plant Genetic Conservation- The In Situ Approach The University of Burmingham, Chapman and Hall, London, UK

Rafalski JA (2002) Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci 162: 329-333

Ramkrishna W, Lagu MD, Gupta VS, Ranjekar PK (1994) DNA fingerprinting in rice using oligonucleotide probes specific for simple repetitive DNA sequences. Theor Appl Genet 88: 402-406

Rao SP, Bounphanousay C, Schiller JM, Jackson MT (2002) Collection, classification, and conservation of cultivated and wild rices of the Lao PDR. Genetic Resources and Crop Evolution 49: 75-81

Rana RB, Paude CL, Tiwari PR, Gauchan D, Subedi A, Sthapit BR, Upadhyaya MP, Jarvis DI (2000a) *In situ* crop conservation: findings of agro-ecological, crop diversity and socio-economic baseline survey of Talium eco-site, Jumla, Nepal. NP working paper No. 3/2000. NARC/LIBIRD/IPGRI. Rome, Italy

Rana RB, Rijal DK, Gauchan D, Sthapit BR, Subedi A, Upadhyay MP, Pandey YR, Jarvis DI (2000b) *In situ* crop conservation: findings of agro-ecological, crop diversity and socio-economic baseline survey of Begnas ecosite, Kaski, Nepal. NP working paper No. 2/2000, NARC/LIBIRD/IPGRI. Rome, Italy

Rana RB, Chaudhari P, Gauchan D, Khatiwada SP, Sthapit BR, Subedi A, Upadhyay MP, JarvisDI (2000c) *In situ* crop conservation: findings of agro-ecological, crop diversity and socio-economic baseline survey of Kachorwa ecosite, Bara, Nepal. NP working paper No. 1/2000, NARC/LIBIRD/IPGRI, Rome, Italy

Riek J De, Calsyn E, Everaert L, Van bockstaele E, Loose M de (2001) AFLP based alternatives for the assessment of distinctness, uniformity and stability of sugarbeet varieties. Theor Appl genet 103: 1254-1265

Rijal DK, Rana RB, Sherchand KK, Sthapit Br, Panday YR, Adhikari N, Kadayat KB, Gautam YP, Chaudhary P, Paudel CL, Gupta SR, Tiwari PR (1998) Strengthening the Scientific

Basis for *In situ* Conservation of Agrobiodiversity: Findings of *In situ* selection exercise in Kaski. NP Working paper No. 1/98 NARC/LIBIRD/IPGRI

Roder M, Plaschke J, Koenig SU, Boerner A, Sorrells ME, Tanksley SD, Canal MW (1995) Abundance, variability, and chromosomal location of microsatellites in wheat. Mol Gen Genet 246: 327-333

Rohlf FJ (1990) NTSYS-pc Numerical taxonomy and multivariate analysis system. Version 1.60. Applied Biostatistics, New york

Rohlf FJ (1992) NTSYS-pc Numerical taxonomy taxonomy and multivariate analysis system version 1.8. Owner manual

Rohlf FJ (1993) NTSYS-pc Numerical Taxonomy and multivariate analysis system version 1.8. State University of New York, Stony Brook, NY Exeter Software, Setauket, New York

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor appl genet 90: 43-48

Ruiz M, Varela F, Carrillo JM (1997) Analysis of the discriminating power of agro-morphological and biochemical descriptors in a sample of the Spanish collection of barley (*Hordeum vulgare* L.), Genetic Resources and Crop Evolution 44: 247-255

Saghai-Maroof MA, Biyashev, RM, Yang GP, Zhang O, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA population dynamics. Proc Natl Sci USA 91: 5466-3470

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Santos CAF (1999) *In situ* evaluation of fruit yield and estimation of repeatability coefficient for major fruit traits of umbu tree [*Spondias tuberose* (Anacardaceae)] in the semi-arid region of Brazil. Genetic Resources and Crop Evolution 46: 455-460

Schaal BA, Leverich WJ, Rogstad SH (1991) Comparision of methods for assessing genetic variation in plant conservation biology. pp 123-134 In DA Falk, KE Holsinger (eds). Genetics and Conservation of rare plants Oxford University Press, New York

Schung MD, Hutter CM, Wetterstrand KA, Gaudene MS, Mackay TFC, Aquadro CF (1998) The mutation rates of di-, tri, and tetranucleotide repeats in *Drosophola melanogaster*. Mol Biol Evol 5: 1751-1760

Schut JW, Qi X, Stam P (1997) Association between relationship measures based on AFLP markers, pedigree data, and morphological traits in barley. Theor Appl Genet 95: 1161-1168

Sebastian L, Garcia JS, Quiloy SE, Hipolito LR, Sanchez PL, Calibo MC, Pham JL (1998) Assessment of diversity and identity of farmers' rice varieties using microsatellite markers. Plant and Animal Genome VI Conference, Jan. 18-22. Town and country hotel, San Diego, CA

Second G (1982) Origin of the genetic diversity of cultivated rice (*Oryza sativa* L.): study of the polymorphism scored at 40 isozyme loci. Japanese Journal of Genetics 57:25-57

Senior ML, Heun M (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome 36: 884-889

Senior ML, Murphy JP, Goodman MM, Stuber CW (1998) Utility of SSRs determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci. 38: 1088-1098

Serikawa T, Kuramoto T, Hillbert P, Mori M, Yamada J, Dubay CJ, Lindpainter K, Ganten D, Guenet JL (1992) Rat gene mapping using PCR analyzed microsatellites. Genetics 131: 701-722

Shahi BB, Heu MH (1979) Low temperature problem and research activities in Nepal. pp 61-68. In Report of a rice cold tolerance workshop, IRRI, Philippines

Shahi BB, Morishima H, Oka HI (1969) A survey of variation in peroxidase, acid phosphatase and esterase isozymes of wild and cultivated *Oryza* species. Japanese Journal of Genetics 44: 303-319

Shannon CE (1948) A mathematical theory of communication. Bell System Technology Journal 27:379-423, 623-656

Sherchand KK, Adhikari NP, Khatiwada SP, Shrivastav AC, Bajracharya J, Joshi KD, Kadayat KB, Chaudhary M, Chaudhary P, Bishwakarma SS, Yadav S (1998) Strengthening the Scientific Basis for *In situ* Conservation of Agrobiodiversity: Findings of *In situ* selection exercise in Bara. NP Working paper No. 2/98 NARC/LIBIRD/IPGRI

Shrestha GL, Vaughan DA (1989) The wild relatives of rice in Nepal. In S Lyama and G Takedo (eds) Proceedings of the 6th International Congress of SABRAO Proc I from August 21-25, 1989, Tsukuba, Japan

Shrivastava JP, Damania AB (1989) Use of collections in cereal improvement in semi arid areas. pp 88-104 In AHD Brown, OH Frankel, DR Marshall, JT Williams (eds) The use of plant genetic resources Cambridge University press, Cambridge

Sinha PK, Chauhan VS, Prasad K, Chauhan JS (1991) Genetic diversgence in indigenous upland rice varieties. Indian Journal of Genetics and Plant breeding 50: 47-50

Slatkin M (1989) A comparison of three indirect methods for estimating average levels of geneflow. Evolution 43: 1349-1368

Smith S, Helentjaris T (1996) DNA fingerprinting and plant variety protection. pp 95-110 In AH Paterson (ed) Genome Mapping in Plants, RG Landes, New York

Smith TB, Wayne RK (1996) Molecular genetic approaches in conservation. Oxford University Press, London

Sneath PHA, Sokal RR (1973) Numerical taxonomy. The principle and practice of numerical classification. W.H. Freeman and Co, San Fransisco

Sperling L, Loevinsohn ME, Ntabomvra B (1993) Rethinking the farmers' role in plant breeding: local bean experts and on-station selection in Rwanda. Experimental Agriculture 29:509-519

Srivastava A, Gupta V, Pental D, Pradhan AK (2001) AFLP-based genetic diversity assessment amongst agronomocally important natural and some newly synthesized lines of *Brassica juncea*. Theor Appl Genet 102:193-199

Sthapit BR (1994) Genetics and physiology of chilling tolerance in Nepalese rice. A thesis submitted to the University of Wales for the degree of philosophiae doctor in plant biology. School of Biological Sciences, University of North Wales, Bangor, UK

Sthapit BR, Joshi KD, Witcombe JR (1996) Farmer participatory crop improvement. III: Participatory plant breeding; a case study for rice in Nepal. Experimental agriculture 32: 479-496

Sthapit BR, Joshi KD, Rana RB, Upadhayay MP, Eyzaguirre P, Jarvis DI (2001) Enhancing biodiversity and production through participatory plant breeding: setting breeding goals. In *An exchange of experiences from South and South East Asia*. pp Proceedings of the International Symposium on Participatory Plant genetic Resource Enhancement, Pokhara, Nepal, 1-5 May, 2000 Cali, Colombia: Participatory Research and Gender Analysis Program, Coordination Office; International Center for Tropical agriculture

Stoskopf NC (1985) Variability in cereal grains *In Cereal Grain Crops* Part one. Reston Publishing Company, INC. A Prentice-Hall Company, Reston, Virginia

Sun CQ, Wang XK, Li JC, Yushimura A (2001) Comparision of the genetic diversity of common wild rice (*Oryza rufipogon* Griff.) and cultivated rice (*O. sativa* L.) using RFLP markers. Theor Appl Genet 102: 157-162

Tateoka T (1963) Taxonomic studies of the *Oryza* III Key to the species and their enumeration. Bot Mag Tokyo 76: 165-173

Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci 36: 186-192

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17: 6463-6471

Tautz D, Ranz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucl Acids Res 12: 4127-4138

Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100: 697-712

Templeton AR (1994) Biodiversity at the molecular genetic level: Experiences from disparate macroorganisms. Phil Trans R Soc London 345: 59-64

Teshome A, Baum BR, Fahring L, Torrance JK, Arnason TJ, Lambert JD (1997) Sorghum (*Sorghum bicolour*) landrace variation and classification in North shewa and South Welo, Ethiopia Euphytica 97: 225-263

Thanh ND, Zheng HG, Dong, NV, Trinh LN, Ali ML, Nguyen HT (1999) genetic characterization in root morphology and microsatellite DNA loci in upland rice (*Oryza sativa* L.) from Vietnam. Euphytica 105: 43-51

Thomas C, Vos P, Zabeau M, Jones D, Norcott K, Chadwick B, Jones J (1995) Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato CF-9 gene for resistance to *Cladosporium fulvum* Plant J 8: 785-794

Thomas MR, Scott NS (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence tagged sites (STSs). Theor Appl Genet 86: 986-990

Tilman D (1996) Biodiversity: population versus ecosystem estability. Ecology 77:350-363

Tilman D, Wedin D, Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature 379: 718- 720

Tranquilli G, Appendino ML, Pfluger LA, Bullrich L, Lewis S, Suarez EY (2000) Morphoagronomic and biochemical variation in an Argentinean landrace of wheat. Genetic Resources and Crop Evolution 47: 281-284

Tolbert DM, Qualset CO, Jain SK, Craddock JC (1979) A diversity analysis of a world collection of barley. Crop Sci 19: 789-794

UNCED (1992) Convention on Biological Diversity. United Nations Conference on Environment and Development, Geneva

