STUDIES ON THE ESTABLISHMENT OF A CORE COLLECTION OF PEARL MILLET (*Pennisetum glaucum*) GERMPLASM

Ranjana Bhattacharjee

Thesis submitted to the
CCS Haryana Agricultural University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN
PLANT BREEDING

COLLEGE OF AGRICULTURE
HISAR 125004, HARYANA, INDIA

2000

To
My Parents
For Their Encouragement and Support

CERTIFICATE I

This is to certify that this thesis entitled, **"*Studies on the establishment of a core collection of pearl millet [Pennisetum glaucum (L.) R. Br.] germplasm*"** submitted for the degree of ***Doctor of Philosophy*** in the subject of Plant Breeding of the Chaudhary Charan Singh Haryana Agricultural University, is a bonafide research work carried out by ***Ms Ranjana Bhattacharjee*** under my supervision and that no part of this thesis has been submitted for any other degree.

The assistance and help received during the course of investigation have been fully acknowledged.

(I.S. Khairwal)
Major Advisor

Paula J. Bramel
Co-Major Advisor

CERTIFICATE II

This is to certify that this thesis entitled, "***Studies on the establishment of a core collection of pearl millet [Pennisetum glaucum (L.) R. Br.] germplasm***" submitted by ***Ms Ranjana Bhattacharjee*** to the Chaudhary Charan Singh Haryana Agricultural University in partial fulfillment of the requirements for the degree of ***Doctor of Philosophy*** in the subject of Plant Breeding has been approved by the Student's Advisory Committee after an oral examination on the same, in collaboration with an External Examiner.

5.1.2021
Major Advisor

External Examiner 5.1.20

Head of the Department 6/1

Dean, Post-Graduate Studies

Acknowledgements

I express my deep sense of gratitude and indebtedness to Dr. I.S. Khairwal, Professor, Bajra Section, Department of Plant Breeding, Chaudhury Charan Singh Haryana Agricultural University, Hisar, for his encouragement and support throughout this course of study. It is with pleasure I express my gratitude and reverence to Dr. Paula J. Bramel, Principal Scientist, Genetic Resources Division, International Crop Research Institute for Semi-Arid Tropics (ICRISAT) for her interest and encouragement that greatly helped to complete this work. I thank the members of advisory committee, Dr. P.S. Kakar, Dr. B.D. Mehta, Dr. V.K. Chowdhury, and Dr. K.P. Singh, CCS HAU, Hisar, for their guidance during this course of study.

I acknowledge with gratitude the support of Dr. B.P.S. Lather, Head, Department of Plant Breeding, all Scientists, Bajra Section, CCS HAU, Hisar and Dr. R.Ortiz, Program Director; Dr. C.T. Hash, Dr. K.N. Rai, Dr. V. Mahalaxmi, Dr. N. Kameshwar Rao, Dr. L.J. Reddy, Dr. H.D. Upadhyaya, Dr. N. Seetharama, and Dr. S. Sivaramkrishnan, ICRISAT, for their encouragement and erudite discussions, that greatly contributed for successful completion of this project.

My special thanks to Dr. Subhash Chandra, Mr. Harikrishna and Mr. Anuj Sharma, Statistical Unit, for their precious discussions and suggestions in aspects related to application of various statistical methods. I also acknowledge Dr. B. Diwakar, Dr. C.L.L. Gowda and staff of Learning Systems Unit, for providing an opportunity to work at ICRISAT.

I am thankful to Mr. K.N. Reddy, Mr. Sube Singh, Ms. Santosh Gurtu, K. Seetha Mr. B. Ashok Kumar, Mr. Prabhakar and Mr. L. Vidyasagar for their cooperation and support during the course of this investigation.

I am grateful to my parents, brothers, Rajesh and Samrat and sister-in-law, Priyanka for their affection, moral support and life giving words that greatly helped me to complete this study. My special thanks to Lava Kumar, Dr. S.K. Dasgupta and family who provided invaluable source of strength, inspiration and blessings during my stay at ICRISAT.

I acknowledge my friends and colleagues, Richa, Maria, Sushmita, Subhanu, Sameena, Kiran, Arun, Azaguel, Naveen and Ranjini for their support and endearing help.

I gratefully acknowledge CCS HAU for providing merit scholarship and ICRISAT awarding research fellowship and funding for this research.

Hisar ***Ranjana Bhattacharjee***

Outline

Chapter 1	Introduction	1-4
Chapter 2	Review of Literature	5-27
Chapter 3	Materials and Methods	28-45
Chapter 4	Results	46-124
Chapter 5	Discussion	125-135
Chapter 6	Summary and Conclusions	136-140
Chapter 7	References	141-154
Chapter 8	Appendix	155-162

CONTENTS

1	**Introduction**	**1-4**
2	**Review of Literature**	**5-27**
2.1.	Core Collection	5
2.1.1.	The concept, definition and importance	5-6
2.1.2.	Stratification of the entire collection	7-8
2.1.3.	Sampling strategies	8-9
2.2.	Establishment of core collection from different germplasm collections	9-19
2.3.	Utilization of core collections	19-22
2.4.	Core collections and molecular genetics	22-25
2.5.	Diversity in pearl millet	26-27
3	**Materials and Methods**	**28-45**
3.1.	Studies on base collection	28
3.2.	Establishment of pearl millet core collection	28-31
3.2.1.	Stratification of the collection	32
3.2.2.	Selection	32
3.3.	Verification of core selection method	32-34
3.4.	Establishment of a minimum core	35
3.5.	Effect of different seasons on characterization of core collection	35-39
3.6.	Diversity assessment by molecular markers	39
3.6.1.	Plant material	39
3.6.2	Solutions required for the RFLP procedure	39-40
3.6.3.	Extraction of genomic DNA	40-41
3.6.4.	Restriction enzyme digestion, electrophoresis, and southern blotting	42
3.6.5.	Labeling of probes	42-43
3.6.6.	Hybridization, washing of blots and autoradiography	44
3.6.7.	Scoring of RFLP bands and data analysis	44-45
4.	**Results**	**46-124**
4.1.	Studies on the base collection	46
4.1.1.	Morpho-agronomic diversity in the world pearl milletcollection maintained at ICRISAT	46-50
4.1.2.	Quantitative characters	50-58
4.1.3.	Qualitative characters	58-59
4.1.4.	Diversity in countries of origin	59-64
4.2.	Development of core collection	64
4.2.1.	Stratification of the collection	64
4.2.2.	Selection of accessions from different clusters and geographical regions	64-68
4.3.	Evaluation of selection method used for the development of core collection	68

4.3.1.	Frequency of accessions from each country of origin	68
4.3.2.	Quantitative characters	68-72
4.3.3.	Qualitative characters	72-86
4.3.4.	Assessment of diversity	86-89
4.4.	Influence of Genotype x Season interaction on characterization of core accessions	89-91
4.5.	Assessment of diversity	91-100
4.6.	Application of molecular markers (RFLP) for diversity assessment	100
4.6.1.	Polymorphism detected by various probe-enzyme combinations	100-122
4.6.2.	Diversity among and within pearl millet accessions at molecular level	123-124
5	**Discussion**	**125-135**
5.1.	Studies on the base collection	125-127
5.2.	Establishment of core collection	127
5.2.1.	Geographical stratification and cluster analysis	127-128
5.2.2.	Evaluation of selection method for the development of core collection	128-131
5.3.	Influence of Genotype x Season interaction on characterization of core collection	131-133
5.4.	Application of molecular markers (RFLP) for diversity assessment	133-135
6	**Summary**	**136-140**
7	**References**	**141-154**
8	**Appendix**	**155-162**
I	Distribution of pearl millet in India and Africa	155
II	Domestication and introduction of pearl millet in Africa and Indian subcontinent	156
III	Preparation of buffers and other chemicals for RFLP procedure	157-160
IV	Southern transfer based on Reed and Mann	161
V	Purification of DNA inserts from plasmid DNA	162

List of Tables

Table No	Description	Page No.
1	Pearl millet collection at ICRISAT.	29
2	Pearl millet accessions selected for core collection studies.	30
3	List of quantitative and qualitative characters studied to evaluate the core.	36-37
4	List of probes selected for RFLP analysis.	43
5	Means and variances for eleven quantitative characters in the world pearl millet collection and different geographical regions.	47-49
6	Mean, variance and range of eleven quantitative characters in the world pearl millet collection and different countries of origin.	51-54
7	Frequency distribution of seven quantitative and five qualitative characters in world pearl millet collection and different geographical regions.	55-56
8	Shannon-Weaver diversity indices for qualitative and quantitative characters in the world pearl millet collection and different geographical regions.	60
9	Shannon-Weaver diversity indices for qualitative and quantitative characters in the world pearl millet collection and different countries of origin.	61-63
10	Range of eleven quantitative characters in world pearl millet collection and twenty-five groups formed by cluster analysis.	65
11	Shannon-Weaver diversity indices for eleven quantitative characters in world pearl millet collection and twenty-five groups formed by cluster analysis.	66
12	Proportion of accessions from different geographical regions and different clusters that form the core sample.	67
13	Representation of pearl millet accessions from different geographical regions in various clusters for the minimum core.	69
14	Proportion of accessions from 25 countries of origin in the world collection of pearl millet and deviations in the core samples drawn with proportional strategy.	70
15	Mean, range and variance for quantitative characters in the world pearl millet collection, core sample and the minimum core.	71
16	Proportion of accessions in different frequency classes of ten quantitative characters in the world pearl millet collection and percentage deviations in the core sample and minimum core.	81-82
17	Proportion of accessions for five qualitative characters in the world pearl millet collection and percentage deviation in the core sample and minimum core.	83
18	Proportion of accessions for major qualitative character	.

	classes in the world pearl millet collection and percentage deviations in the core sample and minimum core.	
19	Proportion of accessions for rare qualitative character classes in the world pearl millet collection and percentage deviation in the core sample and minimum core.	85
20	Comparison of frequency distribution over all quantitative and qualitative characters in the world pearl millet collection, core sample and minimum core.	86
21	Range ratio for quantitative characters in the core sample and minimum core in comparison to the world collection.	87
22	Shannon-Weaver diversity indices for eleven quantitative and five qualitative characters in world pearl millet collection, core sample and minimum core.	88
23	REML variance component analysis for different quantitative characters of pearl millet core accessions.	90
24	REML variance component analysis for different quantitative characters to study the effect of method used to define a core.	91
25	Spearman rank order correlations for quantitative characters among two seasons (kharif and rabi) data and characterization data (data collected in different years from 1974-94).	92
26	Mean, range and variance for seven quantitative characters in the minimum core developed from existing characterization data, core evaluated during *rabi* and core evaluated during *kharif*.	93
27	Range ratio for quantitative characters of the selected minimum core evaluated during two different seasons in comparison to the existing core data.	95
28	Shannon-Weaver diversity indices to compare core data collected in two different seasons and the existing evaluation data.	96
29	Frequency distribution of fourteen qualitative characters along with their diversity indices.	97-98
30	Correlation coefficients among quantitative characters observed in the core accessions in different environments.	99
31	Frequency of RFLP bands in the pearl millet core accessions with different probe-enzyme combinations used in the study.	120
32	Frequency of common and rare RFLP bands among the plants within each accession.	121-122

List of Figures

Figure No.	Description	Page No.
1	Area, production and productivity of pearl millet in the world, Indian subcontinent, West Africa, East Africa, South and Central Africa.	2
2	Schematic representation of the strategy adopted to develop a pearl millet core from the global pearl millet collection.	31
3	Frequency distribution of ten quantitative characters in the entire pearl millet collection.	73-74
4	Frequency deviations of ten quantitative characters in the core sample when compared to entire pearl millet collection.	75-76
5	Frequency deviations of ten quantitative characters in the minimum core when compared to the entire pearl millet collection.	77-78
6	Frequency deviations of ten quantitative characters in the minimum core when compared with the core sample of pearl millet collection.	79-80
7	Representation of polymorphism observed in ten core accessions with different probe-enzyme combinations.	119
8	Ward's dendogram of the core accessions.	124

List of Plates

Plate No.	Description	Page No.
1	Variation for anther color observed in the pearl millet core accessions.	101
2	Variation observed for spike and grain characters within a core accession.	101
3	Variation observed for the presence and absence of bristles on the spikes and bristle length within an accession.	102
4	Variation observed for spike characters within a core accession.	102
5	Variation observed for days to flowering, plant height, number of tillers and canopy spread observed in the pearl millet core accessions.	103
6	Pearl millet core accessions representing variation for days to flowering, plant height, maturity, spike length, presence or absence of bristles and grain colour.	103
7	Variation observed for spike length, shape, thickness and grain colour within an accession.	104
8	Variation observed for spike length, shape, thickness, density and grain colour within an accession.	104
9	Variation observed for spike and grain characters in the pearl millet entire collection.	105
10	Variation observed for spike and grain characters in the pearl millet core collection.	105
11	Pearl millet core accessions evaluated during two different seasons, *rabi* (1998-99) and *kharif* (1999-2000).	106
12	Polymorphism observed in core accession (IP3890) with RFLP probe (PSM 214) and *Dra*I enzyme	107
13	Polymorphism observed in accession (IP4828) with RFLP probe (PSM443) and *Dra*I enzyme.	107
14	Polymorphism observed in accession (IP4542 and IP3122) with RFLP probe (PSM305) and *Dra*I enzyme.	108
15	Polymorphism observed in accession (IP3626 and IP3098) with RFLP probe (PSM305) and *Dra*I enzyme.	109
16	Polymorphism observed in accession (IP3890 and P10394)	110

	with RFLP probe (PSM305) and *Dra*I enzyme.	
17	Polymorphism observed in accession (IP4542 and IP8074) with RFLP probe (PSM443) and *Dra*I enzyme.	111
18	Polymorphism observed in accession (IP3626 and IP4542) with RFLP probe (PSM592) and *Hind*III enzyme.	112
19	Polymorphism observed in accession (IP3122 and IP10394) with RFLP probe (PSM592) and *Hind*III enzyme.	113
20	Polymorphism observed in accession (IP3626 and IP8074) with RFLP probe (PSM716) and *Hind*III enzyme.	114
21	Polymorphism observed in accession (IP8074 and IP4828) with RFLP probe (PSM834) and *Hind*III enzyme	115
22	Polymorphism observed in accession (IP8074 and IP4828) with RFLP probe (PSM857) and *Hind*III enzyme.	116
23	Polymorphism observed in accession (IP4542 and IP3890) with RFLP probe (PSM858) and *Dra*I enzyme.	117
24	Polymorphism observed in accession IP10394 with RFLP probe (PSM858) with *Dra*I enzyme.	118
25	Polymorphism observed in accession IP8276 with RFLP probe (PSM857) and *Hind*III enzyme.	118

List of Abbreviations and Symbols

CIAA	Chloroform : Iso-amyl alcohol
cm	Centimeter(s)
CTAB	Cetyltrimethylammonium bromide
cv	Cultivar
dH_20	Distilled water
dATP	Deoxyadenosine 5'- triphosphate
dCTP	Deoxycytidie 5'–triphosphate
dGTP	Deoxyguanosine 5'-triphosphate
DNA	Deoxyribonucleic acid
ds	Double stranded
dTTP	Deoxythymidine 5'-triphosphate
EAC	East African collection
EDTA	Ethylene diamine tetra acetic acid (disodium salt)
et al.	And others
EtOH	Ethanol
g	Grams
ISC	Indian Subcontinent collection
KCl	Potassium chloride
kg/ha	Kilograms per hectare
LSD	Least significant difference
l	litre(s)
M	Molar
m	Meter
max	Maximum
$MgCl_2$	Magnesium chloride
min	Minute(s)
mg	Milligram(s)
ml	Millilitre(s)
mm	Millimeter(s)
mM	Millimolar
mol. wt.	Molecular weight
NaCl	Sodium chloride
NaOAc	Sodium acetate
ng	Nanogram
NH_4Ac	Ammonium acetate
N-K test	Newman-keuls test
nm	Nanometer
No	Number(s)
NS	Non-significant
nts	Nucleotides
OD	Optical density
PCR	Polymerase chain reaction
PEG	Polyethylene glycol

pH	Hydrogen ion concentration
r	Pearson correlation coefficient
r_s	Spearman rank order correlation
RE	Restriction enzyme
RFLP	Restriction fragment length polymorphism
RNase	Ribonuclease
rpm	Revolutions per minute
SCAC	South and Central African collection
SDI	Shannon-Weaver diversity index
SDW	Sterile distilled water
SE	Standard error
S.Em	Standard error of means
ss	Single stranded
SDS	Sodium dodecyl sulphate
sec	Seconds
spp	Species
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
Tris	Tris (hydroxymethyl) amino methane
U	Unit(s)
UV	Ultraviolet
V	Volts
v	Volume
W	Watts
WAC	West African collection
WC	World collection

1. Introduction

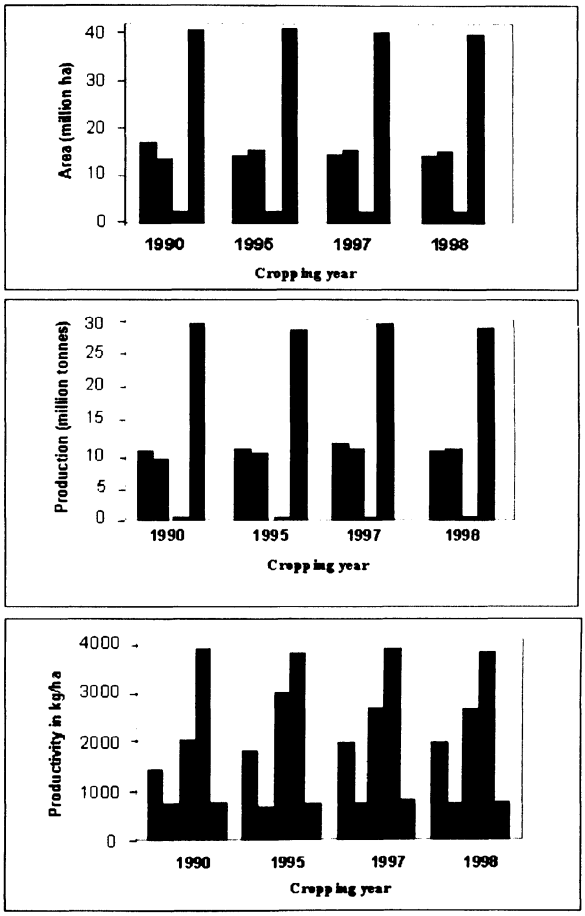
1. INTRODUCTION

Pearl millet (*Pennisetum glaucum* (L.) R. Br.), commonly known as bajra, is the sixth most important cereal and forage crop of the arid and semi-arid tropics of South Asia and Sub-Saharan Africa. It has been used as a cereal crop for nearly 3000 years in Africa and parts of the Near East and is grown on about 24.2 million ha in many countries of southern, eastern western, and central Africa, and a few countries of Asia, particularly India, and in some parts of the Americas, and Australia. India is a major pearl millet producing country with 43.4 per cent of the world area and 42 per cent of pearl millet world production. Five states, Rajasthan, Maharashtra, Gujarat, Uttar Pradesh and Haryana account for nearly 90 per cent of the total cultivated area under pearl millet (10.4 mha). Generally, pearl millet is considered more efficient crop for arid and semi-arid conditions because of its potential in utilization of soil moisture and higher level of heat tolerance than sorghum and maize (Harinarayana *et al.*, 1999).

The genus *Pennisetum* consists of more than 140 species, of which four (*P typhoides*, *P. spicatum*, *P glaucum* and *P. americanum)* are cultivated, 16 wild annual and remaining 120 are wild-perennial species. Pearl millet is a C_4 diploid species with basic chromosome number x = 7. Because of its allogamous nature, pearl millet accessions are highly heterogeneous reflecting high variability within and among the accessions. Protogyny and the time lag between stigma emergence and anther dehiscence favour complete cross-pollination leading to the greatest morphological diversity. The plant features include a diverse range of plant height, time to flowering, tillering, stem thickness, fodder and grain quality, high growth rate and adaptability to varied agro-ecological conditions.

Because of diffused belt of origin, stretching from Senegal to western Sudan, pearl millet was domesticated by humans at multiple locations (Harlan and de Wet, 1971; see

Fig 1. Area, production and productivity of pearl millet in the world, Indian subcontinent, West Africa, East Africa, South and Central Africa



Appendix). Adaptation to varied agro-ecological environments, gene flow from wild progenitor(s), and local selection by farmers pearl millet emerged into to various ecotypes during post-domestication process. The crop further diversified into several ecotypes during its cultivation under diverse agro-climatic conditions, and in differing farming systems (Murty *et al.*, 1967; Brunken *et al.*, 1977; Anand Kumar and Appa Rao, 1987; Anand Kumar and Andrews, 1993; Khairwal *et al.*, 1990, Appa Rao and de Wet, 1999, Appa Rao, 1999).

The rapid spread of high yielding pearl millet cultivars, crop substitution with more remunerative crops, recurrent drought, and urbanization led to gradual erosion of pearl millet landraces. National and international gene banks acquired and conserved pearl millet genetic resources to avoid extinction of native germplasm. Such preservation activities resulted in the assemblage of 21,392 accessions of pearl millet, from over 50 countries at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India and many more at National Agricultural Research Systems (NARS). The collection includes primitive landraces, improved cultivars, breeding stocks, mutants and wild species of genus *Pennisetum,* from all the pearl millet growing countries of the world. However, because of the large collection size precise evaluation of these accessions is a big problem. Furthermore, some of these accessions were poorly described impeding the potential utilization of many of these accessions for crop improvement purpose. Therefore, evaluation of accessions for specific traits is essential for effective utilization of pearl millet germplasm.