Upadhyay MP (1995) Food crop genetic resources. In Plant genetic resources: Nepalese perspectives, Proceedings of the national workshop on plant genetic resources

conservation, use and management organized by NARC at Kathmandu, 28 Nov.-1 Dec 1994

Upadhyay MP, Sthapit BR (1995) Plant Genetic Resource Conservation Programs in Nepal: Some proposals for scientific basis of *in situ* conservation of Agrobiodiversity. Paper presented at the participants' conference on *In Situ* Conservation of Agricultural Biodiversity, Hosted by IPGRI in Rome, 17-19 July, 1995

Vavilov NI (1926) Centres of origin of cultivated plants. Bulletin of Aplied Botany and Plant Breeding, Leningrad 16, 139-248

Vavilov NL (!930) The problem of the origin of cultivated plants and domestic animals as conceived at the present time. Plant Breed Abstr 1(4):10

Virk PS, Ford-Lloyd BV, Jackson MT, Pooni HS, Clemeno TP, Nedwbury HJ (1996) Predicting quantitative variation within rice germplasm using molecular markers. Heredity 76: 296-304

Virk PS, Ford-Lloyd BV, Newbury HJ (1998) Mapping AFLP markers associated with sub specific differentiation of *Oryza sativa* and an investigation of segregation distortion. Heredity 81: 613-620

Virk PS, Newbury HJ, Jackson MT, Ford-Lloyd BV (1995) The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. Teor Appl Genet 90: 1049-1055

Virk PS, Newbury HJ, Jackson MT, Ford-Lloyd BF (2000) Are mapped markers more useful for assessing genetic diversity? Theor Appl Genet 100: 607-613

Virk PS, Zhu J, Newbury HJ, Bryan GY, Jackson MT, Ford-Lloyd BV (2000) Effectiveness of different classes of molecular markers for classifying and revealing variation in rice (*Oryza sativa* L.) germplasm. Euphytica 112: 275-284

Vos P, Hogers R, Bleekers M, Rijans M, Van de Lee T, Hornes M, Frizters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) A new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414

Wang ZY, Tanksley SD (1989) Restriction fragment length polymorphism in *Oryza sativa* L. Genome 32: 1113-1118

Wang ZY, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88: 1-6

Ward J (1963). Hierarchical grouping to optimize an objective function. J Am stat Assoc 58:236-244

Weber JL (1990) Informativeness of human (dC-DA)n (dG-dT)n polymorphisms. Genomics 7: 524-530

Weber JL, May PL (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. AM J Hum Genet 44: 388-396

Weir BS, Cockerham CC (1984) Estimating F statistics for the analysis of population structure. Evolution 38: 1358-1370

Weising K, Nybom H, Wolff, K, Meyer W (1995) DNA fingerprinting in plants and fungi. CRC Press, London

Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18: 7213-7218

Weltzien E (1989) Differentiation among barley landrace populations from the near east. Euphytica 43: 29-39

Weltzien E, Whitakar ML, Dhamatron M (1996) Breeding pearl millet with farmers in Rajasthan. In using Diversity. Proceedings of conference on using diversity and maintening genetic resources on farm. New Delhi

Widawasky D, Rozelle S (1998) Varietal diversity and yield variability in Chinese rice production. pp 159-172 In M Smale (ed). Farmers, Gene Banks, and Crop Breeding: Economic Analysis of Diversity in Wheat, Maize, and Rice. CYMMIT, Kluwer Academic publisher, Boston

William JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18: 6531-6535

Witcombe JR (1999) Do farmer-participatory methods apply more to high potential areas to marginal ones? *Outlook* on Agriculture 28(1): 43-49

Witcombe JR (1999) Does plant breeding leads to a loss of genetic diversity? pp 245-272 In D Wood, JM Lenne (eds) Agrobiodiversity: Characterization, utilization and management. CAB International, UK

Witcombe JR (2001) Participatory varietal selection in high-potential production systems. An exchange of experiences from South and South East Asia. pp 19-28 Proceedings of the international symposium on participatory plant breeding and plant genetic resource enhancement, Pokhara, Nepal, 1-5 May 2000. Cali, Colombia: Participatory Research and Gender Analysis Program, Coordination Office; International Center for Tropical agriculture Cali, Colombia: Participatory Research and Gender Analysis Program, Coordination Office; International Center for Tropical agriculture

Witcombe JR, Gilani MM (1979) Variation in cereals from the Himalayas and the optimum strategy for sampling plant germplasm. Journal of Applied Ecology. 16: 633-640

Witcombe JR, Joshi A, Joshi KD, Sthapit BR (1996) Farmer participatory crop improvement. I: Varietal selection and breeding methods and their impact on biodiversity. Experimental Agriculture 32: 445-460

Witcombe JR, Rao AR (1976) The genecology of wheat in a Nepalese center of diversity. J Appl Ecology. 13: 915-924

Witcombe JR, Subedi M, Joshi KD (2001) Towards a practical participatoruy plant breeding strategy in predominantly self-pollinated crops. In An exchange of experiences from South and South east Asia:Proceedings of the international symposium on participatory plant breeding and participatory plant genetic resource enhancement, Pokhara, Nepal, 1-5 May 2000 Cali, Colombia: Participatory Research and Gender Analysis Program, Coordination Office; International Center for Tropical agriculture

Worede M (1991) Crop genetic resources conservation and utilization: an Ethipion persepective, Science in Africa: Achievements and Prospects. Proceedings of the symposium of the American association for advancement of science, 15 February 1991, Washington DC

Worede M, Tesemma T, Feyissa R (2000) Keeping diversity alive: an Ethipion perspective. pp 143-164 In SB Brush (ed) Genes in the field: On-Farm conservation of crop diversity, IDRC/IPGRI and Lewis publishers, CRC press LLC

Wouw M Van de, Hanson J, Nokoe S (1999) Observation strategies for morphological characterization of forages. Genetic resources and Crop Evolution 46: 63-71

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241: 225-235

Xiao J, Li J, Grandillo S, Ahn SN, McCouch SR, Tanksley SD, Yuan L (1996) A wild species contains genes that may significantly increase the yield of rice. Nature 384: 223-224

Yang GP, Saghai Maroof MA, Xu CGU, Zhang O, Biyashov RM (1994) Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol Gen Genet 24: 187-194

Yee E, Kidwell KK, Sills GR, Lumpkin TAS (1999) Diversity among selected *Vigna angularis* (Azuki) accessions on the basis of RAPD and AFLP markers Crop Sci 39: 268-275

Yu LX, Nguyen NT (1994) Genetic variation detected with RAPD markers among upland and lowlad rice cultivars (*Oryza sativa* L.). Theor Appl Genet 87: 668-672

Zeven AC (1998) Landraces: A review of definitions and classification. Euphytica 04: 127-139

Zeven AC, Zhukovsky PM (1975) Dictionary of cultivated plants and their centers of diversity. PUDOC, Wagenigen, 219

Zhang Q, Saghai Maroof MA, Lu TU, Shen BZ (1992) Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. Theor Appl Genet 83: 495-499

Zhang Q, Yang GP, Dai X, Sun JZ (1994) A comparative analysis of genetic polymorphism in wild and cultivated barley from Tibet using isozyme and ribosomal DNA markers. Genome 37: 631- 638

Zhao X, Kochert G (1992) Characterization and genetic mapping of a short, highly repeated, interspersed DNA sequence from rice (*Oryza sativa* L.). Mol Gen Genet 231: 353-359

Zhao X, Kochert, G (1993) Phylogenetic distribution and genetic mapping of a (GCG) microsatellite from rice (*Oryza sativa* L.). Plant Mol Biol 21:607-617

Zhong GY, Qualset CO (1995) Quantitative genetic diversity and cinservation strategies for an allogamous annual species *Dasypyrum vilosum* (L.) Candargy (Poaceae). Theor Appl Genet 91: 1064-1073

Zhu J (1996) DNA fingerprinting in *Oryza sativa* L. A Ph. D. thesis submitted to University of East Anglia, John Innes Centre, UK

Zhu J, Gale M, Quarrie s, Jackson MT, Bryan G (1998) AFLP markers for the study of rice biodiversity. Theor Appl Genet 96: 602-611

Zhu JH, Stephenson P, Laurie DA, Li W, Tang D, Gale MD (1999) Towards rice genome scanning by map-based AFLP fingerprinting. Mol gen Genet 261: 184-195

Ziegle JS, Su Y, Corcoran KP, Nie L, Mayrand PE, Hoff LB, McBride LJ, Kronick MN, Diehl SR (1992) Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14: 1026-1040

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

Appendix 3.1

General description of the study sites

Variables	Jumla (High hill)	Kaski (Mid hill)	Bara (Terai)
Boundaries of study sites	Tallium (wards 1-9) and Kartikswami (wards 8 & 9)	Begnas (wards 9 & 10) Rupakot (wards 1, 7, & 8)	Kachorwa (1-6 wards)
Latitude	29° 6' N to 29° 19.4' N	28° 5.5' to 28° 13' N	26° 52.08' to 26° 54.1' N
Longitude	82° 5.5' to 82° 14.9' E	84° 1.6' to 84° 11.2' E	85° 8.6' to 85° 10.7' E
Elevation	2240-3000 m	668-1206 m	80-90 m
Climate	Artic to cool temperate	Sub-tropical	Sub-tropical
Mean annual temperature	10.2°C	20.9°C	24.6° C
Mean annual precipitation	866 mm	3979 mm	1515 mm
Total area under rice cultivation	81 ha	363 ha	NA
Average size of land hoiding	0.34 ha	0.73 ± 0.05 ha	0.69 ± 0.06 ha
Total households	759	941	914
No of rice landraces	10	69	55

Rice growing environments	Lekh (land situated at high altitude far from village)	*Pakho tari* – upland	Uncha khet – rainfed
	Bari (Upland around homestead, marginal uplands located far away from homestead)	*Tari* – lowland rainfed	*Samtal khet* – rainfed and irrigated
	Khet (stream irrigated; snow melted river irrigated and swampy lands)	*Kule khet* – lowland irrigated	*Nicha/Khala khet* – irrigated / wetland
		Sim khet – marshy land	*Pokhari/man khet* – water logged

NA = Not available

Appendix 3.2a

(a) Reagents and consumables

- DNA extraction kits – Qiagen LTD., UK
- dNTPs (4 x 40 µm of 200mM) – Promega Express, UK
- *Taq* DNA polymerase with $MgCl_2$ (25 µmM stock) and 10 x PCR buffer 100 U - Promega Express, UK
- Agarose SFR, Amresco Agarose SFR – Anachem LTD, UK
- 1 kb DNA ladder, 500 µl for 100 lanes – Promega Express, UK
- Agarose, 500 g - Amresco, UK.
- 1 kb DNA ladder – Gibco BRL, Life technologies
- 2 x Reddy Mix^{TM} PCR Master mix (3.0 mM $MgCl_2$) – Abgene House, UK
- Ethidium bromide, 10 ml – Promega Express, UK
- Acetic acid glacial, AR, 2.5 litre EM/H&W – HM, India
- Boric acd, AR, 500 g – HM, India
- Bromophenol blue, 5 g – Sigma, USA
- EDTA disodium salt ($Na_2EDTA.2H_2O$), 100 g – HM, India
- Ficoll 400 – Sigma, USA
- Hydrochloric acid, AR, 2.5 litre – Exellar, India
- Sodium hydroxide (NaOH) pellets, 100 g – Exellar, India
- Oligonucleotide primers – Research Genetics, USA
- Tris Free Base, 500 g – HM, India
- Mineral oil – Sigma, USA
- Autoclave tape
- Polaroid high speed, coaterless black and white type 667 camera film
- Eppendorfs (0.5 ml and 1ml)
- Disposable gloves
- Detergent
- Pipette tips (10 µl and 100 µl)
- Tissue papers and alluminium foils
- Orange gel loading buffers (Orange green and glycerol mixed)
- Lambda DNA – Promega Express, UK
- Glycerol – Sigma, USA
- Liquid Nitrogen available from Animal Breeding Division, Department of Agricultural development, Khumaltar, Nepal