Frankel and Brown (1984) introduced the concept of "core collection" for better utilization of large collection of the germplasm. Accessions in core collection derived from total collection represent the genetic diversity of a crop species and its relatives with minimum repetitiveness and are ecologically and genetically distinct from one another. The remaining accessions termed, as a "reserve collection", will be conserved for future needs. However, to derive a core, the accessions in the total collection should have taxonomic integrity (Brown and Spillane, 1999). The Core collection has several advantages in the management and utilization of genetic resources in various germplasm operations. The core serves as a guide for acquisition of new accessions of highest priority

for conservation; the accessions in the core in general represent the whole range of variation in the base collection for evaluation of desirable and complex traits in breeding programs and further, because of the small collection size, the whole process of gene bank management will be less expensive.

Various methods have been adopted for developing a core collection. These mainly depend on the genetic structure and reproductive pattern of crop species, pattern of acquisition of germplasm, priorities and objectives of curators (Brown, 1989a). The ultimate aim of any collection is conservation and evaluation of crop germplasm resources for effective utilization.

In order to deal with world collection of pearl millet and its effective utilization, the establishment of a core collection is one of the ideal approache. Further, an assessment of genetic diversity based on morpho-agronomic characters only might be biased because distinct morphotypes can result from a few mutations and share a common genetic background. Therefore, molecular markers are essential for explaining whether existing genetic variability, assessed by measuring morpho-agronomic characters, is related to genetic diversity. This may be assessed by measuring allelic frequencies using molecular markers. Thus, the present study was undertaken with an objective to establish a pearl millet core collection that represents a complete range of genetic diversity of the world collection and to achieve the following objectives:

- To compare multivariate vs. hierarchical classification for determining core collections that maximizes diversity.

- To evaluate the adequacy of mean characterization data to determine core collection and do diversity assessment in a cross-pollinated crop.

- To determine the extent of diversity within and between accessions of established core collection using morphological traits and molecular markers.

2. Review of Literature

2. REVIEW OF LITERATURE

2.1. Core collection

Most of the germplasm collections of crop and pasture plants consist of large number of entries. These collections were meant for preserving the genetic diversity of the particular species for future needs. The use of genetic resources in collections and its access is important for its effective utilization. However, large collection size and heterogeneous genetic structure hinders the efforts to increase the use of gene bank material for crop improvement. Core collections play an important role in solving these problems.

An attempt has been made in this chapter to present a brief review of literature related to different aspects of work undertaken in this investigation.

2.1.1. The Concept, Definition and Importance

Before the core concept, stress has been given on collecting as many resources as possible for securing their survival in the gene banks, which is a continuous process. Paradoxically, the large numbers of accessions accumulated in the gene banks are often poorly described. Their use for breeding purpose could be greatly increased if more information on the amount and kind of variation in these collections is available. However, in most cases, information to characterize thousands of entries is not available. Recognizing this, Frankel (1984) proposed that a collection could be sampled to a core collection, which would 'represent with a minimum of repetitiveness, the genetic diversity of a crop species and its relatives.' The accessions or entries excluded from the sample would be retained as the reserve collection. The proposal was further developed by Frankel and Brown (1984) and Brown (1989a), who outlined how to achieve the coverage of the collection in the core by using information on the origin and characteristics of the accessions. Consequently, the core concept made a considerable impact on plant genetic resources tasks. Brown (1989a,b) argued that for extensive use of germplasm collections, there should be better, rationalised, refined collections, structured as a small, well-defined and representative

sample, called core collection. Here, the intention is not to replace the existing germplasm collections but to identify a sample of accessions, that would represent the genetic diversity present in the whole collection and make this variation accessible to users. Brown (1989a) also suggested the hierarchical relationship between the core and reserve collection so that the core will be a 'first look' at germplasm collection for researchers to identify quickly the desired trait (s). The core collection will then provide the researchers with a manageable number of accessions to look for desired characters, for detailed evaluation and for application of new techniques.

Based on the neutral allele theory of Kimura and Crow (1964) for finite populations, Brown (1989a) simulated different scenarios for different numbers and frequencies of alleles at each locus variously distributed within a collection. According to this analysis, Brown (1989a) proposed that a core should contain about 10% of the whole collection when the collection does not contain the total genetic variability for the species and a maximum of 3000 accessions for a complete collection. This should result in about a 0.85 probability of including 80% of the alleles that occur in the whole collection. A slightly different approach has been developed by Crossa *et al.* (1993) for cross-pollinated crops. They proposed the use of probability models and determined optimal sample sizes with 95 percent probability of including at least one copy of alleles with a given frequency. For example, if there are 50 loci with four alleles each, 156 individuals are required to retain at least one copy of alleles with 95 percent probability and with a frequency of 0.05.

Mackay (1995) pointed that, the potential use of core collection fall into three main groups, namely, (a) plant breeders requiring quick identification of desirable traits and immediate access to seed samples, (b) germplasm specialists attempting to study genetic diversity and (c) curators who require some knowledge of this diversity to meet the needs of breeders and for improving the efficiency of germplasm management. He also pointed that there are at least two objectives for establishing a core collection. Firstly, to assist gene bank management and secondly, to contribute to the effective and efficient use of available germplasm. These two objectives are tightly linked, since good gene bank management contributes to better germplasm utilization.

2.1.2. Stratification of the entire collection

The clustering of accessions based on multivariate statistical techniques have been suggested by Spagnoletti-Zeuli and Qualset (1987) to form groups, the members of which are likely to be genetically similar Accordingly, the accessions of USDA durum wheat were divided into groups based on their country of origin Some countries with contrasting agroecological conditions grouped together The results argued that country of origin is not a reliable basis for stratification However, Brown (1989a) proposed hierarchical stratification as the most effective strategy for developing a core collection The hierarchical stratification groups the accessions based on the characters shared commonly, before sampling In another study with a large set of data for barley, Peeters and Martinelli (1989) reached a contrary conclusion that country of origin is a reliable general indicator of diversity, even when no passport data are available

Perry and McIntosh (1991) used canonical discriminant analysis to study the variation in USDA soybean germplasm collection based on morphological traits Cluster analysis of the canonical means classified the collection into four regional groups Hence, they suggested the use of geographical patterns of variation for grouping of accessions Further, Peeters *et al* (1993) and Holbrook *et al* (1993) assumed that phenotypic similarity for a limited set of characters is a better indicator of genetic and ecological similarity than country of origin

Therefore, for effective stratification of the collections, reliable characterization data along with passport data with information on country of origin should be available (Spagnoletti-Zeuli and Qualset, 1993, Diwan *et al* , 1994, Mathews and Ambrose, 1994, Hamon *et al* , 1995) to make the grouping of accessions based on phenotypic similarity and their country of origin more effective Combined use of agro-ecological and characterization for stratification was used to establish the CIAT *Phaseolus vulgaris* core collection (Tohme *et al* , 1995) Similarly, Hintum (1995) advocated the use of hierarchical approach for splitting the entire collection into smaller and smaller groups within groups In his study, using a set of Chinese barley landraces with reliable passport data, stratification based on collection site was compared with stratification based on qualitative and quantitative characters The collection site data proved best for clustering, followed by

qualitative data. Clustering based on quantitative data did not improve sampling strategies.

When a complete data set for a germplasm collection is available, principal component analysis and cluster analysis are excellent tools for grouping of accessions by degree of similarity (Basigulp *et al.*,1995). Similar study was carried out by Bisht *et al.* (1999) to stratify the germplasm accessions of Indian sesame into diversity groups, based on well-defined passport and characterization data.

2.1.3. Sampling strategies

Once the collection is divided into distinct groups, the questions that need to be addressed are the number of entries to be chosen from each group and the method of choice within the group (Yonezawa *et al.*, 1995).

Several sampling methods to select entries have been suggested, ranging from random sampling to stratified sampling based on known groups with sample size either constant, logarithmic or proportional to the group size (Brown, 1987; Brown, 1989b; Erskine and Muehlbauer, 1991). In cases of larger groups where there are chances of redundancy, representation in proportion to the logarithm of group size is more conservative (Brown, 1989b). Based on the studies on barley and *Glycine*, he concluded that for smaller groups with less chances of redundancy, representation in proportion of group size is more appropriate. Similarly, Erskine and Muehlbauer (1991) found proportional strategy to be better than taking a constant number from each group while developing a core collection in lentil germplasm. In a different study on composing a core collection of cultivated barley collected in China, Hintum *et al.* (1995) found that proportional and logarithmic allocation scored better than constant representation.

Schoen and Brown (1995) developed two new sampling strategies for constructing diverse and representative core collections in cases of unequal diversity and differentiation among accessions occurring mostly in wild crop relatives. These strategies are 'H' (heterozygosity) and 'M' (maximisation) strategies, which refers to genetic index. A comparison between these two strategies and other strategies proposed by Brown (1989b) showed that, the average ranking of the six strategies (highest rank = 1, lowest rank = 6)

for expected allele retention was M (2.1) > H (3.0) > P (3.1) > L (3.7) > C (4.1) > R (4.7).

Several other studies have been made to evaluate different sampling strategies for establishing a core collection in different crops like durum wheat (Spagnoletti-Zeuli and Qualset, 1993), Indian sesame (Bisht *et al.*, 1999), perennial ryegrass and *Medicago truncatula* (Balfourier *et al.*, 1998), sesame (Hodgkin *et al.*, 1999) etc.

2.2. Establishment of core collection from different germplasm collections

Setting up a core collection implies to facilitate the management of germplasm, reduce the cost of conservation, and promote the diffusion and use of genetic resources. From the time this concept has been developed, there has been extensive awareness in the development of core collections. As a result, many researchers and projects have been involved in developing core collections for major crop species and also their wild relatives. There are also broad differences in the aspects adopted for developing core collections by different researchers for example, different strategies for grouping of accessions, sampling strategies to be used, etc. which are discussed below.

The first core collection established was that of perennial *Glycine* by Brown *et al.*, 1987. The core consisted of 111 accessions developed from a collection of 1400 accessions of twelve different species of *Glycine* at Canberra, Australia. Grouping was made initially, at species level and ecogeographic factors were used to select the entries for the core. The proportion of accessions selected from each category (species x state) varied from less than 5 per cent to 100 per cent in order to achieve some representation. The core so developed included at least a few accessions of each species and geographic coverage of each Australian State. Preference was given to the accessions already used in research and included known morphological, cytological and isozyme variation within species.

Further in 1988, a core collection of wheat was developed by Moody *et al.*, with a specific objective of estimating variation for tolerance to soil boron toxicity in the Australian wheat (*Triticum* spp.) collection. Geographical and soil data was used to select genotypes. This approach of developing core collections with primary objectives promoted better evaluation, utilization and understanding of genetic variation.

Brown (1989a) then studied the effect of sampling strategies on core diversity while working with *Glycine tomentella* and *Hordeum vulgare*. He found stratified sampling in logarithm or absolute proportion to the group size as the best strategy. He further recommended that core entries should not be bulked, and the core should be subjected to evolution as new materials are received or better data is obtained.

Hamon and van Sloten (1989) established a core collection from 2283 okra (*Abelmoschus esculentus*) collection in the joint ORSTOM/IBPGR project, using passport and characterization data of west African accessions. The core collection of 189 accessions was developed primarily to have a manageable collection scaled down to the needs of the breeder including the widest possible range of variability. The data on quantitative characters like date of flowering, plant height and number of internodes and qualitative descriptors like stem color, leaf shape and fruit position was analyzed using univariate and multivariate methods to determine the correlations between characters and geographical distribution of variability.

The structure of variation in 67 barley (*Hordeum vulgare*) landraces from Syria and Jordan was investigated by using various spike characters from the landraces grown in a favorable environment in Syria (Weltzien, 1989). Factor analysis was used to identify trait complexes that accounted for major proportions of the total variation among landraces. The landraces were then clustered into nine distinct groups based on their similarity for all traits as well as geographical similarity of the collection. The samples were then proportionally chosen from each cluster to form the core.

In 1990, Hamon and Noirot proposed a procedure for obtaining a core collection of okra (*Abelmoschus* spp.) using quantitative plant characterization data. They found that the selection of accessions based on multivariate analysis maximized the variation in the core in comparison to the original collection.

To develop a core collection, Mackay (1990) used specific traits like pasting quality of bread wheat, heat shock on bread quality and the spectral quality of anthocyanin

across different altitudes, in a range of *Triticum* spp. accessions, including wild and cultivated diploids, tetraploids and hexaploids. He used ecogeographical data for selection of accessions to include a rational amount of genetic variation in the core.

The European cooperative program for conservation and exchange of crop genetic resources nominated a working group, to report a proposal on the development of a core collection of *Hordeum* spp. (von Bothmer, *et al.*, 1990). The strategy of the proposal was to include less than 2000 accessions consisting of cultivars, landraces, wild species and genetic stocks. The characters included for creating dendrogram of variation were growth habit, ear type and pedigree data for cultivars, ecogeographical data and agricultural systems practiced for landraces while for wild *Hordeum* spp., the use of ecogeographical data was specifically mentioned. The inclusion of genetic and cytogenetic marker stocks in the core was also recommended to increase the quality of the core for further research.

Vaughan (1991) used the evaluation database of world rice (*Oryza sativa*) collection, conserved at the International Rice Research Institute (IRRI), to compare different sampling strategies for choosing the core. A frequent source of resistance to different rice pests like brown planthopper, green leafhopper and whitebacked planthopper, and diseases, bacterial blight and rice blast was captured in the core. The results indicated stratified selection as more reliable than random and sequential selection, in retention of these characters.

Holbrook *et al.* (1993) selected a core collection from the U.S germplasm collection of peanut. Data on peanut collection was obtained from the Germplasm Resources Information Network (GRIN), which included country of origin and observations on plant type, pod type, seed size, testa color, number of seed per pod, and average seed weight. The entire collection consisting of 7432 accessions was first stratified by country of origin and morphological data was analyzed using multivariate methods. The accessions were hierarchically clustered into genetically similar groups. Ten per cent sample was selected randomly from each group to develop the core collection of 831 accessions. The results of this study indicated the representativeness of the core to the entire collection with the preservation of genetic variation for the traits.

From a collection of 3000 durum wheat accessions, selection of a core sample was done by applying five different strategies. The core was further evaluated using four qualitative and eight quantitative spike characters (Spagnoletti-Zeuli and Qualset, 1993). Each strategies (random, random systematic according to chronology of entries into the collection, stratified by country of origin, stratified by log frequency by country of origin and stratified by canonical variables) generated about 500 accessions for the core sample. All the strategies produced representative sample from the whole collection, however stratified canonical sample increased phenotypic variances. They suggested multivariate approach as extremely useful though requires extensive data from the whole collection.

Crossa *et al.* (1993) developed maize (*Zea mays*) core collection, by sub dividing the entire germplasm into non-overlapping groups, based on racial complex and selected ecogeographical criteria. Within each race, accessions were stratified by region. Cluster analysis on morphological and agronomic characters identified groups of similar accessions. Similarly, the practical importance for maintaining germplasm in maize was reported by Crossa *et al.* (1994) taking an example of race *Tuxpeno*, for which 848 accessions were available from 23 races and 3 sub races. The selection of 175 accessions based on cluster analysis, principal component analysis, ecogeographical data, lodging and adaptation data assessed in multi-location trials was used to develop the core.

Radovic and Jalovac (1994) suggested a different approach to identify the divergent population for the selection of a core sample. Hierarchy was established using the combining ability of the testers. Seven percent of the 902 maize population at Yugoslav maize gene bank formed the core consisting of 72 entries. Eighteen morphological traits were used to determine the differences between the entire population and the core, and it was found for only one trait.

Balfourier and Charmet (1994) established a core collection from 547 natural populations of perennial rye grass (*Lolium perenne*). The core consisting of 42 entries was developed on the basis of multivariate analysis, to preserve the diversity considering agronomic characteristics and ecogeographical stratification.

A core collection of 211 accessions for annual *Medicago* species was developed by evaluating 1240 accessions (40% of the total collection) based on 16 agronomic and morphological traits (Diwan *et al.* 1994). The collection maintained at U.S. National Plant germplasm System consisted of 36 species of this crop, which was used to develop the core. Accessions within species were grouped by cluster analysis based on unweighted pair group method with arithmetic averages. One accession per cluster was selected for each species to represent the greatest diversity in geographical regions. The selected core was reevaluated and found to represent the variability of the germplasm collection maintaining the stability between evaluation years.

Mathews and Ambrose (1994) selected a core collection of 157 accessions (6 percent of John Innes *Pisum* collection) using passport data, cross reference with morphological data and descriptive statistics of groups formed, based on species and subspecies, ecotypes and landraces, cultivars, and genetic stocks.

A different procedure was used by Hamon *et al.* (1995) to develop a core collection for coffee germplasm. The genetic organization of the coffee gene pool was critically investigated to develop the core according to genetic history and available genetic knowledge. A new method, Principal Components Score Strategy, was used on data obtained for the cultivated *Coffea liberica*, a well-studied species, for developing a core collection using quantitative data. The use of several different strategies, rather than a single one, to establish a core collection was suggested, depending upon the gene pool as well as the level of knowledge. Further Noirot *et al.* (1995) studied the importance of principal component scoring for sampling stratification and choice of sample size. They recommended the use of quantitative data, but with small changes could be also used for qualitative data, to effectively maximise the sample diversity.

In another study, Basigulp *et al.* (1995) developed a core collection of 200 entries from the U.S. collection of perennial alfalfa (*Medicago sativa*) plant introductions. The sample was taken from 1100 plant introductions collected from 47 countries, based on passport data from the Germplasm Resources Information Network (GRIN) system. Eight

methods were compared for developing the core by non-parametric tests. Two methods, combined cluster analysis based on principal component within each cluster and direct selection of entries within each geographical group, were found best for designating the core collection, which retained the greatest variability for all traits.

Cordeiro *et al.* (1995) suggested the use of different criteria for construction of a Brazilian core collection of cassava (*Manihot esculanta* Crantz). These included stratification of the collection into groups based on category (landraces or improved materials), origin (grouping accessions according to agroecological classification of origins) and characters of importance to breeders (selected according to heritability and reliability of data).

A core collection was established at ICRISAT by stratifying the world sorghum collection of 33100 accessions geographically and taxonomically into sub groups (Prasada Rao and Ramanatha Rao, 1995). Accessions in each subgroup were then clustered into closely related groups based on characterization data, using principal component analysis. A proportional sampling strategy resulted in a core collection of 3475 accessions (approximately 10 per cent of the entire collection).

Diwan *et al.* (1995) developed a core collection for U.S national germplasm of annual *Medicago* spp., comparing eleven methods differing in their use of passport and evaluation data. The core developed by cluster analysis based on evaluation data represented the collection better than the core designed solely based on passport data. Further, random logarithmic selection of accessions generated better core collection than proportional method. In contrary to other studies, 5-10 percent sample size for core collection was found to be insufficient to represent the entire collection.

A barley core collection of 2000 accessions was established by Knupffer and van Hintum (1995) from entire *Hordeum* gene pool, consisting of landraces, improved cultivars and wild relatives from the primary (*H. vulgare* and diploid *H. spontaneum* complex), secondary (*H. bulbosum*) and tertiary gene pools (including 30 different species of *Hordeum*). The core collection was a collaborative international effort that included

germplasm maintained at several institutes, representing the entire gene pool of the crop. Brown (1995) termed this kind of core as 'synthetic core' wherein the accessions are assembled from various collaborative germplasm collections, which is more an aid to germplasm use than to gene bank management.

A core sample was established by Tohme *et al.* (1995) from 24000 accessions, available from the world common bean (*Phaseolus vulgaris* L.) collection, maintained at the Centro Internacional de Agricultura Tropical (CIAT). A three-step process was followed, firstly by prioritizing the regions giving greater weight to traditional bean growing areas. Secondly, germplasm was classified according to agroecological origin. The third step was based on morpho-physiological data of growth habit, seed color and seed size. The environmental parameters like length of growing season, photoperiod, soil type and moisture regime yielded 54 agro-ecological classes. A random selection within these agro-ecological classes resulted in a core collection of 1000 accessions identified from primary centers, an additional 300 from secondary centers along with 40 cultivars, 40 bred lines and 40 genetic stocks.

Grauke and Thompson (1995) evaluated pecan [*Carya illinoinensis* (Wangenh.) K. Koch] germplasm collection to designate a core subset of 26 cultivars using stratified sampling procedures. Cultivars were selected with reference to geographic origin.

The effect of different methods on the representativeness of selected accessions for the core was examined by Galwey (1995), using data from Cambridge (UK) *Phaseolus* bean germplasm collection. Generally, different selection strategies yielded very similar outcomes in term of representativeness and diversity. However, the use of passport, taxonomic and ecogeographical data helped to ensure that the core collection is representative of the whole collection as well as the whole plant taxon under consideration.

A hierarchical clustering procedure with incremental sum of squares as the fusion strategy and Euclidean distance as the dissimilarity measure was used by Crossa *et al.* (1995), for classifying 175 maize accessions of the *Tuxpeno* race complex. Multivariate methods were used to study the phenotypic diversity and to select a core collection.