Appendix 3.2b

(b) Stock solutions for SSR (microsatellite) analysis

- **1M Tris** – 121 g of Tris was dissolved in 800 ml water and adjusted to pH 8.5 with concentrated HCl and made up to 1 l with water and autoclaved.
- **5M NaCl** – 292 g of NaCl was dissolved in 750 ml water and made up to 1 l with water and autoclaved.
- **0.5M EDTA** – 186 g of $Na_2EDTA.2H_2O$ was dissolved in 800 ml water and adjusted to pH 8.0 with NaOH pellets and made up to 1 l with water and autoclaved.
- **100 x TE** – 121 g Tris and 37.2 g $Na_2EDTA\ 2H_2O$ was dissolved in 800 ml water, adjusted to pH 8.0 with concentrated HCl, made up to 1 l and autoclaved.
- **1 x TE** – 5 ml of 100 X TE was added to 495 ml of water and autoclaved.
- **50 x TAE** – 242 g Tris was dissolved in 500 ml water. 100 ml 0.5M EDTA and 57.1 ml glacial acetic acid added and made upto 1 l.
- **1 x TAE** – 200 ml of 50 x TAE was added to 9.8 l water in a 10 l bottle.
- **50 x TBE** – 270 g Tris (base), 138 g boric acid dissolved in 500 ml water. 100 ml 0.5 M EDTA added and made upto 1 l.
- **1 x TBE** – 200 ml of 50 x TBE was added to 9.8 l water in a 10 l bottle.
- **Gel loading buffer** – 100 mg bromophenol blue and 372 mg $Na_2EDTA.2H_2O$ dissolved in 100 ml glycerol and 10 ml of water was added to this and stored in fridge.

Appendix 3.2c

(c) List of equipment

- Horizontal Gel Electrophoresis apparatus, Sigma Comapany, USA.
  Biomax HR – 2025 High Resolution
  Gel dimensions – 20 x 25 cm with 160 samples, 1200 ml buffer capacities.
- UV Transluminator compact with 254/312 UV nM bands
  20 x 20 cm viewing area and 220 V, USA
- Gel cam camera system
  D3-34 with components:
  C0540 Gelcam camera, 20 x 25 cm flat face hood with 7inch diagonal,
  Filter holder and filters for ethidium bromide, coomassie blue, SYBR green, Silver stain and ELISA stain, Sigma Company, ltd, USA.
- Biofuge 15 Heralcus centrifuge,
  Maximum speed – 15,000 rpm
  Timer – 30 minutes
  Dimensions – 28 x 21 x 28 cm, can use 1 ml, 0.5 ml and 2 ml tubes.
- Hettich Zentrifugen – EBA 12/12R, 230 V, Hettich company, Germany
  Time to maxumum speed – 10 seconds
  Time to stop – 10 seconds
- PTC – 100™ programmable thermal controller with cold lid
  96 V Block holding 96 x 0.2 ml tubes or one 96 well V bottom PCR plate
  MJ Research, INC Waltham, MA, USA
- Vortex Genie 2™ (Homogeniser)
  G560E model, Scientific Industries INC., USA
- Ice maker and ice slicer 115 volt– HoshiZaki Electric Co. Ltd, Japan
- Digital, electronic balance with range 1mg to 120 g
- Fridge and freezer with lowest temperature – 86° C.
- Power Pack 3000 – Bio-rad laboratories, 2000, Alfred Nobel Drive, Herculus, CA 94547, USA.
- Water bath (100° C)
- Hewlett Packard 8453 spectrophotometer
- Generater for back up electricity in running PCR thermal cycling.
- Micropipettes of 1000ul, 10-100, and 1-20ul – Proline, Germany.
- Water distillation unit
- Local bio-safety facilities for handling and managing the hazardous waste.

Appendix 3.2d

Protocol: DNeasy Mini Kits (Qiagen Ltd, Crawley, UK)

- Water bath was turn on to 65°C and the supplied AE buffer was incubated.
- 100mg fresh leaf material was ground to a fine powder with pastle and mortar with freezing in liquid nitrogen.
- Ground tissue was put into labeled microfuge tube and 400 µl AP1 buffer was added.
- To this 4 µl Rnase A was added and vortex to remove clumps.
- This was Incubated to 65°C for 10 minutes to lyse cells and mixed the content by inverting tube 2-3 times.
- 130 µl AP2 buffer was added , mixed and incubated on ice for 5 minutes.
- The supernant was transferred to liliac QIAshredder spin column in a collection tube and spinned for 2 minutes at max. speed of 13000 rpm.
- Flow through but not the pellet was transferred to a clean microfuge tube labeled with genotype name.
- 225 µl AP3 buffer and 450ul 100% ethanol was added to cleared lysate very gently and mixed with tip.
- 650 µl of the mixture applied into the Dneasy column (clear) in a 2 ml collection tube and spinned for 1 minute at 8000 rpm.
- Remaining sample was repeated and reused the collection tube to spin again for 2 minutes at 8000 rpm and discarded the flow through.
- Dneasy column was kept in a new 2 ml tube and added with 500 µl AW buffer and spinned for I minute at 8000 rpm. Flow through but was discarded and collection tube with pellet was kept.
- 500 µl AW buffer was added to Dneasy column and spinned for 2 minutes at maximum speed of 13000 rpm to dry column membrane.
- Column was removed carefully and discarded the collection tube and contents.
- DNeasy column was transferred to a microfuge tube and pipet out 100 µl preheated AE buffer directly onto column membrane and incubated at room temperature for 5 minutes.
- It was spinned for 1 min at 8000 rpm.

Last two steps were repeated with a clean microfuge tube for a second 100ul of AE buffer for elution.

Appendix 3.2e

(i) Microsatellite Protocol with Taq polymerase (25ul)

1. Make up DNA Mix (for each DNA sample):

32 µl dNTPs stock mix (2.5mM)
32 µl 10x PCR buffer
215 µl SDW
5 µl DNA (to give 4ng/ml in PCR reactions)

DNA mix can be kept in freezer and defrosted each time PCR reactions set up.

2. Make up PCR reactions to 25 µl total volume by combining the appropriate mixes.

*Multiply these volumes to make enough for the number of PCR reactions required.

Primers Mix (for 2 reactions*):	0.5 µl of 20uM primer F
	0.5 µl of 20uM primer R
	4 µl water (SDW)
Taq Mix (for 2 reactions*):	5 µl $MgCl_2$
	3.5 µl water
	1µl 10x PCR buffer
	0.5 µl Taq DNA polymerase
For each reaction mix (25ml):	2.5 µl primers mix
	17.5 µl DNA mix
	5 µl *Taq* mix

3. PCR Programme : Touchdown Programme with low annealing temperature.

(ii) Microsatellite protocol with Ready mix master (25ul Reddy mix PCR)

1. Make up of DNA dilution for each DNA sample

   3 µl of stock Qiagen extract

   97 µl of sterilized distilled water (SDW)

2. Primer Mix for 1 PCR reaction

   0.25 µl of 20uM primer F

   0.25 µl of 20uM primer R

   2 µl of SDW

3. Reddy Mix PCR (25 µl)

   2.5 µl primer mix

   10.0 µl DNA dilution

   12.5 µl Reddy mix

Appendix 3.3

Details of microsatellite primer sequences, repeat motif, information content and expected product size (Wu and Tanksley, 1993; Panaud *et. al.,* 1996; Chen *et. al.,* 1997; Cho *et al.,* 2000; and Temnykh *et al.*, 2000).

Primer designation	Chromosome location	Forward_primer	Repeatt motif	PIC	Size range (bp)	Expected product size (bp) in IR36
RM226	1	agctaaggtctgggagaaacc	(AT)38	0.82	264-342	274
RM246	1	gagctccatcagccattcag	(GA)20	0.81	97-118	116
RM5	1	tgcaacttctagctgctcga	(GA)15	0.77	108-130	113
RM211	2	ccgatctcatcaaccaactg	(GA)18	0.56	144-163	161
RM213	2	atctgtttgcaggggacaag	(GA)17	0.63	127-141	139
RM48	2	tgtcccactgctttcaagc	(GA)17	0.82	199-221	204
RM203	2	cctatcccattagccaaacattgc	TATT(AT)21CCCCC			203
RM60	3	agtcccatgttccacttccg	(AATT)5	0.14	167-171	165
RM232	3	Ccggtatccttcgatattgc	(GA)24	0.78	142-166	158
RM22	3	ggtttgggagcccataatct	(GA)22	0.79	183-201	194
RM148	3	atacaacattagggatgaggctgg	CTCTAT(GT)12TTT	0.64	129-135	129
RM349	4	ttgccattcgcgtggaggcg	(GA)16	0.70	132-146	136
RM164	5	tcttgcccgtcactgcagatatcc	(GT)16TT(GT)4GAG	0.87	246-304	246
RM122	5	gagtcgatgtaatgtcatcagtgc	TAA(GA)7A(GA)2A(GA)11TTGC	0.63	227-233	227
RM26	5	gagtcgacgagcggcaga	(GA)15	0.60	102-112	112
RM3	6	acactgtagcggccactg	(GA)25	0.72	118-148	145
RM234	7	acagtatccaaggccctgg	(GA)25	0.83	133-163	156
RM248	7	tccttgtgaaatctggtccc	(GA)25	0.82	80-104	102
RM11	7	tctcctcttccccgatc	(GA)17	0.83	123-147	140
RM351	7	ccatcctccaccgcctctcg	(CCG)9(CGAAG)4	0.50	129-134	134
RM350	8	tgatcgtcgcgattcccggc	(CT)10	0.67	200-210	208
RM223	8	gagtgagcttgggctgaaac	(GA)25	0.77	139-163	165
RM201	9	ctcgttattacctacagtacc	(GA)17	0.65	144-158	158
RM215	9	caaaatggagcagcaagagc	(GA)16	0.63	147-153	148
RM257	9	cagttccgagcaagagtactc	(GA)24	0.73	121-173	147
RM242	9	Ggccaacgtgtgtatgtctc	(GA)26	0.81	193-225	225
RM228	10	ctggccattagtccttgg	(GA)36	0.83	108-154	154
RM222	10	cttaaatgggccacatgcg	(GA)18	0.76	199-215	213
RM244	10	ccgactgttcgtccttatca	(GA)8	0.63	157-165	163
RM167	11	gatccagcgtgaggaacacgt	GGAA(GA)16GGGG	0.70	127-159	128

RM206	11	cccatgcgtttaactattct	(GA)21	0.88	128-202	147
RM229	11	cactcacacgaacgactgac	(GA)11	0.83	106-131	116
RM224	11	atcgatcgatcttcacgagg	(GA)13	0.88	124-158	157
RM17	11	tgccctgttattttcttctctc	(GA)21	0.63	162-184	184
RM21	11	acagtattccgtaggcacgg	(GA)21	0.88	132-170	157
RM120	11	cacacaagccctgtctcacgacc	GT(GA)9TAG(ATC)4	n.a	n.a.	173
RM202	11	cagattggagatgaagtcctcc	(GA)30	0.82	158-186	189
RM247	12	tagtgccgatcgatgtaacg	(GA)16	0.84	130-176	131
RM20	12	atcttgtccctgcaggtcat	(ATT)14	0.82	162-198	234

Appendix 4.1

Sample of HHs selection for baseline survey in ecosites (Rana *et al.*, 2000)

	Name of farmer group	Wealth categories (population)			Wealth categories (sample)			Total sampled HH†
		Rich	Medium	Poor	Rich	Medium	Poor	
Tallium, Kartikswami, Jumla								
1	Khalla silam	11	18	13	3	4	4	11
2	Gharti-kami-damai	4	10	30	0	0	9	9
3	Shreedhuska bahun	16	33	11	5	8	3	16
4	Shreedhuska thar	13	22	17	3	5	4	12
5	Baya katiya	12	8	27	3	2	6	11
6	Sarki bada 'Ka'	2	22	17	2	3	5	10
7	Sarki bada 'kha'	3	17	26	1	4	6	11
8	Tallium thar bada	9	28	16	3	6	4	13
9	Lawar bada	6	9	13	1	0	4	5
10	Rokaya bada	11	18	14	3	6	2	11
11	Tallo rokaya bada	9	12	6	2	3	1	6
12	Damai bada	10	10	16	2	2	4	8
13	Dharala bada	12	13	10	3	3	2	8
14	Bhandari lawar	7	9	15	2	2	3	7
15	Budha bohara	7	15	20	2	3	5	10
16	Thapa lawar	7	9	30	2	1	8	11
17	Bhandari thapa	7	15	21	2	3	5	10
18	Siyal bada	3	16	24	1	4	6	11
	Total	**149**	**284**	**326**	**40**	**59**	**81**	**180**
Begnas, Rupakot, Kaski								
1	Unnatsil	14	20	4	3	4	1	8
2	Devisthan	12	27	9	4	5	2	11
3	Mahila sabisa	5	16	24	1	3	5	9
4	Upallo talbesi	12	5	7	3	1	1	5
5	Rupakot jamalkuna	5	9	5	1	2	1	4
6	Sundaridanda	14	42	8	3	9	2	14
7	Kholabesi	8	9	6	2	2	1	5
8	Paurakhe	24	12	5	5	3	1	9
9	Chaur	15	5	13	3	1	3	7
10	Kalimati	9	3	3	2	1	1	4
11	Archalathar	19	13	19	4	3	4	11
12	Viveksil aanmasamuha	29	27	18	6	6	4	16
13	Rupa sirjana aama	42	25	6	9	5	1	15
14	Punithar	12	4	9	3	1	2	6
15	Paudelthar	13	18	13	3	4	4	11
16	Aduwabari	31	34	14	7	7	3	17
17	Simalpata	13	7	15	3	1	3	7
18	Kotbari	18	23	17	4	5	4	13
19	Adhikarithar	7	17	4	1	4	1	6
20	Majthar	14	13	18	3	3	4	10
21	Dandathar	26	15	17	5	3	5	13
22	Bisaunathar	8	5	8	2	1	2	5
	Total	**350**	**349**	**242**	**77**	**74**	**55**	**206**
Kachorwa, Bara								
1	One ward	11	55	84	3	12	20	35
2	Two ward	8	69	86	1	14	21	36

3	Three ward	32	68	115	6	15	21	42
4	Four ward	12	24	42	3	6	8	17
5	Five ward	14	55	78	4	13	18	35
6	Six ward	19	53	89	4	13	20	37
	Total	**96**	**324**	**494**	**22**	**71**	**109**	**202**

Source: Findings of agro-ecological, crop diversity and socio-economic baseline survey reports of three sites, 2000).; † HH = households

Appendix 4.2

Use values of rice landraces

Landraces		Preferred use value	Not preferred use value
Begnas, Kaski ecosite (1999-2000) only reported landraces were included in the list (Sthapit et al., 2000)			
Larger area and many HHs†	Ekle	Good taste, long straw and good milling recovery	High water demanding, prone to insects in storage, high nutrient demanding and late maturing
	Madhese	Good taste, good milling recovery and good yield potential	Poor straw yield, high water and nutrient demanding
	Kathe gurdi	Adapted to low-input rainfed conditions, good taste and early maturing	Poor straw yield, poor milling recovery and low yield potential
	Mansara	Early maturing, adapted to very poor soil and low input conditions	Poor taste, poor straw yield and poor milling recovery
	Jethobudho	Good quality for aroma and soft, high price in market, quality straw for mat	High input demanding, low yield
	Thulo gurdi	Good taste, good milling recovery, good straw yield	High water and nutrient demanding
	Panhele	Qood quality for aroma, high price in market, good straw yield	Low yield potential, prone to insect and diseases, water demanding
	Jhinuwa	Good taste, good straw value and high price in market	High nutrient demanding, poor milling recovery, poor straw yield
Larger area and few HHs†	Sano madhese	High yield	Poor straw yield
	Gurdi	Adapted to low-input conditions, relatively good taste, good milling recovery, good straw yield	Poor yielding
	Rato anadi	Good with food culture value for local preparations, medicinal value,	Poor milling recovery, high input requirement, Low yielding
	Seto anadi	Good for local preparations, sticky rice	Coarse grain
	Jerneli	Adapted to low-input rainfed conditions, good taste, Early maturing	Low yielding, Poor milling recovery, high input demanding
	Bayarni	Good quality rice with aroma and soft, Medicinal value and straw good for mat	Low yielding, high input demanding
	Aanga	Adapted to very poor soil and rainfed plots, medicinal value and good fodder	Poor taste, low yielding, red pericarrped, more likely to wild traits.
Smaller area and few HH†	Lahare gurdi	Good taste, long straw, good milling recovery, adapted to cold water	High water demanding, Late maturity, high nutrient demanding

	Thulo madhese	Good straw yield, good taste, adapted to marginal lands, better milling recovery	Late maturing, prone to insect pest, sterility problem
	Sano gurdi	Adapted to rainfed conditions, adapted to shaded area, good milling recovery, good taste	Low yielding
	Naulo madhese	Long straw, drought tolerant	High input demanding
	Kalo gurdi	Long straw, good taste, adapted to shaded area,	Difficult to thresh, prone to false smut, high input demanding
	Tunde	Drought tolerant	
	Seto gurdi	Tolerant to moths, good milling recovery	
	Manamuri	Adapted to low-input conditions, easy to thresh, medicinal value	Low yielding
	Pakhe jerneli	Medicinal value, lodging tolerant, good taste, long straw, aroma, early maturing	Low milling recovery
	Gauriya	Good taste, fine grain, high yielding, good milling recovery, adapted to sandy soil,	Late maturing, awned grain
	Naltumme	Good taste, good for shaded area, early maturing, lodging and shattering tolerant, good milling recovery	Low straw yield
	Dhabe jerneli	Good taste, aromatic, long straw, good food culture value	Poor milling recovery, high nutrient demanding
	Ramani	Good quality with aroma, long straw, less prone to insects	Late and low yielding, poor quality straw
	Kalo jhinuwa	Good quality, good straw quality, aroma, adapted to water logging, shaded area, good food value	Low yield, late, difficult threshing
	Jhinuwa	Good quality, medicinal value, adapted to shaded area, low input requirement	
	Thapachini	Good for local preparation, adapted to marginal lands, good for beaten rice	Poor taste
	Masino jhinuwa	Fine grain, long straw, good taste	Poor yield
	Seto jhinuwa	Good quality, good milling recovery, long straw and aroma	Poor straw quality, high input demanding, prone to rodent damage
	Barmeli	High yielding and long straw	Poor milling recovery, poor eating quality, high input demanding
	Chobo	High yielding, good for local preparation	High nutrient demanding
	Juya bayarni	Good taste, good milling recovery, long straw, aromatic, long grain	Low yielding
	Kalo bayarni	Good yield, aromatic, good milling recovery, black grain and green rice	Low yielding

	Rate	Good taste, good milling recovery,	Difficulty in threshing
	Pakhe ramani	Good taste, aromatic, long straw, high tillering	Poor straw, low milling recovery,
	Seto bayarni	Good taste, good milling recovery, long straw, less shattering	High input demanding, prone to leaf roller attack
	Biramphool	Good quality rice, aromatic, long straw	Low yield, difficulty in threshing
	Basmati	Good quality, long straw, good milling recovery	Low yielding
Kachorwa, Bara ecosite (1999-2000) only reported landraces were included in the list (Sthapit *et al.*, 2000)			
Larger area and many HHs†	Basmati	Good quality rice, soft and taste, Aromatic, religepus and cultural significance, good price in market, good for local food culture	Lodging problem, susceptible to insect and diseases, low yielding, light grain weight
	Sotawa	Quality rice soft and tasty, medicinal value, less prone to insect and disease	Lodging problem, less responsive to irrigation
	Nakhisaro	Good quality rice, soft and tasty, respond well in low input conditions	Lodging problem, less responsive to to irrigation
Smaller area and few HHs†	Sathi	Religious and cultural significance, good price in market, good for local preparations	Low yielding
Larger area and few HHs†	Mutmur	Good taste, good under rainfed condition, low input, good sized grains	Lodging problem and less responsive to irrigation
Tallium, Jumla ecosite (1999-2000) only reported landraces were included in the list (Sthapit *et al.*, 2000)			
	Rato marshi	Tasty good quality, low disease problem, content feeling after consumption, religious value	Low yielding, dwarf,
	Seto marshi	Good quality rice with good taste, content feeling after consumption, easy milling, easy threshing, religious use	High shattering, less milling recovery, high disease incidence
	Kalo marshi	Good quality rice, more productive, easy milling, content feeling after consumption	High shattering, disease incidence

† HHs = households

Appendix 4.3a

Distribution of agro-ecological domains and rice landraces at Begnas, Rupakot, Kaski.