Accessions were randomly selected from each group formed by cluster analysis. The core sample consisted of 48 accessions.

The development of one or many core collection has been discussed by Mackay (1995) to meet non-specific requests made by researchers. He also discussed the advantages of using an attribute based core collection to address individual requirements of germplasm user and effectiveness with which a collection can be subsampled for a specific range of genetic variation.

Mahajan *et al.* (1996) proposed a technique to establish a representative core set of South Asian okra (*Abelmoschus esculentus*) germplasm collection, maintained at the National Bureau of Plant Genetic Resources (NBPGR), using characterization data. A total of 260 representative accessions with diverse geographical background was selected for the study. A non-hierarchical cluster analysis was performed using Euclidean distance based on nine quantitative descriptors. The incremental sum of square was used as clustering criterion. Principal component analysis for quantitative data was used to select the accessions from groups. Shannon-Weaver diversity index for qualitative data was also used to design a core sample of 53 accessions.

A core collection was developed by Pederson *et al.* (1996) to determine how representative the core subset was of the entire U.S white clover germplasm collection for cyanogenesis. This confers resistance to many species of leaf feeding insects and molluscs. A simple technique of geographic stratification and random selection was used to develop the core that was representative for total cyanogenesis and distribution of cyanogenic plants in the entire collection.

To examine the performance of different cluster strategies, to propose a classification method for germplasm accessions and to verify the usefulness of the classification methods in forming a core subset, different hierarchical algorithms like UPGMA, centroid, median, Ward's method and density search methods were used by Franco *et al.* (1997) to classify Mexican maize accessions. Normix density search method followed by Ward's method of classification was most appropriate strategy for grouping

accessions into relatively homogeneous groups.

Taba *et al.* (1998) evaluated Caribbean maize (*Zea mays* L.) accessions from the CIMMYT maize germplasm bank to design a core subset based on morphological and agronomic characters. The sequential strategy proposed earlier by Franco *et al.* (1997) was used to stratify the entire collection. A pattern of phenotypic diversity of the accessions in each cluster was also established by canonical discriminant analysis. The selection of entries for the core was based on yield (mg/ha), ear rot (%), and moisture (%), calculated for each accession. The upper 20 per cent of the accessions representing the phenotypic diversity within clusters with high selection indices were chosen for the subset.

A procedure for establishing a sesame core collection was systematically studied by Xiurong *et al.* (1999) representing over 4000 accessions and genetic diversity for the crop in China. Different hierarchical clustering procedures were used with data on 14 traits, grouping the accessions into seven ecotypes. 453 randomly selected accessions formed the core collection with good representativeness, confirmed by zymogram characteristics.

In another study that involved diversity in the Indian sesame collection (3129 accessions) representing all ecogeographical regions, was carried out by Bisht *et al.* (1999), for a range of morphological and agronomic characters. The accessions were classified into seven discrete clusters. Simple random sampling along with six stratified random sampling strategies, were compared for varying sample fractions ranging from 5-30 per cent. Stratified random sampling was superior to simple random sampling and a 10 per cant sample fraction was invariably determined as a suitable fraction in forming the core.

Balfourier *et al.* (1999) compared different sampling strategies for developing a core sample in large sized natural populations of perennial ryegrass (*Lolium perenne* L.) and medic (*Medicago truncatula* Gaertn.), for their ability to restore the spatial or geographic structure of the initial collection, their capacity to capture the phenotypic diversity of the whole collection and their effect on conservation of neutral alleles. The comparisons of two species found spatial structure of diversity to represent the best core

collection, although there was a minimal effect on the mean number of neutral alleles.

Several methods of developing core collections using agronomic data for *Poa pratensis* L. was studied by Johnson *et al.* (1999). From a total collection of 228 accessions, a core collection representing 10 per cent of the collection was developed using random sampling, hierarchical cluster analysis (Ward's or UPGMA), and stratification by broad geographic regions using agronomic data. The cores developed from cluster analysis resulted in increased variances and ranges of agronomic traits, turf quality and seed production, compared to the core developed without cluster analysis.

A method was developed by Skinner *et al.* (1999) to develop a core collection by maximising the diversity, measured as mean Euclidean distance, from within groups of accessions defined by species, sub species and geographic origin. The effectiveness of the method was tested on a collection of annual *Medicago* evaluated at the Australian *Medicago* Resource Center in Adelaide, South Australia. A core collection of 1705 accessions (10.4 per cent) was obtained by grouping sequentially, such that species formed groups first followed by subspecies within each species, and geographical origin within each species-subspecies group. The core represented 74 per cent of the extreme values of 27 characters, covering entire range in most cases.

Holbrook (1999) developed a core collection for the U.S. *Arachis hypogaea* germplasm collection consisting over 7000 accessions with great amount of genetic diversity. The collection was first stratified by country of origin and then grouped into nine sets based on morphological characters using multivariate approach. The random selection procedure resulted in a core sample of 831 accessions representing the entire range of genetic variation present in the original collection.

A Peruvian sweetpotato core collection was developed by Huaman *et al.* (1999) on the basis of morphological, ecogeographical, and disease and pest reaction data. The entire collection of sweetpotato (5000 cultivated accessions) maintained at the International Potato Center (CIP) was grouped into distinct clusters following Unweighted Pair-Group Method using an Arithmetic average (UPGMA) based on the above morphological

descriptors. The sampling was based on square root of the number of accessions for each cluster, which resulted in a core collection consisting of 85 accessions. This sampling was further validated by partial assessment of the core for resistance to diseases and pests, tolerance to salt, storage root dry matter content and vegetative period.

Several variation in the adoption of approaches and procedures has been apparent from various core collections that have been developed so far. The basic idea in each of these studies was to make a better use of the germplasm, selecting a set of accessions, with different objectives in mind mostly to represent the broad genetic variation available for a total crop genome or species, a geographical region and entire germplasm maintained by a gene bank. The collection may be based on single international or national gene bank germplasm as in winter wheat, *Glycine* and okra or may involve collaborative international effort as in barley and rice. Though the objective in most of the studies was to have a representative sample but the methodology used varied significantly, as does the emphasis given to different kinds of information or the relative size of the resulting cores. In most studies, ecogeographical data was used extensively, while in okra importance was given to characterization and evaluation data.

The common features observed in most of the studies were the need for a germplasm reference collection, stratification of the collection following hierarchical or non-hierarchical strategies, multivariate or principal component analysis and a sampling of 5-10 percent from the defined groups following different sampling strategies as described by Brown (1989a), so that the core sample represent the genetic spectrum in the whole collection.

2.3. Utilization of core collections

Once a core has been established the first question that will be of concern is the extent to which it achieves its aims. Several efforts have been made in this regard to utilize the core sample for specific objectives.

Charmet *et al.* (1993) evaluated the core collection of French perennial ryegrass (*Lolium perenne*) for agronomic traits at different locations of France. There was

significant environmental interaction for most of the characters. Mapping of regression coefficient to identify populations tolerant to specific factors of environment by plant breeders was suggested.

The genetic diversity within a core collection of *durum* wheat was estimated by Spagnoletti-Zeuli and Qualset (1995) to select potential exotic lines for breeding programs. They described the use of information obtained from genetic evaluation of a core collection to exploit the whole collection more critically.

The peanut (*Arachis hypogeae* L.) core collection was evaluated by Holbrook and Anderson in 1995 to identify resistance to late leaf spot. 61 leaf spot resistant accessions were identified in the core that represented 54 percent of the resistant accessions in entire collection. The efficiency of the core collection in improving the peanut germplasm evaluations was well established.

Isleib *et al.* (1995) screened the peanut core collection for resistance to *Cylindrocladium* Black Rot (CBR) and early leaf spot. In a greenhouse screening trial, 11 core accessions had greater resistance to CBR than the resistant check and 12 early maturing accessions had a similar level of resistance to early leaf spot.

The use of core collection with particular reference to rice was discussed by Vaughan and Jackson (1995). The core collection for wild species consisted of resistant sources to the tungro virus complex. Out of 208 accessions for 19 species tested, 15 accessions of four species were found resistant to one of the two forms of this virus and had very low infection by rice tungro spherical virus.

Anderson *et al.* (1996) screened the peanut core collection for resistance to tomato spotted wilt virus and found tremendous variability among plant introductions for reaction to natural tomato spotted wilt virus epidemics. Twenty-seven accessions were reported to exhibit significantly greater resistance than the susceptible check. One of the accession was found to have lower disease incidence in comparison to the check cultivar. The accessions with different maturity levels and growth habits were also found to be resistant to this

disease. Further, Holbrook *et al.* (1997) examined the peanut core collection for reaction to peanut root knot nematode. Thirty-six core accessions showed a reduction in root galling, egg-mass rating, egg count per root system, and egg count per gram of root compared to the resistant check.

Ellis *et al.* (1998) evaluated the core collection of *Brassica oleracea* accessions for resistance to the cabbage aphid, *Brevicoryne brassicae*. Four hundred and ten accessions representing the genetic and geographic diversity of the species were evaluated, of which 98 accessions were identified as moderately resistant or partially resistant. The most promising gene pool was from kale, where a higher than expected resistant accessions were found.

Forty accessions, forming a core collection of bush type of common bean (*Phaseolus vulgaris* L.) germplasm in the Netherlands, were evaluated for 14 qualitative and quantitative traits at the Agricultural University, Wageningen (WAU). An additional 117 Dutch accessions, collected in private home gardens, were also evaluated for morphological and agronomic traits at the International Center for Tropical Agriculture (CIAT), Cali, Columbia. Multivariate and principal component analysis at both WAU and CIAT indicated existence of one large group (31 accessions) with no discernable patterns among collections of landraces, garden forms and cultivars (Zeven *et al.*, 1999).

Miklas *et al.* (1999) used a sub-sample of the core collection of *Phaseolus vulgaris* accessions representing the active USDA National Plant Germplasm System collection of 1698 accessions from Central and South America to identify new sources of resistance to the fungal pathogen, *Sclerotinia sclerotiorum*, causing white mold disease. Eleven of 89 core accessions were identified as potential sources of physiological resistance to white mold. The resistant accessions detected in the core collection guided them to expand their search for resistance in the larger active collection. The results indicated that 57 per cent of the active collection had significantly higher levels of resistance.

The most agronomically acceptable portion of the core collection of peanut was evaluated by Franke *et al.* (1999) for resistance to *Rhizoctonia* limb rot. This subset of the

core collection consisted of 66 accessions having a spreading growth habit. Six core accessions had a high level of resistance to this disease.

All these reports on use of core collection for some specific purpose provide information on the efficiency of developing a representative sample of the total collection in evaluating a wide array of genetic diversity for resistance to complex diseases and tolerance to adverse environmental conditions. Once the core accessions are identified for a desirable character, an expanded search for that particular character can be taken up among accessions in the active collection.

2.4. Core collections and molecular genetics

Molecular markers have provided a new tool for breeders and gene bank managers to search for new variation and to investigate genetic factors controlling quantitatively inherited traits. Before the introduction of RFLPs (Restriction Fragment Length Polymorphism) and PCR (Polymerase Chain Reaction)-based markers, isozymes were widely used in inter and intra specific diversity studies. The molecular markers were extensively used in the studies of complete gene pool, taxonomy, centers of diversity of crop species, path of domestication, and the relation between diversity and environment.

The use of morphological data for larger samples, molecular data for smaller samples and studies that combine both was discussed by Hillis (1987). The areas where molecular techniques can be extensively used for germplasm characterization are identification of genotypes including duplicate accessions, fingerprinting of genotypes, analysis of genetic diversity in the collections, and developing a core collection (Dodds and Watanbe, 1990). Brown (1990) specified the advantages of isozymes in designing sampling strategies for in depth diversity analysis.

Marker based analysis of morphological and agronomic traits to determine the genetic diversity in cultivated common bean (*Phaseolus vulgaris* L.) was studied by Singh *et al.* (1991). A multivariate analysis, using phaseolin and allozymes as initial classification criterion, followed by analysis of morpho-agronomic traits, suggested the existence of subgroups within each major Andean and Mesoamerican groups, with distinct

morphology, adaptation, and disease resistance. BR 62775

Beyermann *et al.* (1992) discussed the benefit of DNA fingerprinting with specific oligonucleotide probes for simple repetitive DNA sequences in germplasm characterization. Similarly, Pradhan *et al.* (1992) concluded that molecular and taxonomic classifications are incompatible while studying the phylogeny of *Brassica* and allied genera based on variation in chloroplast and mitochondrial DNA patterns.

For characterizing the genetic diversity among flint and dent lines of maize (*Zea mays* L.) germplasm, RFLP markers were used by Messmer *et al.* (1992). Principal coordinate analysis of genetic similarity estimates resulted in separate grouping of flint and dent lines. In a similar way, the genetic basis of diversity and differentiation in the *indica* and *japonica* groups of the cultivated rice (*Oryza sativa* L.) was studied by assaying DNA with RFLP markers (Zhang *et al.*, 1992). The results demonstrated *indica* rice as genetically more diverse than *japonica* type.

Twenty-one celery (*Apium graveolens* L. vanidulce) accessions were classified into three distinct groups based on RAPD (Random Amplified Polymorphic DNA) marker differences (Yang and Quiros, 1993). In a different study for comparing isozymes and RFLP (Restriction Fragment Length Polymorphism), to analysis diversity in wild barley (*Hordeum vulgare* ssp *spontaneum*), Zhang *et al.* (1993) concluded that isozymes demonstrated a larger amount of diversity within population while RFLP resolved higher proportion of variation among populations with greater heterozygosity than isozymes.

Lux and Hammer (1994) suggested the use of PCR-based assays for discrimination of higher genetic differences among accessions to define a hierarchical structure. The importance of DNA fingerprinting in distinguishing closely related forms to eliminate the duplicates was mentioned. In another study, forty-eight cultivars from the European barley (*Hordeum vulgare* L.) germplasm were assayed for RFLP markers to determine the genetic similarity estimates within and between winter and spring germplasm (Melchinger *et al.*, 1994). Seven probe-enzyme combinations could distinguish closely related lines.

The genetic diversity in sorghum was demonstrated by Vierling *et al.* (1994) with RFLP and RAPD markers. Patterns of RFLP markers were proposed as estimators of genetic diversity among a set of nine oat lines (Moser and Lee, 1994). These results support sufficient variation with RFLP markers and proposed for further application of this technology in oats.

Gepts (1995) considered that genetic diversity in germplasm collections are best characterized by the use of molecular and biochemical markers. This information was further used in the selection of a core collection, which was more representative of the entire germplasm. In another study, 12 rice (*Oryza sativa* L.) accessions of IRRI germplasm collection were evaluated by Virk *et al.* (1995) using RAPD marker. They demonstrated the validity of RAPD data for the study of diversity within germplasm collections.

The genetic structure between and within gene pools of a core collection of wild *Phaseolus vulgaris* L. was studied by Tohme *et al.* (1996) using Amplified Fragment Restriction polymorphism (AFLP) technique. This generated a large amount of data in a short time than had been possible with other methods.

In order to compare the potential of isozyme and RFLP markers to investigate genetic diversity within and among populations, ten maize populations were characterized for 20 isozyme loci and RFLP for 35 probe-enzyme combinations (Dubreuil and Charcosset, 1998). The superiority of RFLP markers over isozymes was clear for identifying the origin of a given individual and revealing a relevant genetic structure among populations.

Casas *et al.* (1998) characterized the genetic variation among a representative set of 37 barley cultivars currently grown in Spain, using RFLP markers. Genetic distances based on RFLP band patterns showed moderate consistency with coancestry coefficients, based on pedigree records.

RFLP markers were used to study the genetic diversity among 51 accessions

representing 29 *Paspalum* species (Jarret *et al.* 1998). Species affinities based on RFLP data were in close agreement with previously determined relationships based on both morphological and cytological characteristics. A different study was made on Mexican common bean to compare the core and reserve collection using RAPD markers and characterization data (Skroch *et al.* (1998). The enrichment of diversity was not detected in the core but the genetic structure was representative of the reserve collection at CIAT.

The combined use of microsatellite, isozyme and AFLP profiles for rapid identification of potential duplicate accessions in the cassava core collection was studied by Chavarriaga-Aguirre *et al.* (1999). In another study, the genetic diversity of East Asian accessions of the barley core collection was surveyed using six isozyme loci to provide information for the further development of an optimal core collection in barley (Liu *et al.*, 1999). Genetic variation was found in both cultivars and landraces.

In a recent study, Huaman *et al.* (2000) investigated the genetic structure in both entire collection and respective core subset of *Solanum tuberosum* ssp. *andigena* (potato cultivars) using nine isozyme marker that have been genetically characterized. The allozyme frequency distribution for each locus was tested for homogeneity between the entire and core collection by Chi-square tests. The most frequent allozymes in the entire collection also showed highest frequencies in the core collection. This study revealed that the sampling strategy followed to develop the core collection of tetraploid Andean potato cultivars was adequate to capture a representative sample of the entire collection.

All the reported studies on the use of various molecular markers in germplasm management shows their efficiency in characterization of germplasm, analysis of genetic diversity and to distinguish among closely related forms to eliminate duplicates. However, ideal markers are those that are codominant, numerous, easy to use and non-specific. Codominant molecular markers provide more detailed information on the genetic structure of existing diversity within groups. Thus, the potential of RFLP markers that are codominant in nature, was also demonstrated in various studies, in assessing inter and intra specific variation. Further, these markers would prove beneficial in adjusting the core collection for elimination of duplicates, or inclusion of variability that was absent in the

original core.

2.5. Diversity in pearl millet

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is widely grown in arid to semi-arid regions of Africa and Indian subcontinent. The crop is particularly adapted to Sahelian West Africa where landraces have evolved in different ecological niches. These landraces have accumulated inter-population diversity that has not been characterized. Evaluation of genetic diversity is a pre-requisite for successful germplasm exploitation (Anand Kumar and Appa Rao, 1987).

Wilson *et al.* (1990) evaluated the diversity among pearl millet landraces collected in Central Burkina Faso using morphological characteristics. Ward's cluster analysis grouped the landraces into ten clusters of phenotypic diversity. Some correspondence was found between the geographic collection sites of landraces and their inclusion in particular clusters. Analysis of morphologic and disease resistance traits was useful in identifying the duplicates within the collection.

Diversity among African Pearl millet landrace population was studied by Ouendeba *et al.* (1995) using morphological and disease resistance characters. In the pooled analysis, all landraces were significantly different for one or more of the characters evaluated. Ward's cluster analysis and principal component analysis were used to investigate the nature and degree of divergence in the landrace populations. The landraces from Niger showed very less variation than the other African landrace populations.

Most of the pearl millet growing areas in Central African Republic (CAR) was explored by ICRISAT in collaboration with the Ministry of Rural Development, CAR, Bangui. Considerable diversity was observed for several characters. Cluster analysis categorized accessions into similarity groups facilitating sampling a limited number of entries to represent the diversity of pearl millet germplasm from CAR (Reddy *et al.*, 1996). Out of 589 pearl millet germplasm accessions from Sudan assembled at ICRISAT, 581 were evaluated for morphological and agronomic characters by Reddy *et al.* (1996) for diversity assessment.

Pearl millet germplasm from Cameroon was collected jointly by ICRISAT and the National Cereals Research and Extension Project (NCRE)/ Institute of Agronomic Research (IAR), Maroua, Cameroon (Appa Rao *et al.*, 1996). Millet germplasm from Cameroon was found to be a good source for more number of productive tillers, large compact spikes, and larger ivory and cream colored grain besides its potential for forage.

3. Materials & Methods

3. Materials and Methods

3.1. Studies on base collection

Pearl millet collection at ICRISAT constitutes 21,392 accessions from 50 countries. The characterization data consisting of information on twelve quantitative and eight qualitative characters (IBPGR and ICRISAT, 1993) for most of these accessions are available in the Genetic Resource Division, ICRISAT (Table 1). The data recorded by evaluating the accessions in batches of 1000-2000, during *rabi* (post-rainy) and *kharif* (rainy) seasons from 1974-94 in alfisol (red soil) at ICRISAT, Patancheru, located at 17°25'N latitude and 78°E longitude. Of the 21,392 accessions, 20,642 are cultivated types.

In the present study, 16,063 cultivated accessions were selected after eliminating the uncharacterized germplasm (Table 2). Most of these accessions consisted of landraces and a few breeding lines. Since information on breeding lines (especially their origin) is limited, a few lines were included to avoid redundancies in the final core collection. Passport data pertaining to the exact site (longitude and latitude) of collection is also lacking for most of these accessions, but country of origin is known. Based on geographical origin 16,063 pearl millet accessions were classified into 4 groups; the Indian subcontinent collection (ISC; 5585 accessions), East African collection (EAC; 951 accessions), South and Central African collection (SCAC; 4377 accessions), and West African collection (WAC; 5150 accessions). Accessions from each of these geographical regions were used for further analysis to establish a pearl millet core that represents the diversity of the entire collection (Table 2).

3.2. Establishing the pearl millet core collection

The methodology adopted to establish a core collection of pearl millet is schematically represented in Fig 2.

Table 1. Pearl millet collection at ICRISAT*

Country	No. of accessions
AFRICA	
1. Algeria	5
2. Benin	46
3. Botswana	82
4. Burkina Faso	868
5. Cameroon	1001
6. Cape Verde Islands	2
7. Central African Republic	156
8. Chad	134
9. Congo	8
10. Ethiopia	3
11. Gambia	15
12. Ghana	283
13. Kenya	98
14. Lesotho	4
15. Malawi	311
16. Mali	1177
17. Mauritania	36
18. Morocco	4
19. Mozambique	33
20. Namibia	1128
21. Niger	1269
22. Nigeria	1913
23. Senegal	413
24. Sierra Leone	60
25. Somalia	4
26. South Africa	165
27. Sudan	614
28. Tanzania	508
29. Togo	515
30. Tunisia	6
31. Uganda	120
32. Zaire	14
33. Zambia	163
34. Zimbabwe	1395
Total African accessions 12,553	

Country	No. of accessions
ASIA	
1. India	7940
2. Korea (South)	1
3. Lebanon	108
4. Myanmar	10
5. Pakistan	169
6. Russia & CISs	16
7. Sri Lanka	2
8. Turkey	2
9. Yemen	293
Total Asian accessions 8541	
EUROPE	
1. France	11
2. Germany	3
3. UK	32
Total European accessions 46	
AMERICAS	
1. Brazil	2
2. Mexico	11
3. USA	231
Total American accessions 244	
OCEANIC	
Australia	8
Total number of pearl millet accessions at ICRISAT 21,392	

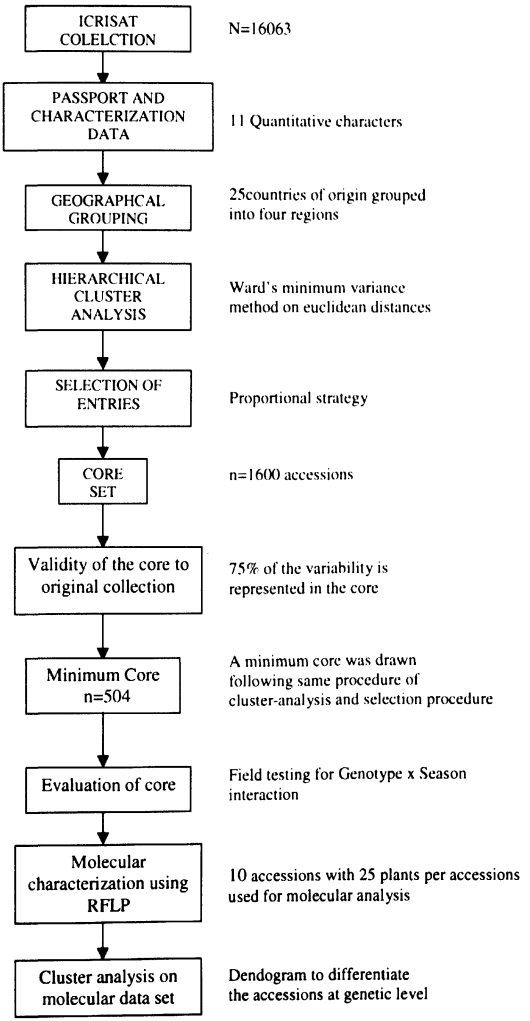
**As on 31st December, 1997*

Table 2. Pearl millet accessions selected for core collection studies

	Grouping based on geographical origin	Grouping based on country of origin	No. of accessions
I	Indian subcontinent collection	1. India	5373
		2. Pakistan	151
		3. Yemen	61
II	East African collection	4. Kenya	75
		5. Malawi	292
		6. Sudan	584
III	South and Central African collection	7. Botswana	82
		8. Cameroon	918
		9. Central Africa	146
		10. Chad	38
		11. Namibia	1052
		12. South Africa	138
		13. Tanzania	467
		14. Uganda	86
		15. Zambia	92
		16. Zimbabwe	1354
IV	West African collection	17. Benin	46
		18. Burkina Faso	841
		19. Ghana	283
		20. Mali	1024
		21. Niger	904
		22. Nigeria	1120
		23. Senegal	349
		24. Sierra Leone	59
		25. Togo	515
Total			**16063**

FIG 2

SCHEMATIC PRESENTATION OF DEVELOPMENT OF A CORE SAMPLE FROM WORLD COLLECTION OF PEARL MILLET



3.2.1. Stratification of the collection

The 16063 pearl millet accessions were further classified into 25 groups based on the country of origin (Table 2). A hierarchical cluster analysis was performed on mean values of the eleven standardized quantitative variables, namely, 1. days to flowering (*kharif*); 2. days to flowering (*rabi*); 3. plant height (*kharif*); 4. plant height (*rabi*); 5 number of productive tillers; 6. spike exertion; 7. spike length (*kharif*); 8. spike length (*rabi*); 9. spike thickness (*kharif*); 10. spike thickness (*rabi*) and 11. 1000-grain weight. The Ward's (1963) minimum variance method was used as per the PROC-CLUSTER program in SAS (SAS Institute, 1989a). The method computes matrix of Euclidean distances among group means and produce a dendrogram depicting successive fusion of individuals, where all the individuals of the same group form a cluster. The distance between two cluster is the ANOVA sum of square of the distances between two clusters added over all the variables and expressed as semipartial R^2 i.e. the sum of square between clusters just joined is divided by the corrected sum of square (Romesburg, 1984). The number of groups in cluster analysis depends on the size of collection, the intended size of the core and the dissimilarity of the groups at the lowest level of sorting.

3.2.2. Selection

Following proportional sampling procedure, 10% accessions were selected from each cluster formed by multivariate analysis and country of origin (Brown, 1989b). The proportional allocation incorporates alleles with lower variance than simple random sampling.

3.3. Verification of the core selection method

The selected pearl millet accessions for core collection should retain full range of variation and resemble the pattern of world collection. These features of the core collection and its representation were assessed by various statistical methods as discussed below:

The frequency distributions observed for country of origin in the selected sample was compared to the expected frequency distribution from the world collection using chi-square test as per the method of Spagnoletti-Zeuli and Qualset (1993). The Chi-square value was calculated as:

$$\text{Chi-square } (\chi^2) = \sum_{i=1}^{n} (O_i - E_i)^2 / E_i$$

where, O_i = Observed frequency, E_i = Expected frequency

Mean values of eleven quantitative characters of the core was compared with that of world collection using Newman-Keuls test for post-hoc comparison of means. Similarly, variances for all these characters were compared by Levene's test of homogeneity of variances (SAS, 1989a).

The frequency distribution (expressed in per cent) for all the quantitative and qualitative characters for the core was compared with the expected frequency distribution (expressed in per cent) in the world collection (Spagnoletti-Zeuli and Qualset, 1993). The frequency deviations for these characters in the core sample were computed. The deviations for country of origin were also compared by Chi-square analysis.

Range ratio was calculated to determine the proportion of range retained in the selected core from the world collection. The range ratio is the average ratio of range of the sample over all characters to the range of the original collection. It was calculated using the equation:

$$\text{Range ratio} = \sum_{i=1}^{n} (RC_i / RO_i) / n$$

where, RC_i = Range for the i^{th} character in the sample, RO_i = Range of the i^{th} character in the original collection and n = Number of characters compared. Sample was considered to be representative of the original collection if the percentage of range retained by the core was at least 70% of the range of the original collection (Diwan *et al.*, 1995).

Because of the categorical nature of the qualitative characters, a non-parametric statistical procedure was used to compare the distributions in core collection and the

original collection. Mann-Whitney test is a two-sample t-test that involves the states of two samples to be compared. It was calculated using the equation:

$$t = \frac{(\bar{X} - \bar{Y})\ [mn\,(N-2)/N]^{1/2}}{[\Sigma\,(X_i - \bar{X})^2 + \Sigma\,(Y_i - \bar{Y})^2]^{1/2}}$$

where, X and Y are the states of two different samples with N-2 degrees of freedom, m is the number of individuals in the population, n is the number of individuals in the sample and N is the number of individuals in the population and sample.

Global diversity was estimated by the Shannon-Weaver Diversity Index (SDI; Hutchenson, 1970). The equation for SDI is:

$$SDI = (-\sum_{i=1}^{n} P_i \times \log_e P_i) / \log_e n$$

where, n = number of phenotypic classes for a trait, P_i = proportion of the total number of entries in the i^{th} class and $\log_e$ = logarithmic of the exponential.

The phenotypic diversity is based upon the frequency distributions for the different classes of qualitative characters. This index was also estimated for the seven quantitative characters after transforming each character into four phenotypic classes as four quartiles of the world collection. Shannon-Weaver indices obtained for each character were pooled for the core and entire collection. Means and standard errors were then calculated over all the characters separately for the core and compared with the means of world collection diversity. The distribution of diversity in the core was also compared with that of original collection for all the traits using chi-square test.

3.4. Establishment of a minimum core

Following the same procedure of cluster analysis and sampling strategy described in previous sections, a minimum core consisting of 504 accessions was established from the core sample. The minimum core was verified using statistical methods that were used to validate the core sample (given in section 3.3) for determining its representation in world collection (16,063 accessions) and core sample (1600 accessions).

3.5. Effect of different seasons on characterization of core collection

The minimum core accessions were grown on alfisol (red soil) at ICRISAT, Patancheru, India during 1998-99 *rabi* (post-rainy) and 1999-2000 *kharif* (rainy) seasons to study within and between variability, and the influence of environment on characterization. The accessions were sown in an augmented design (Federer, 1961). Each accession was sown in two rows of 5 m long ridge with a spacing of 75 cm between rows and 10 cm between plants. HHB67, a short duration pearl millet hybrid was sown at an interval of eight accessions as a 'check line'. Recommended package of practices was followed to raise the crop and irrigation was given at regular intervals during *rabi* season.

Observations on 60 plants for each accession were recorded on seven quantitative and 14 qualitative traits during *rabi* and seven quantitative traits during *kharif* as per the descriptors for pearl millet (Table 3; IBPGR and ICRISAT, 1993). Observation on early vigor was taken on 18 days old seedlings. Anther colour was recorded before anther dehiscence. For the observation on days to flowering, stigma emergence was considered. The data on characters like plant height, node, sheath and blade pigmentation, spike length, thickness, shape and exertion, bristle length, and number of productive tillers were recorded at dough stage. Yield potential was recorded considering spike number, size and density, seed number and size compared to a standard control and fodder yield potential by considering tillering ability, leafiness and bulk at the time of flowering. The observations on node and sheath pubescence were taken at flowering and head emergence, respectively. Spike density was recorded at maturity stage of the plant. The observations on seeds like seed color, 1000-seed weight and seed shape was recorded after threshing on mature dried seeds. Observations on HHB 67 from 10 random blocks were recorded on the basis of 60 plants for each of the characters described above.

Table 3: List of Quantitative and qualitative characters considered to evaluate the core

	CHARACTER	DESCRIPTION	CLASSIFICATION
Quantitative characters			
1	Days to flowering (*kharif* and *rabi*)	Number of days from field emergence to 50% of the plants flower	Early flowering – Accessions flowering before 50 days Late flowering – Accessions flowering after 120 days
2	Plant height (*kharif* and *rabi*) cm	Height of the plant from ground level to the tip of the spike at dough stage	Short – Mean height <100cm Tall – Mean height >350 cm
3	Number of productive tillers	Number of spikes that bear panicles at dough stage	Poor tillering – Mean < 3 tillers Good tillering – Mean >10 tillers
4	Spike exertion cm	The length between the ligule of the flag leaf and the base of the spike of primary tiller, measured at dough stage	Poor exertion – Mean exertion < 0 cm Good exertion – Mean exertion >20 cm
5	Spike length (*kharif* and *rabi*) cm	Length of the spike measured at dough stage	Short spike – Mean spike length <20 cm Long spike – Mean spike length >100 cm
6	Spike thickness (*kharif* and *rabi*) mm	Diameter of the spike, measured at maximum thickness, excluding bristles	Thin spike – Mean thickness <10 mm Thick spike – Mean thickness >40 mm
7	Thousand grain weight g	Thousand grain weight in grams at 12% moisture level	Small grain – Mean grain weight <5 g Large grain – Mean grain weight >15 g
II. Qualitative characters			
1	Early vigour	Seedling growth after 18 days of emergence	3 = Low 5 = Intermediate 7 = High
2	Anther color	Colour of the anther	Y = Yellow P = Purple C = Cream
3	Node pigmentation	Node pigmentation at dough stage	NP = Pigmented (purple) NNP = Non-pigmented (green)
4	Blade pigmentation	Pigmentation of blade at dough stage	BP = Pigmented (purple) BNP = Non-pigmented (green)
5	Sheath pigmentation	Pigmentation of sheath at dough stage	SP = Pigmented (purple) SNP = Non-pigmented (green)
6	Node pubescence	Presence or absence of hairs on the nodes during flowering	NH = Hairy NNH = Non-hairy

		stage	
7	Sheath pubescence	Presence or absence of hairs on the sheath during head emergence	SH = Hairy SNH = Non-hairy
8	Spike shape	Shape of the spike	1 = Cylindrical; 2 = Conical 3 = Spindle; 4 = Club 5 = Candle; 6 = Dumb-bell 7 = Lanceolate; 8 = Oblanceolate 9 = Globose
9	Spike Density	Density of the spike recorded at the maturity	3 = Loose 5 = Intermediate; 7 = Compact
10	Bristle length	Bristle length recorded at dough stage	3 = short (bristles below the level of the apex of seed) 5 = Medium (bristle length between 0 and 2 cm above the seed) 7 = Long (bristles longer than 2 cm above the seed)
11	Grain shape	Shape of the grain	1 = Obovate; 2 = Oblanceolate 3 = Elliptical; 4 = hexagonal 5 = globular
12	Grain color	Grain colour recorded after threshing	1 = Ivory (yellow-white group) 2 = Eream (orange-white group) 3 = Yellow (yellow group) 4 = Grey (grey group) 5 = Deep grey (black group) 6 = Grey brown (brown group) 7 = Brown (brown group) 8 = Purple (purple group) 9 = Purplish black.
13	Yield potential	Number of spikes, their size and density, number of seeds and their size compared with a standard check	3 = Low 5 = Intermediate 7 = High.
14	Fodder yield potential	Tillering ability, leafiness and bulk at flowering stage	3 = Poor 5 = intermediate 7 = Good

The data obtained from *rabi* and *kharif* seasons were analyzed separately by augmented design analysis to determine the block effects. Descriptive statistics for the seven quantitative characters for two seasons, such as mean, variance and range were calculated and compared with the world collection. Levene's test of homogeneity of variances and Newman-Keul's test for post-hoc comparison of means were also calculated for the core sample as well as world collection. Student's t-test and F-test was done to further test the homogeneity in means and variances for these characters. Comparison of frequency distribution through Chi-square test applied on per cent frequency for qualitative characters was calculated for the minimum core, core sample as well as world collection.

To study the effect of environment on stability of accessions of the core sample, the morphological data collected over two seasons, was analysed by estimating the variances due to main effects of seasons, accessions and the interaction effect of accession x season in a linear mixed model using REML (Residual Maximum Likelihood) analysis.

Spearman rank order correlations were calculated between the two seasons data and also between core evaluation data and data obtained during regeneration from 1974-94 for all the characters to confirm the stability of the characters over different seasons. Spearman's coefficients of rank correlation applies in the form of ranks and calculated as:

$$r_s = 1 - \frac{6 \Sigma_i d_i^2}{(n-1)\, n\, (n+1)}$$

where d_i = difference for the i^{th} pair and n = number of differences. The estimate was tested by comparing with the Student's 't' test with (n-2) degrees of freedom (Steel and Torrie, 1980).

Pearson correlation coefficients between different characters were also estimated for regeneration and core evaluation data and it is calculated as:

$$r_{12} = \Sigma (y_{11} - \bar{y}_1) \times (y_{12} - \bar{y}_2) / [\Sigma (y_{i1} - \bar{y}_1)^2 \times \Sigma (y_{i2} - \bar{y}_2)^2]^{1/2}$$

Range ratio was calculated for the core sample data recorded during *kharif* to determine the proportion of range retained. Shannon-Weaver diversity indices were also calculated for all the characters for both season data and compared to the diversity of the existing evaluation data.

3.6. Diversity assessment using RFLP markers

3.6.1. Plant material

Ten highly diverse pearl millet accessions (IP4542, IP10394, IP3890, IP4828, IP3122, IP3626, IP3098, IP8276, IP13608, IP8074) of Indian origin were selected through cluster analysis from a set of core accessions using quantitative characters following random selection within each cluster. These accessions were sown in the field in a two-row plot with 50 seeds per row. The plants were selfed to obtain pure seed from 50 randomly selected plants in each accession. A random sample of about 50 seeds per accession was sown in a pot (11.5 cm diameter) in low-nutrient peat and sharp sand compost in greenhouse. Leaves from 50 seedlings within each accession were harvested at 3-5 leaf stage (13-15 days old plants), frozen in liquid nitrogen and stored at –80°C till analyzed.

3.6.2. Solutions required for the RFLP procedure

1. Liquid nitrogen
2. 1 M Tris. Chloride (121.1 gm of Tris. EDTA in 800 ml of distilled water. pH adjusted to 8.5 with conc. HCl)
3. 5 M NaCl (292.2 gm of NaCl in 750 ml of water)
4. 0.5 M EDTA (186.1 gm of $Na_2.EDTA.2H_2O$ in 800 ml of water. pH adjusted to 8.0 with NaOH pellets)
5. 20% SDS (200 gm of SDS in 1 L of warm water
6. S Buffer (100 mM Tris-HCl, 100 mM NaCl, 50 mM EDTA, 2% SDS)
7. 10 mg/ml proteinase K (100 mg of protienase K dissolved in 10 ml of water)
8. 100x TE (121.1 gm of Tris. and 37.2 gm of $Na_2.EDTA.2H_2O$ in 800 ml of water. pH adjusted to 8.0 with conc. HCl)

9. Phenol
10. Chloroform and isoamyl alcohol in 25:1 ratio
11. Phenol-Chloroform (Equal volume of phenol and Chloroform + isoamyl alcohol)
12. Isopropanol
13. 70% Ethanol (700 ml of 100% ethanol in 300 ml of water)
14. 10 mg/ml RNase (100 mg of RNase in 10 ml of water)
15. 3 M Sodium acetate (408.24 gm of Na. acetate.$3H_2O$ in 600 ml of water. pH adjusted to 5.2 with glacial acetic acid)
16. Kesara's loading buffer (0.10 gm of bromophenol blue, 0.10 gm of xylene cyanol, 10 ml of glycerol and 0.372 gm of $Na_2EDTA.2H_2O$, volume adjusted to 20 ml by adding 1x TAE)
17. 50x TAE (242 gm of Tris base in 500 ml of water, 100 ml of 0.5 M EDTA with pH = 8.0 and 57.1 ml of glacial acetic acid)
18. 0.25 M HCl (10.75 ml of conc. HCl in 500 ml of water)
19. 4 M NaOH (96 gm of NaOH pellets in 2 L of water)
20. 20x SSC (877 gm of NaCl and 441 gm of sodium citrate in 4 L of water)
21. 5x HSB (175.3 gm of NaCl, 30.3 gm of PIPES and 7.45 gm of $Na_2EDTA.2H_2O$ in 800 ml of water. Adjust pH to 6.8 with 4 M NaOH)
22. Denhardt's III (2 gm of gelatin, 2 gm of ficoll, 2 gm of PVP-360, 10 gm of SDS and 5 gm of sodium pyrophosphate in 100 ml of water)
23. Salmon's DNA (5 gm of Salmon testes DNA in 1 L of water)
24. Stripping solution (50 ml of 20% SDS and 10 ml of 20x SSC in 1940 ml of water)
25. Wash solution I (200 ml of 20x SSC and 100 ml of 20% SDS in 1700 ml of water)
26. Wash solution II (20 ml of 20x SSC and 100 ml of 20% SDS)

3.6.3 Extraction of DNA

Total genomic DNA was isolated following the method of Sharp *et al.* (1988) with minor modifications.

1. Five g of frozen leaf sample was ground to fine powder in Waring blender and transferred into a 50 ml polypropylene tube.

2. To this 15 ml of S-buffer and 100 µl of Proteinase-K from stock solution was added and incubated at 65°C for 30 min., followed by incubation at 55°C for 1.5 h with gentle shaking.
3. The samples were then extracted with equal volumes of phenol-chloroform-isoamyl alcohol (24:24:1v/v) mixture and the emulsion was separated by centrifugation at 5,000 rpm for 20 min at 4°C in a Sorvall HB7 rotor.
4. The upper aqueous phase was reextracted with equal volumes of chloroform-isoamyl alcohol (24:1v/v) and emulsion separated by centrifugation at 5,000 rpm for 20 min.
5. To the aqueous phase equal volumes of isopropanal was added and gently mixed and held at -20°C for 30 min.
6. Precipitated DNA was spooled with a hook and dissolved in 2 ml of RNase-$T_{50}E_{10}$ buffer and incubated overnight at room temperature.
7. To inactivate RNase, samples were reextracted with equal volumes of phenol-chloroform as described in steps 3 and 4.
8. To the aqueous phase 2.5 volumes of ethanol was added and held at -20°C for 1 h to precipitate DNA.
9. DNA was spooled and washed with 70% ethanol. DNA pellets were air dried before dissolving in 300 µl of $T_{10}E_1$ and stored at 4°C.
10. DNA concentration was estimated based on the spectrophotometer measurement of UV absorption at 260 nm, assuming 1 OD at 260 nm is equal to 50 µg of DNA (Maniatis *et al.*, 1982). The ratio of OD_{260} to OD_{280} was calculated to check the purity of DNA sample. Pure DNA preparation shows an OD_{260} to OD_{280} ratio ideally between 1.7 and 1.8 (Maniatis *et al.*, 1982).
11. DNA was analyzed in 0.8% TAE-agarose gel to test the integrity as described in Maniatis *et al.* (1982). Gels were stained with ethidium bromide and viewed on UV-transilluminator before photographing with camera fitted with UV filter.