Domain	Rank of domain†	Productivity‡	Cultivated landraces
Pakho tari (Rainfed upland)	I	IV	Pakhe jhinuwa, Kathe gurdi, Mansara, Aanga
Tari (Rainfed land)	II	III	Jhinuwa, Pakhe jhinuwa, Kathe gurdi, Pakhe jarneli, Tunde, Pakhe tuned, Pakhe gurdi, Manamuri, Rate, Bhatte, Chhote.
Mule khet/Kule khet (Irrigation with seasonal canal)	III	I	Kalo jhinuwa, Panhelo jhinuwa, Lamcho jhinuwa, Seto jhinuwa, Masino jhinuwa, Andhere jhinuwa, Jhinuwa, Lahare gurdi, Thulo gurdi, Seto gurdi, Sano gurdi, Lahare kalo gurdi, Gurdi, Thulo kalo gurdi, Bayarni, Kalo bayarni, Seto bayarni, Gajale bayarni, Junge bayarni, Seto anadi, Rato anadi, Sano anadi, Naulo anadi, Dudhe anadi, Madhese, Thulo madhese, Sano madhese, Naulo madhese, Dabhe jarneli, Ramani, Aanpjhutte, Sano aanpjhutte, Gauriya, Ekle, Jethobudho, Rate, Krishnabhog, Thapachini, Bale, Dhabe gauriya, Battisara, Karna jira, Pani barmeli.Makaikhole
Sim khet (Marshy land with no drainage system)	IV	II	Kalo jhinuwa, Panhelo jhinuwa, Jhinuwa, Lamcho jhinuwa, Seto jhinuwa, Masino jhinuwa, Tarkaya jhinuwa, Junge jhinuwa, Andhere jhinuwa, Lahare gurdi, Thulo gurdi, Seto gurdi, Sano gurdi, Gajale gurdi, Gurdi, Thulo kalo gurdi, Bayarni, Kalo bayarni, Seto bayarni, Gajale bayarni, Junge bayarni, Seto anadi, Rato anadi, sano anadi, Dudhe anadi, Madhese, Thulo madhese, Sano madhese, Naulo madhese, Dhabe jarneli, Ramani, Kartike marshi, Panhele marshi, Seto marshi, Chiniya marshi, Aanpjhutte, Sano aanpjhutte, Gauriya, Ekle, Naltumme, Biramphool, Basmati, Chobo, Palungtare, Jyagmdikhole, Krishnabhog, Thapachini, Bale, Makaikhole, Dhabe gauriya, Barmeli, jadan, Battisara, Karnajira, Pani barmeli.

Source: 1998-99 Baseline Report (Rana *et al.*, 2000)

† = Ranking of domains is based on the availability of water.

‡ = Productivity of domains is based on the production potential of the landraces under respective domains

Appendix 4.3b

Distribution of agro-ecological domains and rice landraces at *Kachorwa,* Bara.

Domain	Rank of domain†	Productivity‡	Cultivated landraces
Ucha khet (Rainfed land for early maturing rice varieties)	I	III	Mutmur, Sotawa, Sokan, Sathi
Samtal khet (Irrigated land for Bhadaiya and aghani rice)	II	I	Lalka pharam, Nakhisaro, Sathi, Bhadaiya basmati, Khera, Aanga, Ujala pharam, Sotawa, Sokan, Dudhisaro, Kariya kamodh, Madhumala, Basmati, Karma
Nicha/Khalar khet (Irrigated for Aghani rice)	III	II	Basmati, Lajhi, Mansara, Karma, Madhumala, Batsar, Ratrani, Pharam, Kariya kamodh
Pokhari/Man (Pond for only Aghani deep root rice)	IV	IV	Bhatti, Silhat

Source: 1998-99 Baseline Report (Rana *et al.,* 2000)

† = Ranking of domains is based on the availability of water.

‡ = Productivity of domains is based on the production potential of the landraces under respective domains

Appendix 5.1a : **Correlation matrix of qualitative and quantitative traits with high repeatability on 292 rice accessions from the Kaski ecosite.**

Traits		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Seedling vigour	1	1.00																		
2nd leaf length	2	-0.06	1.00																	
2nd leaf width	3	-0.22	0.15	1.00																
Length / width ratio	4	0.12	0.50	-0.76	1.00															
Blade pubescence	5	-0.19	-0.04	0.49	-0.43	1.00														
Flag leaf angle	6	0.32	0.24	-0.35	0.44	-0.41	1.00													
Ligule length	7	0.03	0.29	0.32	-0.11	0.05	0.11	1.00												
Culm length	8	0.15	0.44	0.10	0.17	-0.14	0.44	0.33	1.00											
Culm number (tillers / plant)	9	-0.23	-0.19	-0.22	0.07	-0.20	-0.11	-0.15	-0.23	1.00										
Internode colour	10	0.32	0.06	-0.39	0.35	-0.34	0.36	-0.20	0.15	-0.14	1.00									
Culm strength	11	-0.23	-0.09	-0.30	0.21	-0.18	-0.02	-0.20	-0.03	0.35	0.09	1.00								
Panicle length	12	0.25	0.35	0.07	0.16	-0.04	0.44	0.29	0.56	-0.39	0.12	-0.31	1.00							
Secondary branching	13	0.27	0.23	0.18	-0.01	0.08	0.17	0.26	0.18	-0.54	0.06	-0.43	0.37	1.00						
Panicle exsertion	14	-0.37	-0.30	0.12	-0.28	0.28	-0.50	-0.27	-0.50	0.40	-0.26	0.18	-0.42	-0.47	1.00					
Panicle axis	15	0.32	0.26	0.27	-0.09	0.08	0.20	0.38	0.37	-0.62	0.11	-0.51	0.52	0.67	-0.54	1.00				
Shattering	16	-0.23	-0.24	-0.27	0.09	-0.16	-0.14	-0.35	-0.30	0.41	0.02	0.31	-0.39	-0.47	0.36	-0.57	1.00			
Threshability	17	-0.16	-0.18	-0.33	0.16	-0.26	-0.02	-0.26	-0.26	0.38	0.13	0.31	-0.35	-0.38	0.24	-0.47	0.71	1.00		
Days to heading	18	0.04	0.40	0.30	-0.01	0.13	-0.10	0.12	0.08	-0.32	-0.06	-0.17	0.15	0.17	-0.12	0.14	-0.20	-0.29	1.00	
Days to maturity	19	0.03	0.39	0.31	-0.03	0.15	-0.13	0.13	0.06	-0.30	-0.07	-0.17	0.15	0.17	-0.10	0.15	-0.20	-0.30	0.99	1.00
1000 grains weight	20	-0.53	-0.13	0.40	-0.40	0.48	-0.40	-0.13	-0.09	0.06	-0.34	0.12	-0.10	-0.28	0.53	-0.18	0.18	0.02	-0.07	-0.05
Well filled grains / panicle	21	0.45	0.34	-0.09	0.28	-0.15	0.38	0.15	0.29	-0.46	0.21	-0.30	0.41	0.56	-0.54	0.45	-0.36	-0.23	0.21	0.19
Empty hulls / panicle	22	0.15	0.37	0.05	0.19	-0.15	0.37	0.07	0.31	-0.19	0.12	-0.14	0.33	0.24	-0.37	0.23	-0.19	-0.14	0.25	0.22
Grains weight / panicle	23	-0.02	0.19	0.40	-0.22	0.38	-0.14	0.16	0.15	-0.46	-0.15	-0.34	0.34	0.43	-0.08	0.41	-0.34	-0.32	0.16	0.17
Yield / plot	24	-0.09	0.09	0.00	0.05	0.10	-0.17	-0.09	-0.15	0.00	0.08	0.00	-0.10	0.10	0.08	-0.04	0.04	0.08	0.19	0.19
Awning	25	0.18	-0.07	-0.01	-0.06	-0.16	0.19	0.16	0.25	0.14	0.02	0.12	0.07	-0.12	-0.18	-0.07	-0.14	-0.14	-0.05	-0.06
Apiculus colour	26	0.40	0.10	-0.17	0.20	-0.29	0.47	0.08	0.20	-0.25	0.31	-0.22	0.30	0.34	-0.53	0.34	-0.23	-0.12	0.11	0.10
Lemma and palea colour	27	-0.35	0.12	0.30	-0.17	0.31	-0.23	0.14	0.05	-0.11	-0.42	-0.03	0.06	0.08	0.20	0.11	-0.12	-0.12	-0.12	-0.11
Lemma and palea pubescence	28	0.01	-0.19	0.15	-0.25	0.09	-0.11	0.00	-0.23	-0.10	0.13	-0.24	-0.09	0.20	0.03	0.18	0.04	0.09	-0.08	-0.07
Sterile lemma colour	29	0.29	-0.20	-0.08	-0.08	-0.20	0.14	0.07	-0.14	-0.07	0.12	-0.14	0.00	0.22	-0.22	0.18	-0.07	0.05	-0.09	-0.08
Sterile lemma length	30	-0.27	-0.04	0.57	-0.46	0.48	-0.38	0.15	-0.10	-0.08	-0.59	-0.21	-0.01	0.09	0.24	0.09	-0.12	-0.28	0.16	0.18
Grain size (Rice length)	31	-0.30	0.07	0.19	-0.11	0.23	-0.12	0.06	0.17	0.00	-0.07	0.25	0.06	-0.20	0.15	-0.11	-0.05	-0.16	-0.16	-0.16
Grain shape (Rice)	32	-0.08	-0.06	0.02	-0.05	0.09	0.03	0.11	0.09	0.11	-0.04	0.24	0.07	-0.20	0.07	-0.17	-0.08	-0.13	-0.18	-0.17
Grain type (Paddy)	33	-0.12	-0.01	0.07	-0.06	0.11	0.04	0.08	0.26	-0.03	-0.02	0.25	0.13	-0.18	-0.04	-0.06	-0.12	-0.18	-0.23	-0.23
Bran colour	34	0.06	0.18	-0.06	0.15	-0.05	0.09	-0.15	0.05	0.13	0.03	0.06	-0.01	-0.10	0.07	-0.14	0.09	0.10	0.10	0.09

Note: Coloured values indicate correlation coefficient (r) at 1% and alphabet in blue at 5% significance levels.

Code	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1															
2															
3															
4															
5															
6															
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20	1.00														
21	-0.52	1.00													
22	-0.33	0.39	1.00												
23	0.28	0.50	-0.01	1.00											
24	0.03	0.06	-0.07	0.17	1.00										
25	-0.16	0.01	0.03	-0.12	-0.20	1.00									
26	-0.46	0.40	0.34	-0.05	-0.08	0.23	1.00								
27	0.49	-0.16	-0.19	0.33	0.01	-0.19	-0.48	1.00							
28	-0.02	-0.07	-0.09	0.07	0.03	-0.08	0.32	-0.20	1.00						
29	-0.32	0.12	-0.05	-0.12	-0.02	0.13	0.63	-0.22	0.48	1.00					
30	0.48	-0.16	-0.11	0.30	-0.04	-0.10	-0.24	0.42	-0.01	-0.16	1.00				
31	0.48	-0.26	-0.25	0.21	-0.06	0.20	-0.34	0.50	-0.23	-0.27	0.11	1.00			
32	0.23	-0.16	-0.19	0.01	-0.15	0.37	-0.16	0.33	-0.26	-0.05	0.01	0.71	1.00		
33	0.29	-0.13	-0.13	0.09	-0.16	0.31	-0.22	0.42	-0.29	-0.13	0.05	0.80	0.80	1.00	
34	-0.06	0.09	0.22	-0.11	-0.04	0.00	-0.01	-0.13	-0.10	-0.11	-0.09	-0.22	-0.11	-0.17	1.00

Appendix 5.1b **: Correlation matrix of qualitative and quantitative traits with high repeatability on 196 rice accessions from the Bara ecosite.**