3.6.4. Restriction enzyme digestion, electrophoresis, and southern blotting

1. Twenty µg of DNA was digested with each of the following restriction enzymes: *Dra*I, *Hind*III, *Eco*RI, and *Eco*RV as per the supplier's instructions (Amersham pharmasia biotech ltd.).
2. Digested DNA was separated by electrophoresis in a 0.8% TAE-agarose gel for 12 h at 40 v/cm in a BioRad DNA Sub Cell™ electrophoresis unit (Owl Separation Systems, Model No.A-1). A *Hind*III-digested Lambda DNA (λ DNA) was used as molecular size marker.
3. The agarose gels were stained in 0.5 µg/ml ethidium bromide and viewed on an UV-transilluminator.
4. Separated DNA fragments from agarose gel were transferred onto Nylon Membrane (Hybond-N+, Amersham pharmasia biotech ltd.) following the procedure of Reed and Mann (1985) (given in Appendix).
5. The transferred membranes were soaked in 2x SSC for 60 sec to neutralize the alkali and air dried and wrapped with cling film and stored at -20°C.

3.6.5. Labeling of probes

Variation in the plant material was studied using 14 DNA probes derived from a pearl millet *Pst*I genomic library (see Table 4; Sharp *et al.*, 1988). These probes were selected so as to spread over all through the seven linkage groups of pearl millet. For this purpose DNA obtained from the select plant material was digested with restriction endonulceases and probed with the selected markers. Details are given below.

Random-primed method of Feinberg and Vogelstein (1983) was used for labeling DNA with ^{32}P. Purified insert DNA (25-50 ng) was denatured by heating at 95°C for 7 minutes, quenched on ice for 5 min and the labeling reaction mixture was added and incubated at 37°C for 3 hours.

[Labeling reaction: 5 µl of oligo-labelling buffer (Amersham pharmasia biotech ltd.), 2 µl equimolar concentrations of dCTP, dGTP and dTTP, 2 µl acetylated BSA (concentration of 10µg/µl), 2-3 µl of 50 uCi ^{32}P-dATP and 2 units of Klenow enzyme]

Table 4: List of probes selected

Probes	Linkage Group	Restriction Enzyme
Downy Mildew Resistance		
Psm 858	I	*Dra*I
Psm 565	I	*Hind*III
Psm 716	II	*Hind*III/ *Dra*I
Psm 305	II	*Hind*III/*Dra*I
Psm 648	II	*Dra*I
Psm 713	VI	*Eco*RI
Psm 618	VII	*Dra*I
Psm 834	VII	*Hind*III
Psm 857	VII	*Hind*III
Drought tolerance		
Psm 214	II	*Dra*I
Psm 025	II	*Eco*RV
Psm 321	II	*Eco*RI
Psm 592	II	*Hind*III
Psm 443	II	*Dra*I

3.6.6. Hybridization, washing of blots and autoradiography

1. Hybridization was performed in a Techne Hybridizer (HB-1D).
2. Membranes were prehybridization at 65°C with 5 ml of prehybridization solution (3 ml of 5x HSB, 1.5 ml of denatured salmon sperm DNA and 1.5 ml of Denhardt's solution and sterile distilled water to 15 ml) C for 6 h for new blots and 1 hour for stripped filters.
3. Labelled probe was added into the prehybridization mixture and incubated at 65°C for 16 h or overnight.
4. The blots were washed four times, each wash for 15 min as shown below.
    - Wash 1 at room temperature with 100 ml of 20x SSC, 25 ml of 20% SDS and distilled water to 1 liter
    - Wash 2 at room temperature with 100 ml of 20x SSC, 25 ml of 20% SDS and distilled water to 1 liter.
    - Wash 3 at 65°C with 10 ml of 20x SSC, 25 ml of 20% SDS and distilled water to 1 liter.
    - Wash 4 at 65°C with 10 ml of 20x SSC, 25 ml of 20% SDS and distilled water to 1 liter.
5. Membranes were dried and wrapped in cling film and exposed to photographic film (KODAK, X-Omat™, XK-5) at –70°C using KODAK intensifying screens in a cassette. Depending on the counts, exposure time was set.
6. The photographic film were developed with KODAK developer for 2 minutes, followed by a stop bath (running tap water) treatment for 1 min, fixed with KODAK fixer for 2 min, washed in running tap water and air-dried. The autoradiograms were photographed using KODAK 100 ASA color films.

3.6.7. Scoring of RFLP bands and data analysis

The bands obtained from RFLP procedure were scored as 1 and 0 for presence and absence of bands, respectively and the data was presented in the form of n x n matrix, where 'n' is the number of accessions. The frequency of polymorphic bands, common bands and rare bands produced were estimated. The distance was measured by per cent disagreement

(GENSTAT Release 4.03) for each pair wise comparison, as the data was categorical in nature. The distance was computed as:

$$\text{Distance } (X, Y) = (\text{number of } X_i \neq Y_i) / i$$

The distance matrix was used to construct dendrogram (Statsoft, 1997) using agglomeration method of hierarchical clustering techniques (Sneath and Sokal, 1973).

4. Results

4. RESULTS

The results obtained in respect of pearl millet world collection, core collection, minimum core and application of molecular markers for diversity assessment are presented below.

4.1. Studies on the base collection

4.1.1. Morpho-agronomic diversity in the world pearl millet collection

The accessions characterized previously at ICRISAT were classified based on geographical regions into four groups. The frequency distribution of accessions was studied by country of origin in each geographical region (See Table 2). Most of these regions were highly represented in the world collection. The largest collection, about 35% were from Indian subcontinent. East African countries were represented with only 951 accessions. The other two geographical regions, WAC and SCAC, were represented as 32 and 27%, respectively.

Significant differences were found in mean and variance for eleven quantitative characters, as tested through Newman-Keul's test for post-hoc comparison of mean and Levene's test for homogeneity of variances among accessions of world collection and different geographical regions (Table 5). In the world collection, accessions from all the African regions showed high variation for days to flowering during seasons, *rabi* and *kharif*. High variation in plant height was observed among accessions from the Indian subcontinent, West Africa and South and Central Africa. The East African and West African accessions contributed to the variation in number of productive tillers and spike exsertion, respectively. The South and Central African accessions accounted for high variability in length and thickness of spike. High variability for grain weight was observed in only West African collection.

Table 5. Mean and variance for eleven quantitative characters in the world pearl collection and different geographical regions

Character	World Collection			ISC				
	Mean	SE	Variance	Mean	SE	N-K Test	Variance	Levene's Test
Days to flowering (K)	76.0	0.20	550.8	61.5	0.20	**	195.1	**
Days to flowering (R)	71.0	0.10	140.9	67.7	0.10	**	84.3	**
Plant height (K)	245.0	0.50	4163.0	210.0	0.80	**	3806.6	NS
Plant height (R)	160.0	0.30	1434.3	155.0	0.50	**	1280.9	**
Number of productive tillers	3.0	0.01	3.3	3.0	0.03	NS	4.8	**
Spike exertion	3.6	0.10	45.1	6.6	0.06	**	22.4	**
Spike length (K)	28.2	0.10	119.1	22.1	0.07	**	29.3	**
Spike length (R)	25.7	0.10	120.4	19.2	0.07	**	26.9	**
Spike thickness (K)	24.0	0.03	25.3	21.7	0.05	**	13.6	**
Spike thickness (R)	23.3	0.04	28.3	20.9	0.05	**	15.4	**
Thousand grain weight	8.6	0.02	5.6	7.4	0.02	**	2.6	**

K = *Kharif* or Rainy Season data; R = *Rabi* or Post-rainy season data

Table 5. Contd.

Character	EAC					WAC				
	Mean	SE	N-K Test	Variance	Levene's Test	Mean	SE	N-K Test	Variance	Levene's Test
Days to flowering (K)	79.0	0.49	**	248.6	**	81.0	0.35	**	638.3	**
Days to flowering (R)	69.0	0.39	**	156.6	**	72.0	0.16	**	132.7	**
Plant height (K)	250.0	1.33	**	1795.8	**	265.0	0.77	**	3005.3	**
Plant height (R)	140.0	0.95	**	907.8	**	160.0	0.57	**	1644.6	**
Number of productive tillers	3.0	0.07	**	5.1	**	2.0	0.01	**	1.3	**
Spike exertion	3.1	0.17	**	29.4	**	1.3	0.11	**	58.4	**
Spike length (K)	23.7	0.14	**	20.3	**	33.8	0.19	**	199.7	**
Spike length (R)	22.1	0.15	**	21.9	**	30.1	0.19	**	198.4	**
Spike thickness (K)	22.4	0.12	**	14.1	**	25.1	0.07	**	25.1	**
Spike thickness (R)	22.7	0.12	**	14.6	**	23.4	0.07	**	25.1	**
Thousand grain weight	8.1	0.06	**	3.9	**	10.2	0.04	**	6.7	**

K = *Kharif* or Rainy Season data; R = *Rabi* or Post-rainy season data; EAC = East African Collection; WAC = West African Collection

Table 5: Contd.

Character	SCAC				
	Mean	SE	N-K Test	Variance	Levene's Test
Days to flowering (K)	86.0	0.36	**	571.2	**
Days to flowering (R)	76.0	0.20	**	179.7	**
Plant height (K)	275.0	0.90	**	3538.9	**
Plant height (R)	175.0	0.52	**	1192.6	**
Number of productive tillers	2.0	0.02	**	1.6	**
Spike exertion	2.5	0.10	**	43.6	**
Spike length (K)	30.2	0.13	**	70.1	**
Spike length (R)	29.3	0.13	**	75.5	**
Spike thickness (K)	26.1	0.08	**	29.5	**
Spike thickness (R)	26.4	0.09	**	34.6	**
Thousand grain weight	8.5	0.03	**	3.5	**

K = *Kharif* or Rainy Season data; R = *Rabi* or Post-rainy season data;
SCAC = South and Central African Collection

The mean, variance and range for eleven quantitative characters in different countries of origin was studied (Table 6). Similarly the percent distribution of these characters and five qualitative characters are presented in Table 7. The results for each character are described separately.

4.1.2. Quantitative characters

Days to flowering

In the world collection, days to flowering ranged from 33 to 159 days, with a mean of 76 days during the rainy season. However, during post rainy season, it ranged from 32 to 138 days with a mean of 71 days. High variation among the accessions between countries and significant variation among accessions within a country was also observed. For example, it is 36 to 150 days in Ghana, 36 to 149 in Togo, 35 to 146 in Burkina Faso, and 44 to 152 in Benin, while variation was less in some countries where the range was 140 to 159 days in Sierra Leone, 36 to 73 days Pakistan and 47 to 87 days in Namibia. The range of variation was reduced to a greater extent during post-rainy season. The accessions flowering before 50 days were considered early and that flowering after 120 days were late types. Most of the accessions from West Africa flowered early in post-rainy season. The late types were represented to an extent of 6.19 percent in the world collection. The accessions that flowered very late were from Sierra Leone, Benin, Burkina Faso, Central Africa, Ghana, Mali and Cameroon. There were very few late accessions from the Indian subcontinent and East Africa (0.86 percent of the world collection).

Plant Height

Enormous variation for plant height was observed both within and among accessions. It ranged from 30 to 480 cm with a mean of 245 cm and from 25 to 425 cm with a mean of 160 cm during rainy and post rainy season, respectively in the world collection. The difference in plant height during rainy and post-rainy season was as high as 170 cm in several accessions from Burkina Faso, Chad and Ghana. Interestingly, accessions from all the countries, except Sierra Leone grew taller during *kharif*. The plants were considered short for a height less than 100 cm and tall when it was greater than 350 cm. In the world collection, most of the accessions were of medium height. The tallest materials are

Table 6. ***Mean*, variance and range of eleven quantitative characters in the world pearl millet collection and different countries of origin**

Countries of origin	Days to Flowering(*Kharif*)					Days to Flowering(*Rabi*)					Plant Height (*Kharif*)				
	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max
WC	**76**	**0.20**	**550.8**	**33**	**159**	**71**	**0.10**	**140.9**	**32**	**138**	**245**	**0.50**	**4163.0**	**30**	**480**
Benin	108	5.27	1276.3	44	152	74	1.77	144.2	53	115	245	4.84	1077.7	170	310
Botswana	64	0.97	77.6	43	105	88	1.27	139.3	60	113	225	3.37	930.6	160	300
Burkina Faso	105	0.69	391.4	35	146	71	0.31	80.2	50	109	315	1.99	3195.6	80	480
Cameroon	114	0.62	355.8	56	146	67	0.28	71.2	46	128	290	1.62	2401.0	90	400
Central Africa	128	1.01	144.7	57	140	73	0.81	94.4	57	132	315	6.11	5294.4	140	410
Chad	82	2.50	238.3	46	116	65	1.99	151.3	52	98	300	9.85	3684.0	140	400
Ghana	51	1.06	324.1	36	150	60	0.64	120.3	43	100	215	1.81	938.1	120	320
India	62	0.19	192.9	33	139	68	0.13	81	32	120	210	0.85	3819.2	30	415
Kenya	81	1.61	190.7	53	140	63	1.58	186.5	45	94	240	4.23	1326.1	100	310
Malawi	65	0.75	165.2	52	119	66	0.54	84.5	49	94	245	1.88	1023.8	80	380
Mali	87	0.64	418.6	47	142	66	0.32	104.8	42	132	285	1.54	2427.4	35	460
Namibia	72	0.17	30.8	47	87	65	0.21	47.1	45	102	325	1.02	1102.6	170	430
Niger	76	0.38	130.4	43	122	80	0.27	66.3	55	110	255	1.28	1467.5	110	365
Nigeria	71	0.48	257.6	46	130	71	0.26	77	41	101	255	1.24	1709.4	90	380
Pakistan	47	0.64	61.7	36	73	54	0.27	10.9	48	65	160	2.52	962.1	110	280
Senegal	89	1.08	406.4	49	134	72	0.43	63.3	47	112	235	1.77	1098.0	95	320
Sierra Leone	149	0.55	14.9	140	159	118	1.14	74.7	90	131	150	3.71	715.3	105	190
South Africa	72	0.64	56.9	48	95	83	0.91	114.7	62	130	230	3.15	1366.6	125	325
Sudan	86	0.53	164.9	47	128	71	0.56	181.7	38	120	250	1.94	2188.8	90	420
Tanzania	107	0.69	210	53	135	90	0.47	104.8	65	138	305	1.8	1512.7	190	400
Togo	62	1.36	885.5	36	149	71	0.36	62.6	52	113	225	2.25	2427.2	110	460
Uganda	74	3.27	919.4	47	143	70	1.00	86.0	47	95	205	6.15	3257.3	50	330
Yemen	74	1.13	77.7	53	110	73	1.28	99.5	48	98	285	3.39	700.7	200	320
Zambia	78	1.86	317.4	45	121	72	0.86	68.1	50	95	220	5.4	2677.9	95	380
Zimbabwe	70	0.21	57.6	44	98	85	0.25	83.3	59	110	230	1.11	1681.8	90	360

Table 6. Contd.

Countries of origin	Plant Height (*Rabi*)					Number of Productive tillers					Spike Exsertion				
	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max
WC	**160**	**0.30**	**1434.3**	**25**	**425**	**3**	**0.01**	**3.3**	**1**	**19**	**3.6**	**0.10**	**45.1**	**-45**	**29**
Benin	145	3.75	646.0	100	210	4	0.29	3.8	1	9	1.4	1.13	56.25	-15	11
Botswana	145	2.60	580.2	80	210	3	0.13	1.5	1	7	1.7	0.73	42.79	-25	10
Burkina Faso	145	1.13	1063.5	65	240	2	0.04	10.1	1	11	2.1	0.25	46.97	-16	20
Cameroon	155	1.05	1010.3	75	345	3	0.05	2.1	1	11	5.2	0.17	23.53	-18	17
Central Africa	155	3.08	1382.9	25	310	3	0.09	1.3	1	6	5.2	0.31	13.77	-18	13
Chad	145	3.75	533.7	60	180	4	0.21	1.6	1	6	7.9	0.56	12.03	3	15
Ghana	135	1.40	563.0	50	250	3	0.06	1.1	1	8	8.0	0.25	17.45	-11	22
India	155	0.50	1303.9	25	335	3	0.03	4.9	1	9	6.6	0.07	22.53	-19	29
Kenya	135	2.96	658.9	80	245	6	0.58	24.1	1	19	3.2	0.59	26.2	-11	16
Malawi	160	1.71	848.7	75	230	3	0.09	2.4	1	15	1.5	0.27	20.56	-18	11
Mali	145	1.20	1474.8	60	425	2	0.03	0.7	1	6	2.9	0.19	35.53	-18	20
Namibia	185	0.63	419.1	130	310	1	0.00	0.0	1	4	1.4	0.22	51.36	-20	14
Niger	175	1.24	1390.8	80	285	2	0.05	2.1	1	14	-2.3	0.28	66.85	-45	21
Nigeria	195	1.07	1263.2	50	290	2	0.03	0.9	1	13	-1.0	0.27	75.87	-32	25
Pakistan	155	1.79	484.9	95	230	3	0.09	1.2	1	7	7.5	0.35	18.34	-9	14
Senegal	125	1.73	1044.9	45	220	2	0.04	0.5	1	5	-0.4	0.31	32.57	-22	14
Sierra Leone	165	4.45	1147.2	90	250	4	0.09	0.5	3	6	0.7	0.74	28.6	-9	10
South Africa	170	2.20	672.5	90	290	3	0.19	4.9	1	19	0.5	0.59	45.27	-15	13
Sudan	135	1.09	697.4	60	285	3	0.07	3.4	1	16	3.6	0.24	32.65	-15	17
Tanzania	195	1.71	1368.1	105	285	2	0.03	0.3	1	4	-2.1	0.37	62.81	-22	16
Togo	145	1.01	493.4	65	230	3	0.05	1.2	1	12	5.8	0.22	23.5	-20	16
Uganda	145	3.24	900.8	55	220	3	0.13	1.4	1	7	4.2	0.55	23.81	-15	15
Yemen	160	3.32	671.2	95	205	3	0.19	2.4	1	10	5.8	0.64	23.12	-5	18
Zambia	125	2.72	673.9	70	230	3	0.19	3.4	1	15	-0.6	0.71	41.09	-18	11
Zimbabwe	185	0.85	973.8	40	300	2	0.03	0.9	1	9	3.3	0.15	32.22	-16	15

Table 6. Contd.

Countries of origin	Spike Length (*Kharif*)					Spike Length (*Rabi*)					Spike Thickness (*Kharif*)				
	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max
WC	**28.2**	**0.10**	**119.1**	**5**	**120**	**25.7**	**0.10**	**120.4**	**4**	**125**	**24.0**	**0.03**	**25.3**	**8**	**58**
Benin	29.7	1.85	156.9	14	60	28.3	2.09	201.8	10	75	24.9	0.50	11.5	19	33
Botswana	24.4	0.76	47.2	9	44	24.3	0.89	68.9	8	45	26.1	0.66	36.2	13	55
Burkina Faso	30.8	0.34	111.6	15	88	25.6	0.31	78.9	11	80	23.3	0.13	12.6	14	40
Cameroon	23.4	0.15	20.7	12	70	22.3	0.16	22.4	12	80	21.7	0.10	10.0	13	38
Central Africa	20.2	0.34	16.4	11	35	20.9	0.31	14.3	8	34	19.2	0.25	8.7	14	33
Chad	22.6	0.59	13.3	16	30	20.9	0.78	22.8	10	30	21.1	0.48	8.6	15	27
Ghana	19.6	0.23	15.7	12	36	16.8	1.20	11.0	9	30	29.6	0.23	14.5	17	42
India	22.1	0.07	29.4	5	85	19.2	0.07	27.1	5	70	21.6	0.05	13.7	8	47
Kenya	20.9	0.41	12.2	15	30	20.4	0.45	15.1	8	30	23.5	0.47	16.6	17	35
Malawi	25.9	0.21	12.7	16	42	24.3	0.25	18.8	15	40	22.3	0.17	8.5	16	36
Mali	29.6	0.21	42.8	15	56	25.2	0.18	34.1	11	63	25.6	0.14	20.9	10	44
Namibia	38.4	0.22	49.0	13	66	33.9	0.18	35.2	15	65	30.9	0.15	22.2	20	58
Niger	45.2	0.54	261.4	6	114	39.9	0.49	218.6	4	95	24.0	0.16	23.6	12	55
Nigeria	40.5	0.47	245.0	12	120	39.9	0.51	283.8	12	125	24.3	0.11	13.1	10	40
Pakistan	21.6	0.40	24.5	12	34	19.5	0.40	24.5	12	34	20.2	0.20	5.6	14	26
Senegal	33.1	0.36	44.7	15	55	27.3	0.39	53.3	11	55	20.9	0.19	12.3	13	40
Sierra Leone	24.9	0.58	17.5	17	42	24.4	0.42	10.1	17	34	21.8	0.48	11.9	16	33
South Africa	31.6	0.55	41.6	13	50	32.6	0.65	58.8	16	57	24.6	0.41	23.4	12	45
Sudan	23.1	0.19	20.7	13	42	21.4	0.19	20.3	9	46	22.0	0.17	16.1	13	37
Tanzania	31.0	0.34	55.2	13	54	33.4	0.49	111.4	10	75	25.6	0.19	17.0	12	55
Togo	21.4	0.18	14.9	9	50	18.1	0.18	15.6	7	35	31.5	0.24	27.1	15	53
Uganda	23.9	0.58	29.0	12	45	20.5	0.45	17.2	11	33	23.7	0.40	13.9	16	35
Yemen	22.1	0.63	24.1	15	33	20.8	0.69	28.7	13	40	24.7	0.41	10.4	14	33
Zambia	25.1	0.48	21.1	15	35	22.3	0.53	25.3	12	35	23.7	0.39	13.8	16	34
Zimbabwe	30.4	0.17	40.9	10	60	31.4	0.22	66.8	10	74	26.5	0.12	19.8	14	50

Table 6. Contd.