Traits		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
2nd leaf length	1	1.00																				
Length / width ratio	2	0.76	1.00																			
Ligule length	3	0.81	0.55	1.00																		
Culm lemgth	4	0.74	0.51	0.61	1.00																	
Culm number (tillers/plant)	5	-0.05	0.28	-0.16	-0.10	1.00																
Panicle length	6	0.51	0.32	0.46	0.47	-0.31	1.00															
Panicle exsertion length	7	0.16	0.05	0.27	0.36	-0.26	0.37	1.00														
1000 grains weight	8	0.44	0.15	0.36	0.40	-0.22	0.02	0.01	1.00													
Days to heading	9	0.76	0.74	0.74	0.59	0.10	0.40	0.15	0.35	1.00												
Days to maturity	10	0.73	0.71	0.73	0.53	0.10	0.39	0.11	0.35	0.97	1.00											
Well filled grains / panicle	11	0.21	-0.11	0.32	0.26	-0.51	0.54	0.46	0.04	0.06	0.11	1.00										
Empty hulls / panicle	12	0.06	-0.04	0.11	0.00	-0.19	0.22	0.06	-0.07	0.01	0.05	0.31	1.00									
Grains weight / panicle	13	0.43	0.00	0.51	0.44	-0.54	0.41	0.41	0.52	0.27	0.30	0.79	0.17	1.00								
Yield / plot	14	0.41	0.33	0.47	0.40	0.18	0.26	0.13	0.27	0.47	0.48	0.21	-0.01	0.33	1.00							
Blade pubescence	15	0.35	0.11	0.53	0.36	-0.24	0.25	0.27	0.33	0.34	0.35	0.37	0.14	0.50	0.37	1.00						
Blade colour	16	0.28	0.11	0.42	0.30	-0.25	0.29	0.44	0.33	0.35	0.35	0.29	0.01	0.49	0.17	0.49	1.00					
Flag leaf angle	17	0.25	0.25	0.21	0.14	0.11	-0.11	-0.10	0.13	0.31	0.33	-0.11	0.15	0.01	0.20	-0.12	0.04	1.00				
Collar colour	18	-0.15	-0.14	-0.09	-0.20	-0.10	-0.03	-0.08	0.00	-0.10	-0.09	-0.02	0.09	-0.04	-0.09	-0.06	0.07	-0.03	1.00			
Node colour	19	0.11	-0.18	0.05	0.22	-0.43	0.03	0.18	0.52	-0.10	-0.09	0.21	0.06	0.43	-0.02	0.16	0.31	-0.09	-0.03	1.00		
Culm angle	20	-0.23	-0.03	-0.39	-0.15	0.24	-0.41	-0.27	-0.04	-0.16	-0.15	-0.33	-0.33	-0.27	-0.03	-0.25	-0.37	0.12	-0.07	-0.11	1.00	
Culm strength	21	-0.20	0.03	-0.33	-0.16	0.16	-0.10	-0.16	-0.18	-0.11	-0.14	-0.31	-0.18	-0.35	-0.25	-0.17	-0.26	-0.20	0.05	-0.18	0.41	1.00
Panicle type	22	-0.03	0.32	-0.17	-0.11	0.46	-0.19	-0.19	-0.17	0.18	0.16	-0.44	-0.27	-0.42	0.02	-0.28	-0.27	-0.01	0.02	-0.31	0.43	0.35
Panicle axis	23	0.04	-0.18	0.03	-0.05	-0.28	0.33	-0.06	-0.03	-0.13	-0.08	0.30	0.27	0.17	-0.04	0.06	0.04	-0.01	-0.13	0.16	-0.23	-0.06
Secondary branching	24	0.24	-0.09	0.31	0.25	-0.32	0.17	0.20	0.29	0.15	0.16	0.30	0.05	0.47	0.16	0.22	0.29	0.20	-0.05	0.11	-0.31	-0.35
Awning	25	0.15	0.23	0.13	0.23	0.17	0.08	0.12	0.03	0.26	0.25	0.01	-0.08	0.10	0.18	0.13	0.40	0.37	-0.05	-0.07	0.21	0.03
Apiculus colour	26	-0.04	0.00	-0.04	0.16	0.06	-0.17	0.14	0.09	-0.04	-0.12	-0.21	-0.20	-0.06	-0.19	-0.08	0.22	-0.06	0.00	0.18	-0.16	0.16
Lemma and palea colour	27	-0.01	-0.04	0.06	0.04	0.09	0.03	-0.02	-0.12	-0.05	-0.02	0.10	-0.14	0.06	0.09	0.14	0.11	0.11	0.21	-0.17	0.12	-0.06
Lemma and palea pubescence	28	-0.11	-0.14	-0.27	-0.03	-0.05	-0.12	-0.18	0.21	-0.29	-0.32	-0.21	-0.20	-0.07	-0.27	-0.25	-0.24	-0.54	-0.03	0.21	0.03	0.24
Sterile lemma colour	29	0.17	-0.02	0.05	0.13	-0.17	0.01	-0.07	0.29	0.04	0.02	-0.03	-0.08	0.12	0.01	0.02	-0.04	0.09	-0.04	0.28	0.09	0.12
Grain size (Length) (Rice)	30	0.31	0.24	0.36	0.11	-0.07	0.34	0.02	-0.08	0.40	0.41	0.03	0.23	-0.01	0.10	0.15	0.23	0.43	0.00	-0.22	-0.32	-0.10
Grain shape (Rice)	31	0.33	0.29	0.41	0.20	0.05	0.35	0.15	-0.17	0.46	0.44	0.11	0.16	0.02	0.15	0.15	0.30	0.31	0.00	-0.20	-0.33	-0.07
Grain type (Paddy grain)	32	0.25	0.26	0.34	0.14	0.11	0.27	0.15	-0.22	0.37	0.35	-0.01	0.13	-0.09	0.06	0.11	0.28	0.24	-0.01	-0.25	-0.41	-0.02
Bran colour	33	-0.09	-0.06	-0.19	0.00	0.24	-0.37	-0.09	0.16	-0.21	-0.23	-0.18	-0.25	-0.05	0.04	-0.31	-0.36	-0.05	-0.09	-0.04	0.39	-0.01

Note: Coloured values indicate correlation coefficient (r) at 1% and alphabet in blue at 5% significance levels.

Code	22	23	24	25	26	27	28	29	30	31	32	33
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22	1.00											
23	-0.47	1.00										
24	-0.57	0.19	1.00									
25	0.10	-0.03	-0.06	1.00								
26	0.06	-0.45	-0.03	0.15	1.00							
27	-0.17	0.12	0.23	0.37	-0.15	1.00						
28	0.03	-0.02	-0.09	-0.33	0.21	-0.14	1.00					
29	0.01	0.09	0.15	-0.20	0.19	-0.16	0.21	1.00				
30	-0.14	0.35	0.26	0.10	-0.27	0.04	-0.43	0.08	1.00			
31	-0.13	0.23	0.25	0.13	-0.16	0.15	-0.52	-0.13	0.83	1.00		
32	-0.09	0.14	0.21	0.08	-0.03	0.15	-0.40	-0.19	0.79	0.95	1.00	
33	0.08	-0.30	-0.10	0.01	0.07	0.10	0.33	-0.14	-0.60	-0.48	-0.45	1.00

Appendix 5.2: Data for figure 5.4 (a)

Principal Component Analysis outputs based on 12 significant qualitative and quantitative traits

JUMLA ECOSITE:

Eigenanalysis of the Correlation Matrix

Eigenvalue	5.3619	2.5166	1.1870	0.8804
Proportion	0.447	0.210	0.099	0.073
Cumulative	0.447	0.657	0.755	0.829

Variable	PC1	PC2	PC3
LWIDTH	-0.049	0.118	0.744
SECBRA	-0.397	0.168	-0.184
SHATTER	0.367	-0.174	0.194
LEAFSENE	-0.397	0.168	-0.184
AWNING	-0.120	-0.578	0.100
STIGCOL	0.251	0.456	-0.006
LPCOL	0.396	0.035	0.130
GRAINSP	-0.336	0.142	0.292
HULLSP	-0.194	0.185	-0.138
GWP	-0.327	0.058	0.391
STLCOL	0.234	0.482	-0.070
BRANCOL	0.016	-0.256	-0.229

Principal Component Analysis outputs using 8 qualitative traits for Figure 5.4 (b)

Eigenanalysis of the Correlation Matrix

Eigenvalue	4.2229	2.3442	0.8964	0.2176
Proportion	0.528	0.293	0.112	0.027
Cumulative	0.528	0.821	0.933	0.960

Variable	PC1	PC2	PC3
SECBRA	0.431	0.283	0.089
SHATTER	-0.401	-0.287	-0.060
LEAFSENE	0.431	0.283	0.089
AWNING	0.201	-0.570	-0.152
STIGCOL	-0.338	0.401	0.120
LPCOL	-0.458	-0.079	0.005
STLCOL	-0.321	0.440	0.127
BRANCOL	0.014	-0.268	0.963

Principal Component Analysis outputs using 4 quantitative traits for Figure 5.4 (c)

Eigenanalysis of the Correlation Matrix

Eigenvalue	2.1109	0.9059	0.8538	0.1294
Proportion	0.528	0.226	0.213	0.032
Cumulative	0.528	0.754	0.968	1.000

Variable	PC1	PC2	PC3
LWIDTH	-0.296	0.843	0.448
GRAINSP	-0.627	-0.061	-0.350
HULLSP	-0.326	-0.531	0.781
GWP	-0.643	-0.059	-0.260

Appendix 5.3: Data for figure 5.5

Hierarchical Cluster Analysis of Observations outputs using 12 qualitative and quantitative traits

Pearson Distance, Ward Linkage

Final Partition

Cluster composition and their relative distance from the centroid

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	64	178.643	1.435	5.398
Cluster2	72	227.106	1.650	3.042
Cluster3	11	98.124	2.795	4.935
Cluster4	1	0.000	0.000	0.000

Mean values of clusters;

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Grand centrd
LWIDTH	0.1791	-0.2759	-0.0624	0.2762	-0.0595
SECBRA	-0.0822	-0.0822	-0.0822	12.0833	-0.0000
SHATTER	0.0253	0.1106	0.1106	-10.8003	-0.0000
LEAFSENE	-0.0822	-0.0822	-0.0822	12.0833	-0.0000
AWNING	-0.2536	-0.2536	3.1584	-0.2536	0.0000
STIGCOL	0.2382	0.2382	-2.5670	-4.1634	0.0000
LPCOL	0.1564	0.1704	-1.0682	-10.5271	0.0000
GRAINSP	0.0888	-0.2007	0.0488	8.2299	-0.0000
HULLSP	-0.0329	-0.0108	-0.2383	5.5032	-0.0000
GWP	0.1631	-0.3204	0.5020	7.1076	0.0000
STLCOL	0.2960	0.2960	-3.3551	-3.3551	0.0000
BRANCOL	-0.9035	0.6995	0.7631	-0.9326	0.0000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4
Cluster1	0.0000			
Cluster2	1.7613	0.0000		
Cluster3	6.1109	5.9343	0.0000	25.7920
Cluster4	26.5556	26.8666	25.7920	0.0000

Appendix 5.4: Data for figure 5.6 (a)

Principal Component Analysis out puts based on 32 qualitative and quantitative traits over 292 rice accessions from Kaski

Eigenanalysis of the Correlation Matrix

Eigenvalue	6.1154	4.8968	3.2448	2.1763
Proportion	0.197	0.158	0.105	0.070
Cumulative	0.197	0.355	0.460	0.530