Countries of origin	Spike Thickness (*Rabi*)					1000-Grain Weight				
	Mean	S.E.	Variance	Min	Max	Mean	S.E.	Variance	Min	Max
WC	**23.3**	**0.04**	**28.3**	**9**	**61**	**8.6**	**0.02**	**5.6**	**1.5**	**21.33**
Benin	26.8	0.62	17.9	18	40	12.2	0.42	8.1	6.8	19.32
Botswana	28.3	0.78	52.3	15	52	8.1	0.19	3.1	4.0	12.60
Burkina Faso	22.1	0.15	19.4	13.5	55	12.2	0.07	3.7	6.0	21.33
Cameroon	21.5	0.11	13.1	13	42	9.7	0.05	2.0	5.2	18.63
Central Africa	21.6	0.29	11.9	14	30	9.1	0.16	3.6	4.7	12.75
Chad	20.5	0.60	13.7	12	28	8.9	0.23	2.1	5.5	11.5
Ghana	27.4	0.26	20.2	15	40	13.0	0.10	3.1	5.5	18.64
India	20.8	0.05	15.5	9	40	7.3	0.02	2.6	1.5	16.50
Kenya	22.9	0.47	16.3	16	35	6.3	0.23	3.8	3.1	11.48
Malawi	22.3	0.19	10.3	13	46	6.6	0.06	1.1	3.3	9.71
Mali	23.9	0.18	31.4	13	49	9.5	0.05	2.7	5.0	16.38
Namibia	28.5	0.14	21.3	10	50	9.8	0.04	1.3	7.0	14.20
Niger	22.5	0.16	24.5	12	40	9.0	0.05	2.6	4.3	14.03
Nigeria	22.4	0.11	12.2	13	38	8.8	0.04	1.7	4.4	14.18
Pakistan	20.6	0.23	8.1	14	30	7.6	0.06	0.6	5.7	9.90
Senegal	20.1	0.21	15.9	12	40	7.4	0.06	1.3	3.2	11.51
Sierra Leone	19.8	0.36	7.3	16	25	4.2	0.06	0.2	3.2	5.25
South Africa	24.4	0.40	22.4	12	43	6.8	0.16	3.4	2.9	11.36
Sudan	22.6	0.17	16.3	12	38	8.7	0.06	2.4	2.9	14.90
Tanzania	28.2	0.28	35.7	15	52	7.0	0.06	1.8	2.5	11.55
Togo	28.2	0.18	14.8	14	40	13.7	0.08	3.5	4.1	18.45
Uganda	23.3	0.43	15.6	17	40	8.3	0.2	3.3	4.3	12.90
Yemen	24.5	0.46	12.7	17	33	11.1	0.27	4.4	6.9	15.81
Zambia	21.0	0.37	12.3	13	32	5.6	0.17	2.5	2.7	14.95
Zimbabwe	28.7	0.16	32.4	11	61	7.5	0.04	2.0	3.5	19.32

Table 7. Frequency distribution of seven quantitative and five qualitative characters in world pearl millet collection and different geographical regions

Characters	Modality	WC (%)	ISC (%)	RATIO	AC (%	RATIO	EAC (%)	RATIO	CAC (%	RATIO
Quantitative										
All Traits		10.7	10.4		14.0		9.1		10.9	
Days to flowering	**Early**	9.93	44.3	4.3	18.4	1.3	0.7	0.1	2.4	0.2
	Late	6.19	0.06	0.0	21.0	1.5	0.8	0.1	15.4	1.4
Plant Height	**Short**	1.78	1.39	0.1	3.3	0.2	0.7	0.1	1.2	0.1
	Tall	4.91	1.06	0.1	3.8	0.3	0.7	0.1	10.0	0.9
Number of productive tillers	**Poor**	71.5	44.7	4.3	71.3	5.1	58.1	6.4	67.7	6.2
	Good	0.97	1.08	0.1	0.14	0.0	4.6	0.5	0.4	0.0
Spike Exertion	**Poor**	27.4	7.94	0.8	38.6	2.8	24.7	2.7	30.8	2.8
	Good	0.05	0.04	0.0	0.1	0.0	-	-	-	-
Spike Length	**Short**	19.9	40.7	3.9	18.6	1.3	26.7	2.9	16.8	1.5
	Long	0.02	-	-	0.03	0.0	-	-	-	-
Spike Thickness	**Thin**	1.21	0.52	0.1	3.2	0.2	0.6	0.1	0.8	0.1
	Thick	0.46	0.03	0.0	0.4	0.0	-	-	0.7	0.1
1000-Grain Weight	**Small**	3.88	3.92	0.4	10.7	0.8	8.3	0.9	6.1	0.6
	Large	1.56	0.05	0.0	6.7	0.5	1.2	0.1	0.1	0.0

Table 7. Contd.

Qualitative										
All Traits		7.2	0.17		0.17		0.17		0.17	
Bristle Length	**Short**	20.9	0.3	1.76	0.15	0.88	0.19	1.12	0.15	0.88
	Medium	65.6	0.63	3.71	0.64	3.76	0.65	3.82	0.57	3.35
	Large	13.5	0.06	0.35	0.19	1.12	0.15	0.88	0.27	1.59
Spike Density	**Loose**	57.0	0.58	3.41	0.48	2.82	0.48	2.82	0.68	4
	Medium	39.1	0.4	2.35	0.45	2.65	0.44	2.59	0.28	1.65
	Compact	3.9	0.02	0.12	0.07	0.41	0.08	0.47	0.04	0.24
Spike Shape	**Cylindrical**	28.0	0.26	1.53	0.32	1.88	0.48	2.82	0.19	1.12
	Conical	8.4	0.12	0.71	0.09	0.53	0.07	0.41	0.02	0.12
	Spindle	2.0	0.03	0.18	0.04	0.24	0.01	0.06	.	.
	Candle	47.4	0.5	2.94	0.43	2.53	0.29	1.71	0.51	3
	Lanceolate	11.6	0.09	0.53	0.05	0.29	0.14	0.82	0.22	1.29
Grain Shape	**Obovate**	19.7	0.29	1.71	0.11	0.65	0.32	1.88	0.15	0.88
	Oblanceolate	20.7	0.28	1.65	0.09	0.53	0.18	1.06	0.26	1.53
	Elliptical	12.398	0.11	0.65	0.15	0.88	0.06	0.35	0.14	0.82
	Hexagonal	17.9	0.19	1.12	0.21	1.24	0.13	0.76	0.14	0.82
	Globular	29.4	0.13	0.76	0.34	2	0.31	1.82	0.31	1.82
Grain Color	**Ivory**	53.501	0.39	2.29	0.63	3.71	0.45	2.65	0.61	3.59
	Cream	35.8	0.46	2.71	0.29	1.71	0.44	2.59	0.27	1.59
	Yellow	6.0	0.1	0.59	0.03	0.18	0.06	0.35	0.04	0.24
	Grey	1.8	0.03	0.18	0.01	0.06	.	.	0.02	0.12

represented to an extent of five percent and they were mostly from South and Central African countries. In general, late maturing accessions grew taller in both the seasons.

Number of productive tillers

The number of productive tillers ranged from 1 to 19, in the world collection, with a modal distribution of poor and good tillering ability. Most of the accessions in the world collection (71.49 percent) were found to have poor tillering ability. The accessions with good tillering (0.97 percent) were mainly from the East African countries. High variability was observed for tiller number in accessions from India, Sudan and Cameroon.

Spike Exsertion

The relative length of peduncle and leaf sheath of the boot leaf determines spike exsertion. This length varied from -45 to 29 cm in the world collection. Negative exsertion was considered poor as part of the spike remained within the boot leaf, leading to poor seed set. Accessions from Ghana, Pakistan, India, Yemen and Togo showed positive exsertion. Most of the accessions (72.55 percent) in the world collection had medium exsertion of 10 cm.

Spike Length

Spike length is a yield-related character that varied in the world collection from 5 to 120 cm with a mean of 28 cm in *kharif,* and 4 to 125 cm with a mean of 26cm in *rabi*. Spike length was found to be relatively stable across seasons, *rabi* and *kharif*, with little variation. However, variation in length was high among accessions within a country. The range was more among accessions from Nigeria, Niger and Burkina Faso. Usually, longer spikes attribute for higher yield. The modal distribution for this character was slightly skewed towards the lower values. Indian subcontinent (40.68%) and East African countries (18.6%) represented mostly short spikes. There was very poor representation of the longer spikes (0.02%) in the world collection, mostly from Niger, Nigeria and Namibia. Usually, the late and medium maturing accessions produced longer spikes.

Spike Thickness

Spike thickness depends mainly on the length of involucre, rachis thickness, grain size and spike length. This varied between 8 to 58 mm with a mean of 24 mm during *kharif* and 9 to 61 mm with a mean of 23 mm during *rabi*, depicting stability of the character over seasons. Early maturing accessions with more number of tillers produced thin spikes. Most of the accessions were of medium thickness in the collection. Accessions from Togo, Namibia and Ghana produced thick spikes.

1000-Grain Weight

In the world collection, 1000-grain weight varied from 1.5 to 21.3 g with a mean of 8.66 g. The early maturing accessions from West Africa, mostly Ghana, Togo and Benin were the main source of large grains. Smallest seeds were found in accessions from India, Malawi, Sierra Leone, Tanzania and Zimbabwe contributing 3.88 per cent of the world collection.

4.1.3. Qualitative characters

Bristle Length

Most of the accessions in world collection had short or no bristles. However, the late maturing accessions from Senegal produced very long bristles.

Spike Shape and Density

The collection varied widely for spike shape and were categorized in a 1-9 score. Cylindrical and candle-shaped spikes were common in most of the region. The occurrence of club, oblanceolate, and globose spikes was considerably rare. The accessions from India and Zimbabwe represented all the nine spike shapes, indicating large amount of diversity. Accessions with loose spikes and stiff branches were largely represented in the collection. Compact spikes were found only in West African collection.

Grain Colour and Shape

Grain colour varied greatly among accessions within a country and among countries. The grain colour in pearl millet was predominantly grey (48%). Grains with yellow endosperm constituted only 3.3% of the collection. Grains classified as ivory or cream constituted only

5% of the world collection. Some of the West African accessions had purplish-black grains.

The grain shapes are categorized in a 1-5 score. Variation in grain shape was relatively low within different spikes of the same plant and plants of the same accession. This was found to be a stable character to distinguish different accessions. The common shapes found in the collection are globular (29%), oblanceolate (21%), obovate (20%), hexagonal (18%) and elliptical (12%). Most of the accessions from Sierra Leone produced elliptical grains.

4.1.4. Diversity in countries of origin

The diversity in the world pearl millet collection was compared with all geographical regions and countries of origin by Shannon-Weaver diversity indices (Table 8 & 9) for quantitative and qualitative characters independently, and also on the basis of all the characters. No significant differences were observed for diversity among quantitative characters over geographical regions. Individually, high amount of diversity was found for all the characters except in number of productive tillers, for all the regions. The mean diversity represented by the qualitative characters in different geographical regions was significantly higher than that of the world collection. However, the trend was not the same for different countries of origin. The accessions from Ghana, Malawi, Central Africa and Uganda represented less diversity for days to flowering. The mean diversity over qualitative characters was found to be significantly less for the accessions from Benin, Central Africa, Ghana, Burkina Faso, Kenya, Malawi, Namibia, Pakistan, Sierra Leone, Chad, Yemen and Zimbabwe. Diversity over all quantitative characters in collections from Botswana (0.58 ± 0.03), India (0.59 ± 0.07), Mali (0.58 ± 0.05), Niger (0.58 ± 0.02) and Sudan (0.60 ± 0.04) were not different from world collection diversity (0.59 ± 0.06). However, the diversity in collection from Senegal (0.64 ± 0.02) was significantly higher than the world collection diversity. The diversity represented by other countries of origin were significantly less than the world collection diversity (Table 9). The Chi-square analysis revealed significant differences for diversity over all characters in countries like Sierra Leone, Chad, Togo, Namibia and Pakistan.

Table 8. Shannon-Weaver diversity indices for quantitative and qualitative character

Characters	WC	ISC	WAC	EAC	SCAC
Quantitative					
Days to flowering (*kharif*)	0.59	0.53	0.59	0.6	0.58
Days to flowering (*Rabi*)	0.6	0.59	0.61	0.62	0.61
Plant Height (*Kharif*)	0.63	0.62	0.63	0.6	0.62
Plant Height (*Rabi*)	0.63	0.6	0.61	0.62	0.6
Number of productive tillers	0.4	0.47	0.51	0.46	0.47
Spike Exertion	0.59	0.62	0.62	0.61	0.61
Spike Length (*Kharif*)	0.56	0.61	0.6	0.61	0.58
Spike Length (*Rabi*)	0.57	0.59	0.58	0.62	0.59
Spike Thickness (*Kharif*)	0.61	0.61	0.61	0.6	0.57
Spike Thickness (*Rabi*)	0.58	0.62	0.58	0.47	0.59
1000-Grain Weight	0.62	0.63	0.62	0.61	0.63
Mean + S.E.	0.6+0.05	0.6+0.07	0.6+0.06	0.6+0.08	0.6+0.06
Qualitative					
Spike Shape	0.59	0.93	0.91	0.91	0.92
Spike Density	0.59	0.93	0.95	0.95	0.92
Bristle Length	0.65	0.93	0.9	0.9	0.89
Grain Shape	0.47	0.69	0.91	0.91	0.89
Grain Colour	0.68	0.88	0.7	0.7	0.7
Mean + S.E.	0.6+0.04	0.9+0.07	0.9+0.05	0.9+0.05	0.8+0.04
Total Mean + S.E.	0.6+ 0.03	0.7+ 0.0	0.7+0.04	0.7+0.07	0.7+0.05
Chi-square (d.f = 15)		0.704 **NS**	0.937 **NS**	0.944 **NS**	0.853 **NS**

*, ** P < 0.05 and P < 0.01, respectively; NS = P > 0.05; WC = World Collection;
ISC = Indian subcontinent collection; WAC = West African Collection;
EAC = East African Collection; SCAC = South and Central African Collection

Table 9. Shannon-Weaver diversity indices for eleven quantitative and five qualitative characters in world pearl millet collection and different countries of origin

Character	WC	BEN	BWA	CAF	CMR	GHA	HVO	IND	KEN
Quantitative									
Days to flowering (K)	**0.59**	0.47	0.57	0.42	0.57	0.24	0.62	0.56	0.52
Days to flowering (R)	**0.60**	0.59	0.61	0.33	0.52	0.58	0.61	0.60	0.52
Plant height (K)	**0.63**	0.60	0.60	0.56	0.63	0.62	0.61	0.62	0.60
Plant height (R)	**0.63**	0.56	0.61	0.55	0.57	0.63	0.62	0.63	0.58
Number of productive tillers	**0.40**	0.58	0.51	0.61	0.55	0.54	0.46	0.45	0.41
Spike exertion	**0.59**	0.57	0.57	0.60	0.56	0.60	0.57	0.61	0.61
Spike length (K)	**0.56**	0.50	0.58	0.61	0.58	0.61	0.54	0.61	0.57
Spike length (R)	**0.57**	0.51	0.59	0.60	0.57	0.63	0.54	0.60	0.57
Spike thickness (K)	**0.61**	0.54	0.56	0.51	0.61	0.59	0.63	0.63	0.58
Spike thickness (R)	**0.58**	0.60	0.55	0.57	0.62	0.61	0.58	0.64	0.58
Thousand grain weight	**0.62**	0.61	0.60	0.61	0.61	0.60	0.62	0.62	0.59
Mean ± SE	**0.58**	0.56	0.58	0.54	0.58	0.57	0.58	0.60	0.56
	±0.05	±0.08	±0.04	±0.12	±0.09	±0.04	±0.02	±0.07	±0.11
Qualitative									
Spike Shape	**0.59**	0.40	0.64	0.61	0.52	0.55	0.44	0.56	0.51
Spike Density	**0.65**	0.49	0.47	0.31	0.74	0.57	0.57	0.63	0.57
Bristle Length	**0.47**	0.29	0.77	0.38	0.37	0.36	0.18	0.50	0.48
Grain Shape	**0.68**	0.32	0.63	0.49	0.53	0.35	0.34	0.67	0.46
Grain Colour	**0.59**	0.55	0.33	0.56	0.59	0.61	0.58	0.43	0.34
Mean ± SE	**0.60**	0.41	0.57	0.45	0.55	0.49	0.42	0.56	0.47
	±0.09	±0.15	±0.08	±0.22	±0.12	±0.22	±0.32	±0.12	±0.22
Total Mean ±SE	**0.59**	0.51	0.58	0.51	0.57	0.55	0.53	0.59	0.53
	±0.06	±0.10	±0.05	±0.15	±0.10	±0.10	±0.11	±0.09	±0.14
Chi-square (d.f. = 15)		0.499	0.405	0.572	0.154	0.470	0.412	0.069	0.226
		NS	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**

K = Kharif; R = Rabi; BEN = Benin; BWA = Botswana; CAF = Central Africa; CMR= Cameroon; GHA = Ghana; HVO = Burkina Faso; IND = India; KEN = Kenya

Table 9. Contd.

Character	WC	MLI	MWI	NAM	NER	NGA	PAK	SDN	SEN
Quantitative									
Days to flowering (K)	**0.59**	0.62	0.34	0.62	0.60	0.52	0.54	0.63	0.54
Days to flowering (R)	**0.60**	0.61	0.59	0.60	0.62	0.61	0.63	0.60	0.57
Plant height (K)	**0.63**	0.62	0.54	0.64	0.63	0.63	0.54	0.63	0.60
Plant height (R)	**0.63**	0.53	0.63	0.63	0.62	0.63	0.57	0.62	0.63
Number of productive tillers	**0.40**	0.53	0.20	0.46	0.32	0.46	0.57	0.49	0.43
Spike exsertion	**0.59**	0.62	0.60	0.55	0.62	0.61	0.45	0.58	0.62
Spike length (K)	**0.56**	0.62	0.60	0.59	0.62	0.59	0.60	0.61	0.61
Spike length (R)	**0.57**	0.58	0.62	0.62	0.62	0.59	0.59	0.62	0.61
Spike thickness (K)	**0.61**	0.61	0.62	0.55	0.60	0.62	0.59	0.63	0.59
Spike thickness (R)	**0.58**	0.62	0.59	0.58	0.61	0.62	0.56	0.63	0.61
Thousand grain weight	**0.62**	0.63	0.59	0.60	0.63	0.63	0.60	0.63	0.63
Mean ± SE	**0.58**	0.60	0.56	0.59	0.59	0.59	0.57	0.61	0.64
	±0.05	±0.18	±0.07	±0.12	±0.09	±0.14	±0.12	±0.07	±0.11
Qualitative									
Spike Shape	**0.59**	0.56	0.43	0.15	0.47	0.43	0.17	0.51	0.51
Spike Density	**0.65**	0.66	0.58	0.47	0.63	0.64	0.30	0.63	0.73
Bristle Length	**0.47**	0.30	0.35	0.11	0.41	0.35	0.16	0.46	0.68
Grain Shape	**0.68**	0.58	0.38	0.50	0.67	0.63	0.40	0.56	0.66
Grain Colour	**0.59**	0.52	0.49	0.50	0.66	0.57	0.05	0.69	0.61
Mean ± SE	**0.60**	0.52	0.45	0.35	0.57	0.52	0.22	0.57	0.64
	±0.09	±0.15	±0.28	±0.32	±0.12	±0.18	±0.35	±0.02	±0.02
Total Mean ±SE	**0.59**	0.58	0.53	0.52	0.58	0.57	0.39	0.60	0.64
	±0.06	±0.17	±0.14	±0.18	±0.10	±0.15	±0.19	±0.06	±0.08
Chi-square (d.f. = 15)		0.157	0.459	0.741	0.072	0.102	1.437	0.087	0.136
		NS	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**

K = Kharif; R = Rabi; MLI = Mali; MWI = Malawi; NAM = Namibia;
NER = Niger; NGA = Nigeria; PAK = Pakistan; SDN = Sudan; SEN = Senegal

Table 9. Contd.