Variable	PC1	PC2	PC3
SEEDVIG	0.238	0.073	-0.032
L_LENGTH	0.164	-0.141	-0.105
L_WIDTH	-0.078	-0.337	0.075
LBRATIO	0.160	0.195	-0.124
BLAPUB	-0.134	-0.247	0.091
FLANGLE	0.213	0.090	-0.157
LIGLEN	0.095	-0.171	-0.130
C_LENGTH	0.165	-0.109	-0.300
TILLERS	-0.188	0.233	-0.002
INTNCOL	0.115	0.250	-0.135
CULMSTRE	-0.151	0.172	-0.138
P_LENGTH	0.209	-0.175	-0.201
SECBRA	0.191	-0.116	0.122
PANEXSE	-0.274	-0.044	0.105
PANAXIS	0.225	-0.225	-0.034
SHATTER	-0.135	0.138	0.111
THRESH	-0.118	0.223	0.069
DH	0.126	-0.195	0.193
DM	0.119	-0.201	0.197
TGW	-0.289	-0.181	-0.017
GRAINSP	0.303	-0.084	-0.035
HULLSP	0.212	-0.040	-0.003
GWP	0.046	-0.312	-0.021
AWNING	0.033	0.058	-0.261
APICOL	0.289	0.075	0.002
LPCOL	-0.172	-0.238	-0.151
STLCOL	0.136	0.106	0.048
STLLEN	-0.142	-0.291	0.129
SIZE	-0.190	-0.119	-0.375
SHAPE	-0.132	-0.040	-0.415
GTYPE	-0.128	-0.081	-0.452
YPLOT	0.032	0.091	-0.128

Principal Component Analysis outputs of 292 rice accessions using 17 qualitative traits for Figure 5.6 (b)

Eigenanalysis of the Correlation Matrix

Eigenvalue	3.9378	2.7798	2.4004	1.1423
Proportion	0.232	0.164	0.141	0.067
Cumulative	0.232	0.395	0.536	0.604

Variable	PC1	PC2	PC3
SEEDVIG	0.294	-0.051	0.225
BLAPUB	-0.206	0.302	0.022
FLANGLE	0.236	-0.133	0.227
INTNCOL	0.199	-0.364	-0.011
CULMSTRE	-0.151	-0.347	-0.156
SECBRA	0.188	0.281	0.131
PANAXIS	0.164	0.288	0.335
SHATTER	-0.074	-0.155	-0.349
THRESH	0.006	-0.215	-0.324
AWNING	0.005	-0.276	0.281
APICOL	0.373	-0.038	0.224
LPCOL	-0.341	0.193	0.137
STLCOL	0.270	-0.053	0.157
STLLEN	-0.224	0.380	0.018
SIZE	-0.364	-0.158	0.276
SHAPE	-0.280	-0.258	0.360
GTYPE	-0.310	-0.223	0.372

Principal Component Analysis outputs using 15 quantitative traits for Figure 5.6 (c)

Eigenanalysis of the Correlation Matrix

Eigenvalue	3.9732	2.7358	1.7824	1.3200
Proportion	0.265	0.182	0.119	0.088
Cumulative	0.265	0.447	0.566	0.654

Variable	PC1	PC2	PC3
L_LENGTH	0.341	-0.016	-0.138
L_WIDTH	0.077	0.519	0.145
LBRATIO	0.143	-0.456	-0.217
LIGLEN	0.198	0.114	0.249
C_LENGTH	0.298	-0.060	0.312
TILLERS	-0.299	-0.150	-0.123
P_LENGTH	0.329	-0.004	0.283
PANEXSE	-0.275	0.261	-0.115
DH	0.282	0.256	-0.480
DM	0.275	0.268	-0.477
TGW	-0.194	0.384	0.121
GRAINSP	0.370	-0.130	0.069
HULLSP	0.288	-0.115	-0.041
GWP	0.207	0.303	0.220
YPLOT	0.020	0.092	-0.342

Appendix 5.5: Data for figure 5.7

Hierarchical Cluster Analysis of Observations outputs based on 31 qualitative and quantitative traits

Cluster composition and their relative distance from centroid

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	84	1659.715	4.307	10.111
Cluster2	43	552.754	3.527	5.441
Cluster3	45	547.996	3.437	5.288
Cluster4	32	672.492	4.561	5.583
Cluster5	68	1711.689	4.915	7.844
Cluster6	20	318.702	3.939	5.026

Mean values of the morphological traits for the clusters

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
SEEDVIG	0.5538	0.2889	0.4857	-0.9147	-0.4099
L_LENGTH	-0.0854	-0.5302	0.9489	-0.1623	0.1565
L_WIDTH	-0.0695	-0.2032	-0.3501	1.9587	-0.2060
LBRATIO	-0.0456	-0.1949	0.9165	-1.6603	0.2630
BLAPUB	-0.2547	-0.2735	-0.2223	1.8469	-0.1786
FLANGLE	0.4851	0.2008	0.5736	-1.1759	-0.3624
LIGLEN	0.3897	-0.1458	-0.1667	0.3971	-0.0539
INTNCOL	0.1468	0.2067	0.2067	-1.5220	0.2067
CULMSTRE	0.0887	-0.4409	-0.3710	-0.5648	0.1654
P_LENGTH	0.4301	-0.2850	0.7898	-0.1797	-0.2362
C_LENGTH	0.5708	-0.7609	0.5845	-0.4415	-0.0278
TILLERS	-0.1796	-0.0973	-0.4125	-0.3909	0.1054
PANEXSE	-0.6746	-0.2295	-0.6063	1.0690	0.5135
SECBRA	0.0139	0.5466	0.6536	0.2300	-0.5441
PANAXIS	0.2707	0.2707	0.2707	0.2707	0.2707
SHATTER	-0.2850	-0.0057	-0.2493	-0.1491	0.1601
THRESH	-0.4080	0.7411	-0.4933	-0.6736	0.2806
DH	-0.0291	-0.3775	0.5615	0.6351	-0.2132
DM	-0.0438	-0.3725	0.5181	0.6890	-0.1901
TGW	-0.4900	-0.5043	-0.5897	1.6039	0.3555
GRAINSP	0.2592	0.0736	1.2702	-0.6192	-0.4701
HULLSP	0.2529	-0.0858	0.8629	-0.4458	-0.4001
GWP	-0.1762	-0.1272	0.3838	1.2338	-0.1193
AWNING	0.9658	-0.4096	-0.4843	-0.3605	-0.5300
APICOL	0.5075	0.9942	0.5392	-0.7912	-0.9125
LPCOL	-0.2399	-0.3488	-0.5194	0.8103	0.5833
LPPUB	-0.3442	1.4907	-0.3590	0.5977	-0.3947
STLCOL	0.3185	1.5913	-0.6066	-0.6066	-0.5345
STLLEN	-0.2559	-0.2559	-0.2559	2.0789	-0.2559
SIZE	0.3467	-1.0716	-0.7845	0.6329	0.3638
SHAPE	0.7457	-1.0108	-0.7627	0.0991	-0.0013

Variable	Cluster6	Grand centrd
SEEDVIG	-1.1829	0.0000
L_LENGTH	-0.9091	-0.0000
L_WIDTH	-0.9170	-0.0000
LBRATIO	0.3103	-0.0000
BLAPUB	-0.1900	0.0000
FLANGLE	-0.6460	-0.0000
LIGLEN	-1.4003	-0.0000
INTNCOL	0.2067	0.0000
CULMSTRE	1.7513	-0.0000
P_LENGTH	-1.8805	0.0000
C_LENGTH	-1.2758	-0.0000
TILLERS	2.1586	0.0000
PANEXSE	1.2347	-0.0000
SECBRA	-1.2222	0.0000
PANAXIS	-3.6815	0.0000
SHATTER	1.4648	0.0000
THRESH	1.3541	-0.0000
DH	-0.6210	-0.0000
DM	-0.6370	-0.0000
TGW	0.6939	-0.0000
GRAINSP	-1.5159	-0.0000
HULLSP	-0.7454	-0.0000
GWP	-1.4183	-0.0000
AWNING	0.2926	0.0000
APICOL	-1.1140	-0.0000
LPCOL	-0.3537	0.0000
LPPUB	-0.5658	-0.0000
STLCOL	-0.6066	0.0000
STLLEN	-0.2559	-0.0000
SIZE	0.3638	-0.0000
SHAPE	0.6032	0.0000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0.0000	4.2116	3.4919	6.6889	3.6417
Cluster2	4.2116	0.0000	4.6200	7.1352	4.6151
Cluster3	3.4919	4.6200	0.0000	7.3755	4.5066
Cluster4	6.6889	7.1352	7.3755	0.0000	5.5759
Cluster5	3.6417	4.6151	4.5066	5.5759	0.0000
Cluster6	8.0490	8.1464	9.2243	9.1905	6.3211

	Cluster6
Cluster1	8.0490
Cluster2	8.1464
Cluster3	9.2243
Cluster4	9.1905
Cluster5	6.3211
Cluster6	0.0000

Appendix 5.6: Data for figure 5.9 (a)

rincipal Component Analysis out puts based on 33 qualitative and quantitative traits over 196 rice accessions from Bara

Eigenanalysis of the Correlation Matrix

Eigenvalue	7.9084	4.1964	3.3808	2.1015
Proportion	0.240	0.127	0.102	0.064
Cumulative	0.240	0.367	0.469	0.533

Variable	PC1	PC2	PC3
L_LENGTH	-0.280	-0.080	0.196
LBRATIO	-0.192	-0.284	0.161
BLAPUB	-0.196	0.128	0.088
BLACOL	-0.210	0.111	0.040
FLANGLE	-0.116	-0.174	-0.026
LIGLEN	-0.301	-0.026	0.111
COLLCOL	0.032	0.020	-0.089
NODECOL	-0.039	0.292	0.168
C_LENGTH	-0.237	0.001	0.243
TILLERS	0.082	-0.347	0.060
CULMANG	0.167	-0.142	0.190
CULMSTRE	0.128	-0.135	0.020
P_LENGTH	-0.231	0.074	-0.098
PANTYPE	0.086	-0.315	0.150
SECBRA	-0.164	0.170	-0.055
PANEXSE	-0.141	0.144	0.020
PANAXIS	-0.057	0.184	-0.247
DH	-0.278	-0.212	0.145
DM	-0.278	-0.198	0.127
TGW	-0.123	0.136	0.340
GRAINSP	-0.170	0.291	-0.078
HULLSP	-0.080	0.105	-0.202
GWP	-0.211	0.300	0.145
YPLOT	-0.181	-0.071	0.151
AWNING	-0.095	-0.127	0.091
APICOL	0.044	0.008	0.170
LPCOL	-0.036	-0.017	-0.074
LPPUB	0.163	0.127	0.193
STLCOL	-0.014	0.083	0.131
SIZE	-0.212	-0.150	-0.297
SHAPE	-0.230	-0.174	-0.290
GTYPE	-0.190	-0.186	-0.295
BRANCOL	0.151	-0.008	0.284

Principal Component Analysis outputs using 14 quantitative traits for Figure 5.9 (c)

Eigenanalysis of the Correlation Matrix

Eigenvalue	5.7310	2.6762	1.2583	0.9725
Proportion	0.409	0.191	0.090	0.069
Cumulative	0.409	0.601	0.690	0.760