Character	SLE	TCD	TGO	TZA	UGA	YEM	ZAF	ZIM	ZMB
Quantitative									
Days to flowering (K)	0.57	0.60	0.32	0.57	0.30	0.53	0.62	0.63	0.46
Days to flowering (R)	0.30	0.35	0.59	0.65	0.63	0.54	0.58	0.64	0.61
Plant height (K)	0.54	0.61	0.53	0.62	0.61	0.59	0.60	0.63	0.59
Plant height (R)	0.62	0.52	0.61	0.62	0.59	0.62	0.63	0.63	0.59
Number of productive tillers	0.57	0.58	0.55	0.48	0.61	0.47	0.29	0.44	0.43
Spike exsertion	0.41	0.57	0.52	0.62	0.55	0.55	0.57	0.58	0.56
Spike length (K)	0.56	0.58	0.59	0.61	0.59	0.55	0.61	0.60	0.60
Spike length (R)	0.62	0.61	0.62	0.60	0.62	0.56	0.62	0.63	0.59
Spike thickness (K)	0.56	0.60	0.63	0.62	0.59	0.57	0.59	0.59	0.59
Spike thickness (R)	0.54	0.61	0.62	0.60	0.58	0.62	0.61	0.62	0.61
Thousand grain weight	0.59	0.57	0.61	0.63	0.60	0.59	0.59	0.59	0.51
Mean ± SE	0.53	0.56	0.56	0.60	0.57	0.56	0.57	0.60	0.56
	±0.15	±0.08	±0.07	±0.02	±0.09	±0.14	±0.12	±0.07	±0.13
Qualitative									
Spike Shape	0.14	0.24	0.52	0.62	0.49	0.58	0.50	0.48	0.54
Spike Density	0.51	0.59	0.54	0.56	0.45	0.56	0.58	0.53	0.70
Bristle Length	0.00	0.05	0.08	0.52	0.63	0.25	0.41	0.49	0.69
Grain Shape	0.19	0.52	0.37	0.63	0.61	0.61	0.64	0.65	0.54
Grain Colour	0.33	0.75	0.40	0.51	0.51	0.42	0.52	0.36	0.50
Mean ± SE	0.23	0.43	0.38	0.57	0.54	0.48	0.53	0.507	0.59
	±0.19	±0.15	±0.22	±0.12	±0.12	±0.28	±0.15	±0.22	±0.02
Total Mean ±SE	0.44	0.52	0.50	0.59	0.56	0.54	0.56	0.57	0.57
	±0.16	±0.10	±0.12	±0.05	±0.10	±0.18	±0.08	±0.12	±0.10
Chi-square (d.f. = 15)	1.615	0.885	0.767	0.064	0.072	0.209	0.087	0.158	0.217
	NS	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**	**NS**

K = Kharif; R = Rabi; SLE = Sierra Leone; TCD = Chad; TGO = Togo;
TZA= Tanzania; UGA = Uganda; YEM = Yemen; ZAF = South Africa;
ZIM = Zimbabwe; ZMB = Zambia

The accessions grouped based on countries of origin were further subjected to hierarchical cluster analysis to establish the core sample.

4.2. Development of core collection

4.2.1. Stratification of the collection

The methodology adopted in developing the core collection is represented in Fig 2. The countries of origin were then subjected to hierarchical cluster analysis based on characters days to flowering (*kharif* and *rabi*), plant height (*kharif* and *rabi)*, number of productive tillers, spike exsertion, spike length (*kharif* and *rabi)*, spike thickness (*kharif* and *rabi)* and 1000-grain weight. The analysis resulted in 25 distinct clusters with 90% variation explained by the data (R^2 = 0.9). All those accessions that were predominantly late maturing, growing over 300 cm tall, producing very long spikes and the early maturing accessions that are ecologically and geographically similar were grouped together. The accessions from Mali, Sudan, Senegal, Niger and Nigeria were clustered together except, Namibia. All the germplasm from Sierra Leone with distinct day-length and temperature sensitive accessions with small elliptical grains formed a group. The range for each character used in the study for 25 clusters are described in Table 10. The late flowering accessions with medium height from Sierra Leone and other countries were recovered in cluster 10 that formed a distinct group. Considerable variation was observed within and between clusters for most of the characters. The early maturing accessions with high tillering ability that produced loose heads were grouped together. The variation present in the world collection for each character was well depicted in all clusters. The selection of accessions from these clusters represented the whole range of world collection diversity as described in Table 11. No significant differences were observed in diversity among different clusters when compared with world collection.

4.2.2. Selection of accessions from different clusters and geographical regions

About 10% accessions were proportionally selected from each of 25 clusters to form the core sample of 1600 accessions. At least one accession was selected even from the clusters with less than 10 accessions, in order to represent world genetic diversity. The accessions selected from different geographical regions and clusters are presented in Table 12. A

Table 10. Range of eleven quantitative characters in the world pearl millet collection and twenty-five groups formed by cluster analysis

Clusters	Days to flowering (K)	Days to flowering (R)	Plant Height (K)	Plant Height (R)	Number of productive tillers	Spike exertion	Spike Length (K)	Spike Length (R)	Spike Thickness (K)	Spike Thickness (R)	1000-Grain Weight
WC	**33-159**	**32-138**	**30-480**	**25-425**	**1-19**	**-45-29**	**5-120**	**4-125**	**8-58**	**9-61**	**1.5-21.33**
1	33-86	33-78	90-290	55-230	0-6	-5-29	11-50	9-43	35-13	10-38	3.0-12.37
2	41-112	47-101	90-365	100-290	1-9	-28-20	12-100	10-88	12-38	12-36	4.5-13.70
3	45-137	42-96	95-390	50-275	1-5	-25-22	6-85	4-80	13-55	12-45	4.3-14.30
4	48-140	45-94	140-380	75-245	1-9	-18-13	15-75	11-65	16-36	13-46	3.1-14.0
5	33-124	32-138	60-400	45-225	1-7	-19-22	13-58	10-67	12-34	14-41	3.0-11.30
6	42-142	50-115	30-420	25-290	1-11	-33-18	6-115	5-75	8-55	11-50	2.7-17.30
7	44-110	58-100	135-320	65-200	1-6	-13-17	12-40	6-45	15-38	15-35	4.4-14.90
8	36-152	38-115	30-460	25-280	1-12	-45-25	10-100	8-125	9-44	9-49	3.5-19.30
9	44-125	58-107	100-380	60-210	1-12	-15-20	15-80	11-70	13-35	13-35	2.5-21.33
10	39-159	41-131	105-380	60-250	1-19	-23-18	14-90	11-120	10-45	12-35	2.8-17.50
11	38-130	47-105	45-330	45-255	1-15	-32-22	5-120	5-115	12-47	13-40	3.5-14.10
12	39-122	43-132	100-385	100-425	1-8	-22-18	9-90	6-95	10-42	14-55	3.4-18.60
13	37-136	35-108	120-400	90-300	1-19	-16-18	13-70	10-80	15-50	10-61	2.2-16.10
14	80-146	55-85	180-400	100-210	1-11	-18-15	16-35	12-46	15-32	13-33	5.9-18.60
15	47-146	47-107	30-380	90-220	1-17	-12-15	12-85	11-70	14-38	15-42	2.5-15.60
16	35-120	55-95	130-480	90-255	1-17	-5-20	15-60	14-55	10-35	14-35	1.5-14.80
17	40-135	54-111	110-455	110-335	1-11	-19-17	15-52	10-75	14-40	12-50	2.5-15.80
18	49-96	63-110	130-310	110-280	1-6	-10-15	11-50	13-64	15-50	15-60	3.5-11.00
19	49-90	65-110	90-310	100-360	1-4	-10-13	15-50	18-60	16-40	11-44	4.6-9.50
20	45-146	55-113	80-450	60-250	1-5	-11-15	11-60	7-60	15-45	11-50	4.6-15.60
21	36-145	46-88	160-460	25-210	1-8	-10-18	9-50	10-45	15-55	15-40	6.0-18.50
22	36-121	48-120	90-300	75-240	1-7	-10-13	13-55	10-35	15-42	13-40	3.5-16.50
23	47-87	45-78	170-430	140-310	1-4	-15-12	16-60	16-55	20-60	10-50	7.2-14.20
24	52-85	50-80	190-420	130-240	1-3	-13-14	15-66	15-65	24-54	20-45	7.0-13.50
25	60-115	50-102	240-420	120-250	1-3	-20-14	13-65	15-55	19-58	15-50	5.5-12.00

K = *Kharif* or Rainy season data; R = *Rabi* or Post-rainy season data

Table 11. Shannon-Weaver diversity indices for eleven quantitative characters in world pearl millet collection and twenty five groups formed by cluster analysis

Cluster	Days to flowering (K)	Days to flowering (R)	Plant Height (K)	Plant Height (R)	Number of productive tillers	Spike Exertion	Spike Length (K)	Spike Length (R)	Spike Thickness (K)	Spike Thickness (R)	1000-Grain Weight	Mean+SE	Chi-square	
WC	0.59	0.60	0.63	0.63	0.40	0.59	0.56	0.57	0.61	0.58	0.62	0.49+0.12		
1	0.46	0.57	0.62	0.57	0.57	0.51	0.53	0.60	0.47	0.55	0.58	0.49+0.21	0.153	NS
2	0.47	0.56	0.61	0.51	0.54	0.60	0.46	0.42	0.56	0.58	0.59	0.43+0.09	0.163	NS
3	0.51	0.56	0.61	0.60	0.42	0.63	0.48	0.48	0.63	0.60	0.62	0.45+0.21	0.046	NS
4	0.47	0.61	0.58	0.58	0.48	0.59	0.58	0.54	0.59	0.63	0.59	0.57+0.07	0.057	NS
5	0.57	0.58	0.56	0.59	0.56	0.58	0.60	0.57	0.58	0.53	0.60	0.57+0.11	0.085	NS
6	0.54	0.52	0.57	0.63	0.55	0.61	0.61	0.59	0.58	0.62	0.59	0.58+0.05	0.088	NS
7	0.52	0.54	0.52	0.57	0.54	0.64	0.55	0.53	0.53	0.61	0.60	0.56+0.22	0.108	NS
8	0.57	0.62	0.61	0.58	0.49	0.58	0.60	0.63	0.65	0.62	0.63	0.60+0.12	0.041	NS
9	0.53	0.51	0.56	0.57	0.55	0.61	0.50	0.56	0.62	0.56	0.59	0.56+0.05	0.099	NS
10	0.56	0.58	0.60	0.57	0.48	0.58	0.59	0.56	0.59	0.59	0.62	0.57+0.11	0.028	NS
11	0.39	0.56	0.61	0.50	0.40	0.55	0.55	0.56	0.57	0.60	0.55	0.53+0.19	0.112	NS
12	0.59	0.57	0.57	0.59	0.43	0.57	0.60	0.54	0.60	0.54	0.60	0.56+0.08	0.021	NS
13	0.59	0.59	0.61	0.62	0.30	0.62	0.61	0.57	0.57	0.58	0.60	0.57+0.04	0.035	NS
14	0.46	0.58	0.60	0.57	0.18	0.58	0.54	0.48	0.58	0.57	0.54	0.52+0.13	0.185	NS
15	0.34	0.49	0.57	0.64	0.26	0.60	0.55	0.56	0.61	0.54	0.57	0.52+0.09	0.188	NS
16	0.52	0.60	0.63	0.62	0.52	0.60	0.51	0.48	0.60	0.58	0.61	0.57+0.11	0.064	NS
17	0.61	0.56	0.58	0.60	0.49	0.58	0.47	0.51	0.60	0.55	0.62	0.56+0.18	0.052	NS
18	0.58	0.60	0.53	0.57	0.45	0.61	0.56	0.61	0.54	0.57	0.60	0.57+0.07	0.040	NS
19	0.53	0.34	0.56	0.61	0.46	0.60	0.54	0.53	0.53	0.59	0.55	0.53+0.18	0.158	NS
20	0.31	0.58	0.49	0.57	0.52	0.63	0.51	0.58	0.61	0.59	0.58	0.54+0.21	0.216	NS
21	0.43	0.58	0.60	0.58	0.37	0.58	0.58	0.59	0.60	0.59	0.41	0.54+0.22	0.125	NS
22	0.47	0.62	0.57	0.52	0.58	0.56	0.55	0.55	0.56	0.60	0.58	0.56+0.09	0.141	NS
23	0.62	0.61	0.57	0.59	0.33	0.59	0.60	0.54	0.55	0.55	0.59	0.56+0.11	0.036	NS
24	0.57	0.53	0.60	0.60	0.41	0.55	0.63	0.62	0.53	0.57	0.48	0.55+0.15	0.070	NS
25	0.35	0.53	0.59	0.51	0.51	0.55	0.59	0.48	0.56	0.50	0.56	0.52+0.21	0.201	NS

Table 12. Proportion of accessions from different geographical regions and different clusters to form the core sample

Clusters	Geographical Regions				
	ISC	WAC	EAC	SCAC	Total
1	24	24	14	18	80
2	19	16	4	46	85
3	34	52	4	21	111
4	22	33	7	8	70
5	8	18	3	40	69
6	42	32	10	19	103
7	9	9	5	12	35
8	88	40	9	37	174
9	8	42	5	5	60
10	50	3	5	12	70
11	3	9	3	36	51
12	21	23	0	17	61
13	34	28	22	6	90
14	21	2	3	9	35
15	12	15	4	4	35
16	36	16	0	8	60
17	10	46	6	18	80
18	6	29	0	5	40
19	9	9	8	9	35
20	12	4	0	19	35
21	26	25	1	6	58
22	12	14	5	12	43
23	7	14	1	28	50
24	4	33	0	2	39
25	17	6	1	7	31
Total	**534**	**542**	**120**	**404**	**1600**
WC	**5585**	**5150**	**951**	**4377**	**16063**

ISC = Indian Subcontinent Collection; EAC = East African Collection; WAC = West African Collection; SCAC = South and Cetral African Collection; WC = World Collection; CC = Core Collection.

minimum core was later established through hierarchical cluster analysis on 1600 accessions. The minimum core consisted of 504 accessions (3.14% of world collection). The number of accessions selected from different clusters and geographical regions to form the minimum core is given in Table 13. Only a few accessions captured from East African region in the core sample as well as the minimum core.

4.3. Evaluation of the selection method used for the development of core collection

4.3.1 Frequency of accessions from each country of origin

The number and percentages of accessions from each country of origin in the world collection, core sample and minimum core are presented in Table 14, together with deviations from the world collection frequency. The Chi-square values were calculated over all the countries to compare the frequencies in the world collection. The stratification by country of origin was not significantly different from the world collection in core and minimum core. However, some countries represented proportionately more (Nigeria) or less (India, Namibia, Togo, Zimbabwe) accessions in the core sample, than in the world collection. Similarly, accessions from Burkina Faso, Namibia and Zimbabwe were represented more while Cameroon, Nigeria and Sudan was less represented in the minimum core. Significant reduction was observed for the accessions from Yemen. Only one accession was represented from this region in the minimum core. The percentages of accessions selected from different countries ranged from 0.20% for Yemen to 30.16% for India. These countries were represented the least and the most respectively, in the world collection as well as in core samples.

4.3.2. Quantitative characters

The mean, variance and range of eleven quantitative characters were compared among core sample, minimum core and world collection (Table 15). The Newman-Keul's post-hoc comparison for means indicated that for the core sample only characters, plant height (*Kharif* and *rabi*) and spike exsertion had significantly different mean values. All other characters except days to flowering (*kharif*), spike thickness and thousand grain weight had significantly different variance compared to those of world collection. The range for most of the characters was retained in the core sample. Comparison of minimum core with core

Table 13. Representation of pearl millet accessions from different geographical regions in various clusters for the minimum core

Clusters	ISC	WAC	EAC	SCAC	Total
1	1	16	0	5	22
2	0	17	0	15	32
3	9	9	1	18	37
4	12	6	2	7	27
5	9	2	2	5	18
6	1	18	0	8	27
7	15	7	3	5	30
8	15	10	0	5	30
9	15	0	0	1	16
10	14	0	0	3	17
11	5	2	0	1	8
12	1	6	1	1	9
13	6	3	1	2	12
14	9	0	0	0	9
15	6	5	0	0	11
16	0	9	0	0	9
17	0	5	1	5	11
18	0	5	1	8	14
19	6	1	7	5	19
20	14	4	3	6	27
21	13	5	1	3	22
22	2	14	2	7	25
23	0	13	0	14	27
24	5	11	2	15	33
25	5	3	0	4	12
MC	**163**	**171**	**27**	**143**	**504**
WC	**5585**	**5150**	**951**	**4377**	**16063**
CC	**534**	**542**	**120**	**404**	**1600**

ISC = Indian Subcontinent Collection; EAC = East African Collection; WAC = West African Collection; SCAC = South and Cetral African Collection; WC = World Collection; CC = Core Collection; MC = Minimum Core

Table 14. **Proportion of accessions from 25 countries of origin in the world collection of pearl millet germplasm and deviations in the core samples drawn with proportional strategy**

Country	World Collection		Core Subset			Minimum Core		
	COUNT	PERCENT	COUNT	PERCENT	DEVIATION	COUNT	PERCENT	DEVIATION
Benin	46	0.29	6	0.38	-0.09	2	0.40	-0.02
Botswana	86	0.54	7	0.44	0.10	3	0.60	-0.16
Burkina Faso	841	5.24	77	4.81	0.43	39	7.74	-2.93
Cameroon	918	5.71	91	5.69	0.02	21	4.17	1.52
Central Africa	146	0.91	11	0.69	0.22	5	0.99	-0.30
Chad	38	0.24	9	0.56	-0.32	3	0.60	-0.04
Ghana	290	1.81	39	2.44	-0.63	10	1.98	0.46
India	5373	33.45	522	32.63	0.82	152	30.16	2.47
Kenya	75	0.47	9	0.56	-0.09	4	0.79	-0.23
Malawi	292	1.82	35	2.19	-0.37	9	1.79	0.40
Mali	1024	6.37	110	6.89	-0.52	30	5.95	0.94
Namibia	1052	6.55	76	4.75	1.80	38	7.54	-2.79
Niger	904	5.63	91	5.69	-0.06	32	6.35	-0.66
Nigeria	1120	6.97	136	8.50	-1.53	37	7.34	1.16
Pakistan	151	0.94	12	0.75	0.19	5	0.99	-0.24
Senegal	349	2.17	36	2.25	-0.08	10	1.98	0.27
Sierre Leone	59	0.37	9	0.56	-0.19	3	0.60	-0.04
South Africa	138	0.86	13	0.81	0.05	3	0.60	0.21
Sudan	584	3.64	70	4.38	-0.74	13	2.58	1.80
Tanzania	467	2.91	54	3.38	-0.47	15	2.98	0.40
Togo	517	3.22	38	2.38	0.84	13	2.58	-0.20
Uganda	86	0.54	11	0.69	-0.15	8	1.59	-0.90
Yemen	61	0.38	6	0.38	0.00	1	0.20	0.18
Zambia	92	0.57	13	0.81	-0.24	5	0.99	-0.18
Zimbabwe	1354	8.43	119	7.44	0.99	43	8.53	-1.09
Total	16063		1600			504		
Chi-Square					2.62 NS			7.26 NS

*,** P < 0.05 and P < 0.01, respectively; NS = P > 0.05
Chi-square value = Departure from proportions in world collection

Table 15. Mean, range and variance for quantitative characters in the world pearl millet collection, core sample and minimum core

	World Collection				Core Collection						Minimum Core Sample					
Characters	Mean	SE	Range	arianc	Mean	SE	N-K Test	Range	Variance	Levene's Test	Mean	SE	N-K Test	Range	Variance	Levene's Test
Days to flowering (K)	76.0	0.20	33-159	550.8	75.22	0.58	NS	33-157	531.12	NS	76.1	1.03	NS	38-152	531.0	NS
Days to flowering (R)	71.0	0.10	32-138	140.9	71.0	0.31	NS	32-125	148.69	**	71.56	0.55	NS	43-122	148.8	NS
Plant Height (K)	245.0	0.50	30-480	4163.0	243.4	1.66	**	30-450	4426.77	**	245.3	3.14	NS	30-450	4971.3	**
Plant Height (R)	160.0	0.30	25-425	1434.3	159	1.00	**	25-365	1579.6	**	160.6	1.79	NS	25-300	1576.3	NS
Number of productive tillers	3.0	0.01	1-19	3.3	2.68	0.04	NS	1-15	2.78	**	2.63	0.08	**	1-15	10.5	**
Spike Exertion	3.6	0.10	-45-29	45.1	3.23	0.17	**	-32-22	48.92	**	3.36	0.32	NS	-25-19	51.1	**
Spike Length (K)	28.2	0.10	5-120	119.1	28.25	0.29	NS	6-120	135.81	**	28.55	0.53	NS	9-90	139.4	**
Spike Length (R)	25.7	0.10	4-125	120.4	25.73	0.28	NS	5-115	128.03	**	25.77	0.49	NS	9-80	118.2	**
Spike Thickness (K)	24.0	0.03	8-58	25.3	23.84	0.12	NS	10-55	24.63	NS	24.22	0.24	**	10-55	27.9	**
Spike Thickness (R)	23.3	0.04	9-61	28.3	23.2	0.14	NS	10-52	28:84	NS	23.73	0.25	**	12-50	31.7	**
1000-Grain Weight	8.6	0.02	1.5-21.3	5.6	8.59	0.06	NS	2.9-19.	5.62	NS	8.82	0.12	**	3.0-19.3	6.7	**

.K = *Kharif* or Rainy season; R = *Rabi* or Post-rainy season

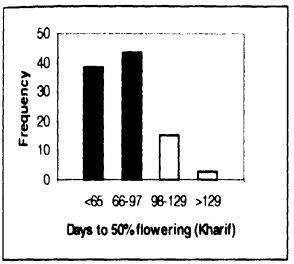
sample showed significantly different mean values for number of productive tillers, spike thickness and thousand-grain weight. The variance was significantly different for almost all the characters studied. Similarity in mean, variance and range for characters was observed in minimum core and world collection (Table 15).