Variable	PC1	PC2	PC3
L_LENGTH	-0.375	-0.114	-0.019
LBRATIO	-0.279	-0.339	0.198
LIGLEN	-0.364	-0.018	-0.003
C_LENGTH	-0.329	-0.009	-0.053
TILLERS	0.068	-0.467	0.176
P_LENGTH	-0.261	0.198	0.389
PANEXSE	-0.153	0.284	0.203
DH	-0.353	-0.234	0.047
DM	-0.349	-0.217	0.040
TGW	-0.199	0.018	-0.695
GRAINSP	-0.176	0.477	0.138
HULLSP	-0.049	0.211	0.385
GWP	-0.259	0.390	-0.277
YPLOT	-0.240	-0.085	-0.057

Principal Component Analysis outputs using 19 qualitative traits for Figure 5.9 (b)

Eigenanalysis of the Correlation Matrix

Eigenvalue	4.4168	2.4496	1.9308	1.7031
Proportion	0.232	0.129	0.102	0.090
Cumulative	0.232	0.361	0.463	0.553

Variable	PC1	PC2	PC3
BLAPUB	-0.173	0.221	0.241
BLACOL	-0.245	0.188	0.356
FLANGLE	-0.194	-0.240	0.204
COLLCOL	-0.008	-0.051	0.037
NODECOL	0.029	0.436	0.135
CULMANG	0.256	-0.292	0.150
CULMSTRE	0.162	-0.195	-0.221
PANTYPE	0.173	-0.415	-0.066
SECBRA	-0.234	0.256	0.078
PANAXIS	-0.150	0.224	-0.216
AWNING	-0.099	-0.216	0.517
APICOL	0.085	0.101	0.164
LPCOL	-0.105	-0.102	0.319
LPPUB	0.293	0.241	-0.200
STLCOL	0.038	0.203	-0.156
SIZE	-0.394	-0.143	-0.246
SHAPE	-0.410	-0.179	-0.152
GTYPE	-0.378	-0.176	-0.184
BRANCOL	0.302	-0.061	0.219

Appendix 5.7: Data for figure 5.10 and 5.11

Hierarchical Cluster Analysis of Observations outputs based on 23 qualitative and quantitative traits for 196 rice accessions from Bara

Cluster composition and their relative distance from centroid

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	16	132.326	2.823	3.974
Cluster2	10	30.345	1.657	2.659
Cluster3	67	941.162	3.707	5.277
Cluster4	15	77.672	2.134	4.499
Cluster5	53	673.182	3.466	5.568
Cluster6	16	80.080	2.197	3.023
Cluster7	19	131.445	2.513	4.188

Mean values of morphological traits of clusters

Variable	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
L_LENGTH	1.0300	-1.3455	0.7170	0.4708	-1.0456
LBRATIO	0.0437	-0.6834	0.6020	1.0238	-1.0052
BLAPUB	1.4493	-0.6865	0.5248	-0.6865	-0.3238
BLACOL	2.5321	-0.5892	-0.1699	0.8674	-0.2948
FLANGLE	0.3421	-1.1634	0.1548	-0.3605	-0.0727
LIGLEN	1.4151	-1.2303	0.7983	0.1339	-0.8042
C_LENGTH	1.4181	-1.0403	0.3495	0.6297	-0.8852
TILLERS	-1.2734	0.9513	0.0974	0.2189	-0.2775
CULMANG	-0.8883	-0.8883	-0.3718	-0.8162	0.1523
CULMSTRE	-1.4356	0.7420	-0.2302	0.8429	-0.2012
P_LENGTH	0.4579	-1.4710	0.3454	1.9247	-0.3257
PANTYPE	-0.9461	0.6254	0.0625	0.5206	-0.5606
PANEXSE	1.5189	-0.1175	-0.1580	0.9394	-0.0317
SECBRA	1.3840	-0.7188	0.1914	-0.7188	-0.0047
DH	0.8526	-1.5645	0.7481	0.9852	-1.1034
DM	0.8423	-1.6630	0.7709	0.8450	-1.0063
TGW	1.8180	-0.5600	0.1594	-1.0354	-0.6695
GRAINSP	1.1592	-1.6991	0.1771	0.1944	0.3832
HULLSP	0.2033	-0.6391	0.0776	-0.0991	0.4118
GWP	2.2096	-1.4324	0.2037	-0.4860	-0.1611
YPLOT	0.6516	-1.3331	0.5811	-0.2383	-0.4651
LPPUB	-0.4333	2.1402	-0.1644	-0.4333	-0.4333
SIZE	-0.2096	-1.2517	0.4903	0.9714	-0.1282

Variable	Cluster6	Cluster7	Grand centrd
L_LENGTH	-0.1471	-0.0187	-0.0000
LBRATIO	-0.8053	0.8741	0.0000
BLAPUB	-0.6865	-0.6865	-0.0000
BLACOL	-0.5892	-0.5892	-0.0000
FLANGLE	-1.1634	1.2453	-0.0000
LIGLEN	-0.8918	-0.4709	-0.0000
C_LENGTH	0.0132	0.0820	0.0000
TILLERS	-0.6612	1.3864	0.0000
CULMANG	1.0717	1.8437	0.0000
CULMSTRE	0.9313	0.7420	0.0000
P_LENGTH	-0.1050	-1.3517	0.0000
PANTYPE	-0.0621	1.4525	0.0000
PANEXSE	-0.5879	-0.8182	0.0000
SECBRA	-0.1931	-0.7188	0.0000
DH	-1.0250	0.6309	-0.0000
DM	-1.1281	0.5375	-0.0000
TGW	0.8918	0.1356	0.0000
GRAINSP	-0.3755	-1.6126	-0.0000
HULLSP	-0.9198	-0.4044	-0.0000
GWP	0.1781	-1.1422	-0.0000
YPLOT	-0.5326	0.0379	-0.0000
LPPUB	2.1402	-0.4333	0.0000
SIZE	-1.2517	-0.2488	0.0000

Distances Between Cluster Centroids

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Cluster1	0.0000	10.5038	5.2272	6.8177	7.5226
Cluster2	10.5038	0.0000	7.2271	7.5759	4.9876
Cluster3	5.2272	7.2271	0.0000	3.5552	4.5814
Cluster4	6.8177	7.5759	3.5552	0.0000	5.5411
Cluster5	7.5226	4.9876	4.5814	5.5411	0.0000
Cluster6	8.3028	4.4429	5.6400	6.6392	4.3358
Cluster7	9.1149	6.3614	4.9951	5.9761	5.6798

	Cluster6	Cluster7
Cluster1	8.3028	9.1149
Cluster2	4.4429	6.3614
Cluster3	5.6400	4.9951
Cluster4	6.6392	5.9761
Cluster5	4.3358	5.6798
Cluster6	0.0000	5.9502
Cluster7	5.9502	0.0000

Appendix 8.1

Proportion of alleles and diversity indices in 12 rice varieties showing within accession variation.

SSR locus	CH† No.	No of alleles	Kaski landraces			Bara landraces			PPB varieties and parental variety			PPB and Parental varieties		
			JB	KG	RA	NS	BAS	LAL	M3	M9	CHH	A200F	ET	KIII
RM246	1	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.00	0.62	1.0
		2	0	0	0	0	0	0	0	0	0	0	0.38	0
RM5	1	1	0.66	1.0	0.88	0.87	0.83	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		2	0.24	0	0.12	0.13	0.17	0	0	0	0	0	0	0
RM48	2	1	0.66	1.00	0.88	0.87	1.00	1.0	0.90	0.70	1.0	1.0	1.0	1.0
		2	0.24	0	0.12	0.13	0	0	0.10	0.30	0	0	0	0
RM213	2	1	1.0	1.0	1.0	1.0	1.0	1.0	0.90	0.80	0.80	0.88	0.62	1.0
		2	0	0	0	0	0	0	0.10	0.20	0.20	0.12	0.38	0
RM232	3	1	1.0	1.0	0.88	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.87	0.88
		2	0	0	0.12	0	0	0	0	0	0	0	0.13	0.12
RM226	4	1	0.66	0.90	1.0	1.0	1.0	1.0	1.0	0.90	0	0.76	1.0	0.66
		2	0.22	0.10	0	0	0	0	0	0.10	1.00	0.24	0	0.24
		3	0.12	0	0	0	0	0	0	0	0	0	0	0
RM164	5	1	1.0	1.0	1.0	1.0	1.0	0.90	1.0	1.0	1.0	1.0	0.63	1.0
		2	0	0	0	0	0	0.10	0	0	0	0	0.37	0
RM3	6	1	1.0	1.0	0.88	0.87	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		2	0	0	0.12	0.13	0	0	0	0	0	0	0	0
RM234	7	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RM248	7	1	1.0	1.0	1.0	0.87	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.88
		2	0	0	0	0.13	0	0	0	0	0	0	0	0.12
RM11	7	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.75	1.0
		2	0	0	0	0	0	0	0	0	0	0	0.25	0
RM223	8	1	0.66	1.0	1.0	0.87	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		2	0.24	0	0	0.13	0	0	0	0	0	0	0	0
RM257	9	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RM242	9	1	1.0	1.0	0.88	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		2	0	0	0.12	0	0	0	0	0	0	0	0	0
RM228	10	1	1.0	1.0	0.88	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.75	1.0
		2	0	0	0.12	0	0	0	0	0	0	0	0.25	0
RM222	10	1	1.0	1.0	0.88	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

		2	0	0	0.12	0	0	0	0	0	0	0	0	0
RM206	11	1	1.0	1.0	0.88	1.0	0.83	1.0	1.0	1.0	0.90	1.0	1.0	1.0
		2	0	0	0.12	0	0.17	0	0	0	0.10	0	0	0
RM21	11	1	1.0	1.0	0.88	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.74	0.88
		2	0	0	0.12	0	0	0	0	0	0	0	0.13	0.12
		3	0	0	0	0	0	0	0	0	0	0	0.13	0
RM17	11	1	1.00	1.00	0.55	0.87	1.00	1.00	1.00	1.00	1.00	1.00	0.87	1.00
		2	0	0	0.33	0.13	0	0	0	0	0	0	0.13	0
		3	0	0	0.12	0	0	0	0	0	0	0	0	0
RM224	11	1	0.66	1.0	1.0	1.0	0.83	1.0	1.0	1.0	1.0	0.88	0.50	1.0
		2	0.24	0	0	0	0.17	0	0	0	0	0.12	0.50	0
RM167	11	1	1.0	1.0	1.0	1.0	0.83	1.0	1.0	1.0	1.0	1.0	1.0	1.0
		2	0	0	0	0	0.17	0	0	0	0	0	0	0
RM229	11	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
RM202	11	1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.80	1.0	0.87	0.88
		2	0	0	0	0	0	0	0	0	0.20	0	0.13	0.12
RM247	12	1	1.0	1.0	0.88	1.0	0.83	1.0	1.0	1.0	0.80	1.0	1.0	1.0
		2	0	0	0.12	0	0.17	0	0	0	0.20	0	0	0
RM20	12	1	1.0	1.0	0.66	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.87	1.0
		2	0	0	0.34	0	0	0	0	0	0	0	0.13	0
H'			1.45	0.15	2.16	1.01	0.99	0.15	0.29	0.62	0.80	0.57	2.64	0.92

JB = Jetho Budho; KG = Kathe Gurdi; RA = Rato Anadi; NS. = Nakhisaro; BAS = Basmati; LAL = Laltenger; M3 = Machhapuchhre 3; M9 = Machhapuchhre 9; CHH = Chhommrong; A200F = Ashoka 200 F; ET = Early tall; and KIII = Kalinga III; † = Chromosome number.