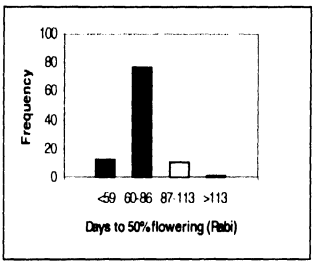
The frequency distribution of ten quantitative characters of the world collection was calculated by classifying the distribution into four different frequency classes taking quartiles into consideration. The Chi-square analysis was done to compare the percent deviations from the world collection frequency in the core sample as well as minimum core (Fig 3,4,5,6 and Table 16). The core sample selected proportionately produced non-significant deviations in the frequency distribution for all the quantitative characters, while the minimum core derived from the core sample, produced significant deviant distributions for most of the characters. The frequency of least represented classes was increased in both core sample and minimum core as described in figures. For characters like days to flowering and spike exsertion, an increase in already well represented class in both cores were observed.

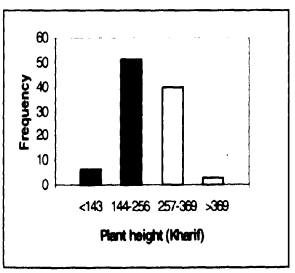
4.3.3. Qualitative characters

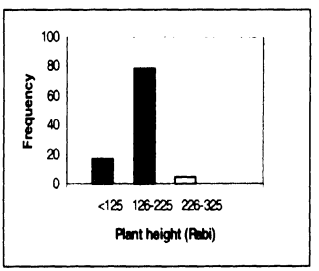
Phenotypic frequencies for five qualitative traits of the core sample and minimum core were compared with the world collection frequencies (Table 17). Most of the classes of these characters appeared at relatively higher frequencies in the whole collection as well the samples. No significant deviation was observed for any of the five qualitative characters in the core sample and in the minimum core. Similarly, the frequency of major classes of qualitative character frequency were compared to frequency of world collection and no significant difference has been observed for any of the characters in both core sample and minimum core (Table 18). The rare variants also showed non-significant deviations when compared to world collection frequencies (Table 19), though there was significant representation of these variants both in world collection and core samples. The pattern of distribution for qualitative characters among world collection and two core samples was significant using Mann-Whitney test.

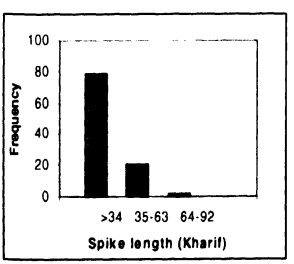
Fig 3. Frequency distribution for ten quantitative characters in the entire landrace collection of pearl millet











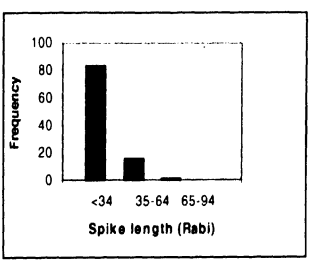
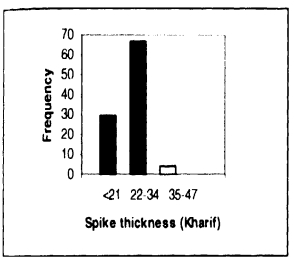
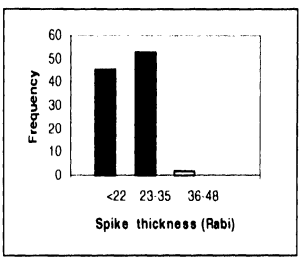
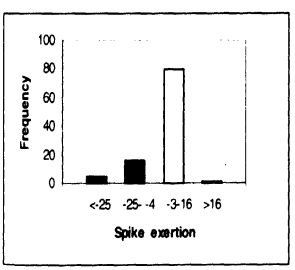


Fig 3. Contd.


Frequency
70
60
50
40
30
20
10
0
<21 22-34 35-47
Spike thickness (Kharif)


Frequency
60
50
40
30
20
10
0
<22 23-35 36-48
Spike thickness (Rabi)


Frequency
100
80
60
40
20
0
<-25 -25- -4 -3-16 >16
Spike exertion

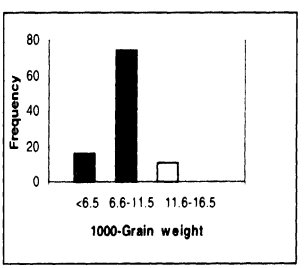
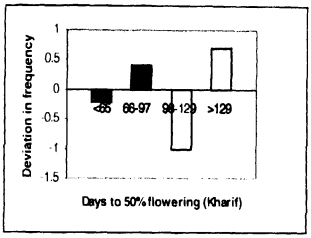
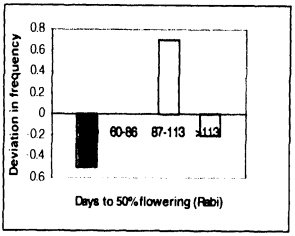
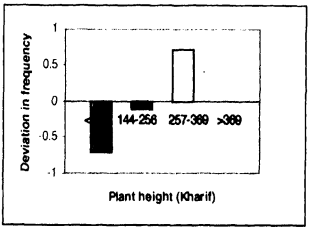
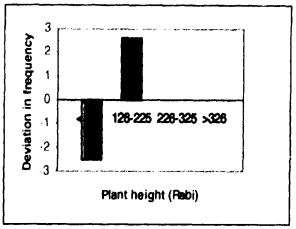

Frequency
80
60
40
20
0
<6.5 6.6-11.5 11.6-16.5
1000-Grain weight

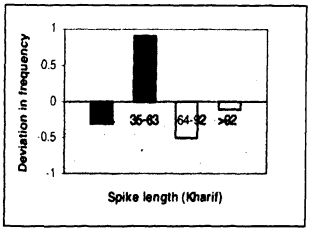
Fig 4. Frequency deviations in ten quantitative characters in the core sample when compared to entire pearl millet landrace collection











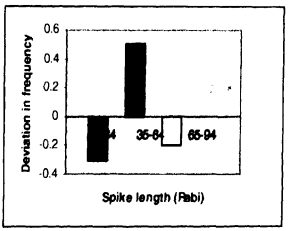
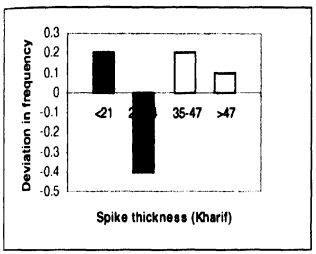
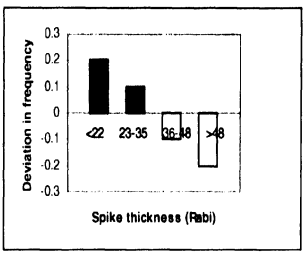
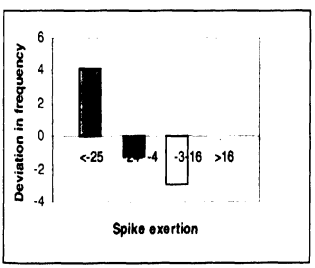


Fig 4. Contd.


Deviation in frequency
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
<21
35-47
>47
Spike thickness (Kharif)


Deviation in frequency
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
<22
23-35
>48
Spike thickness (Rabi)


Deviation in frequency
6
4
2
0
-2
-4
<-25
-3-16
>16
Spike exertion

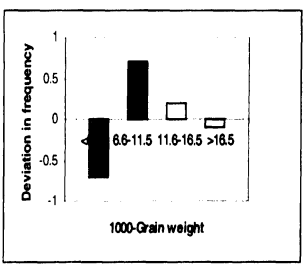
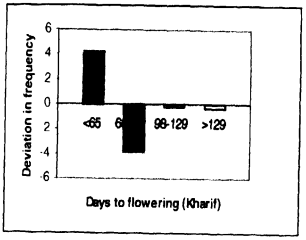
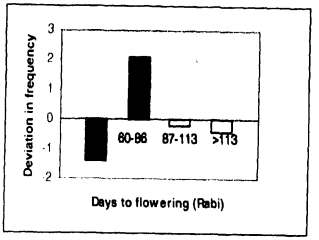
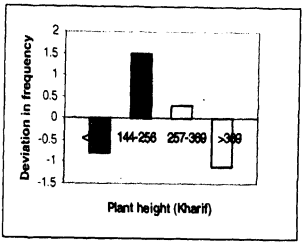
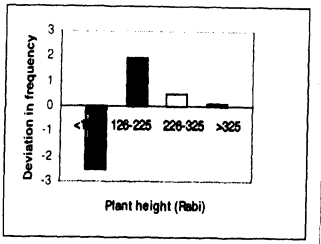

Deviation in frequency
1
0.5
0
-0.5
-1
6.6-11.5
11.6-16.5
>16.5
1000-Grain weight

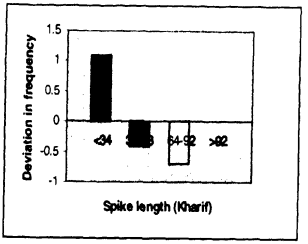
Fig 5. Frequency deviations in ten quantitative characters in the minimum core when compared to entire pearl millet landrace collection











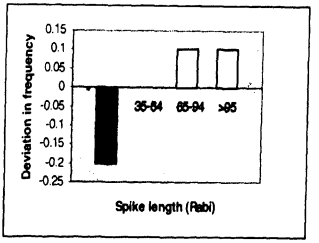
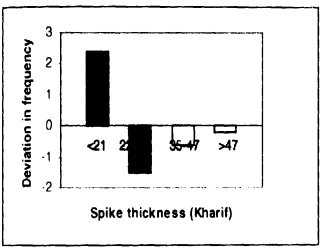
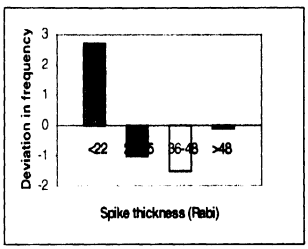
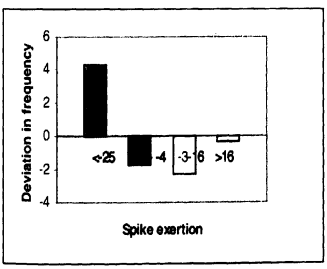


Fig 5. Contd.


Deviation in frequency
3
2
1
0
-1
-2
<21
35-47
>47
Spike thickness (Kharif)


Deviation in frequency
3
2
1
0
-1
-2
<22
36-48
>48
Spike thickness (Rabi)


Deviation in frequency
6
4
2
0
-2
-4
<25
-4
-3-6
>16
Spike exertion

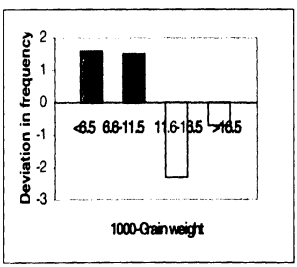
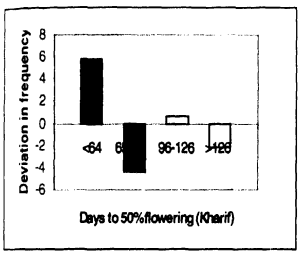
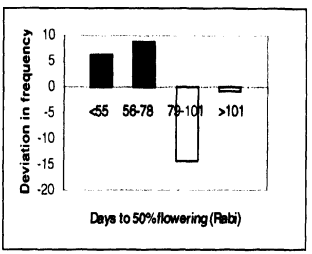
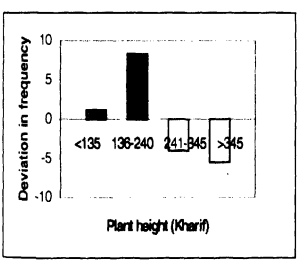
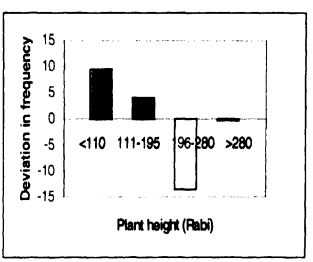

Deviation in frequency
2
1
0
-1
-2
-3
<6.5
6.6-11.5
11.6-13.5
>16.5
1000-Grain weight

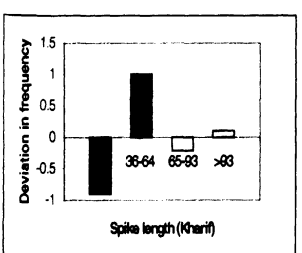
Fig 6. **Frequency deviations in ten quantitative characters in the minimum core when compared to the core sample drawn from the entire collection.**











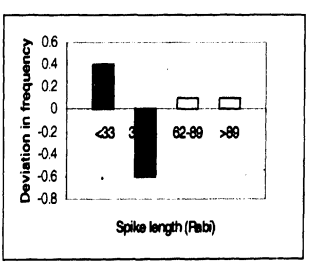
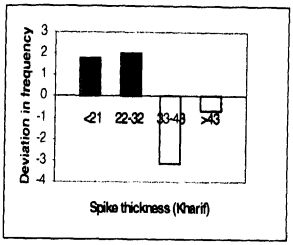
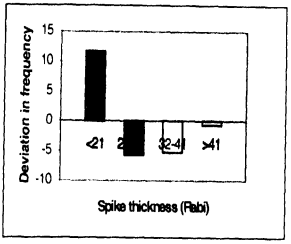
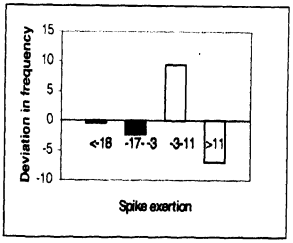


Fig 6. Contd.


Deviation in frequency
3
2
1
0
-1
-2
-3
-4
<21
22-32
33-43
>43
Spike thickness (Kharif)


Deviation in frequency
15
10
5
0
-5
-10
<21
>41
Spike thickness (Rabi)


Deviation in frequency
15
10
5
0
-5
-10
<-18
-17--3
-3-11
>11
Spike exertion

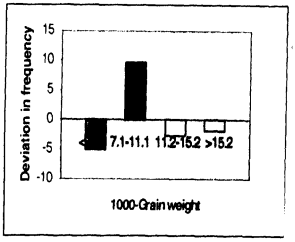

Deviation in frequency
15
10
5
0
-5
-10
7.1-11.1
>15.2
1000-Grain weight

Table 16. Proportion of accessions in different frequency classes of the ten quantitative characters in the world collection and percentage deviation in the core sampl and minimum core.The Chi-square values are departure from proportions in t world collection

	World Collection		Core Subset			Minimum Core		
Character	Count	Percent	Count	Percent	Deviation	Count	Percent	Deviation
DFLK								
1	5958	38.3	594	38.5	-0.2	162	32.7	5.8
2	6735	43.3	662	42.9	0.4	234	47.3	-4.4
3	2379	15.3	251	16.3	-1	77	15.6	0.7
4	471	3	36	2.3	0.7	22	4.4	-2.1
Chi-square					0.27			3.27
					NS			NS
DFLR								
1	1957	12.5	203	13	-0.5	33	6.8	6.2
2	11959	76.6	1194	76.6	0	330	67.8	8.8
3	1609	10.3	150	9.6	0.7	116	23.8	-14.2
4	86	0.6	12	0.8	-0.2	8	1.6	-0.8
Chi square					0.14			25.77
					NS			**
PHTK								
1	967	6	107	6.7	-0.7	27	5.6	1.1
2	8179	51.1	816	51.2	-0.1	208	42.8	8.4
3	6394	39.9	625	39.2	0.7	210	43.2	-4
4	468	2.9	46	2.9	0	41	8.4	-5.5
Chi square					0.09			12.4
					NS			*
PHTR								
1	2687	17.1	306	19.6	-2.5	48	10	9.6
2	12319	78.2	1180	75.6	2.6	342	71.5	4.1
3	727	4.6	72	4.6	0	86	18	-13.4
4	14	0.1	2	0.1	0	2	0.4	-0.3
Chi square					0.46			44.85
					NS			**
SPE								
1	682	4.3	3	0.2	4.1	3	0.6	-0.4
2	2557	16.1	271	17.3	-1.2	95	19.5	-2.2
3	12517	79	1282	81.9	-2.9	354	72.5	9.4
4	80	0.5	10	0.5	0	36	7.4	-6.9
Chi square					4.1			97.38
					NS			**
SPLK								
1	12479	78.3	1248	78.6	-0.3	391	79.5	-0.9
2	3212	20.2	306	19.3	0.9	90	18.3	1
3	243	1.5	32	2	-0.5	11	2.2	-0.2
4	5	0	1	0.1	-0.1	.	.	0.1
Chi square					0.21			0.18
					NS			NS

Table 16. Contd.

Character	World Collection Count	Percent	Core Subset Count	Percent	Deviation	Minimum Core Count	Percent	Deviation
SPLR								
1	13340	83.3	1334	83.6	-0.3	416	83.2	0.4
2	2471	15.5	239	15	0.5	78	15.6	-0.6
3	177	1.1	21	1.3	-0.2	6	1.2	0.1
4	9	0.1	2	0.1	0	.	.	0.1
Chi square					0.06			0.13
					NS			**NS**
SPTK								
1	4550	29.5	450	29.3	0.2	131	27.5	1.8
2	10232	66.4	1024	66.8	-0.4	309	64.8	2
3	613	4	58	3.8	0.2	33	6.9	-3.1
4	26	0.2	2	0.1	0.1	4	0.8	-0.7
Chi square					0.06			7.6
					NS			**NS**
SPTR								
1	6634	45.4	652	45.2	0.2	162	33.6	11.6
2	7694	52.7	758	52.6	0.1	281	58.3	-5.7
3	260	1.8	28	1.9	-0.1	34	7.1	-5.2
4	21	0.1	4	0.3	-0.2	5	1	-0.7
Chi square					0.41			19.46
					NS			**
TGW								
1	2428	15.5	254	16.2	-0.7	105	21.2	-5
2	11567	73.9	1145	73.2	0.7	316	63.7	9.5
3	1623	10.4	160	10.2	0.2	64	12.9	-2.7
4	43	0.3	6	0.4	-0.1	11	2.2	-1.8
Chi square					0.07			11.58
					NS			*

*,**P < 0.05 and P < 0.01, respectively; NS = P > 0.05

DFLK = Days to flowering (*Kharif*); DFLR = Days to flowering (*Rabi*);
PHTK = Plant height (*Kharif*); PHTR = Plant height (*Rabi*); NPT = Number of productive tillers;
SPLK = Spike length (*Kharif*); SPLR = Spike length (*Rabi*); SPTK = Spike thickness (*Kharif*);
SPTK = Spike thickness (*Rabi*); TGW = Thousand grain weight

Table 17. Proportion of accessions for five qualitative characters in the world pearl millet collection and percentage deviation in the core sample and minimum core. The Chi square values are departure from proportions in the world collection

Character	World Collection		Core sample			Minimum Core		
	Count	Percent	Count	Percent	Deviation	Count	Percent	Deviation
Bristle Length								
Short	3267	20.34	317	19.81	0.53	94	18.65	1.16
Medium	10266	63.91	1052	65.75	-1.84	344	68.25	-2.5
Long	2111	13.14	222	13.88	-0.74	62	12.3	1.58
Chi square					0.11			0.34
					NS			**NS**
Spike Shape								
Cylindrical	4415	28.03	471	29.44	-1.41	137	27.18	2.26
Conical	1316	8.35	145	9.06	-0.71	41	8.13	0.93
Spindle	319	2.03	31	1.94	0.09	10	1.98	-0.04
Club	58	0.37	3	0.19	0.18	1	0.2	-0.01
Candle	7469	47.42	747	46.69	0.73	244	48.41	-1.72
Dumb-bell	254	1.61	26	1.63	-0.02	10	1.98	-0.35
Lanceolate	1821	11.56	163	10.19	1.37	53	10.52	-0.33
Oblanceolate	53	0.34	8	0.5	-0.16	3	0.6	-0.1
Globose	47	0.3	6	0.38	-0.08	5	0.99	-0.61
Chi square					0.49			1.42
					NS			**NS**
Spike Density								
Loose	9155	56.99	866	54.13	2.86	271	53.77	0.36
Medium	6277	39.08	661	41.31	-2.23	212	42.06	-0.75
Compact	626	3.9	73	4.56	-0.66	21	4.17	0.39
Chi square					0.38			0.05
					NS			**NS**
Grain Shape								
Obovate	3161	19.68	321	20.06	-0.38	81	16.07	3.99
Oblanceolate	3318	20.66	341	21.31	-0.65	100	19.84	1.47
Elliptical	1991	12.4	192	12	0.4	74	14.68	-2.68
Hexagonal	2872	17.88	292	18.25	-0.37	91	18.06	0.19
Globular	4717	29.37	454	28.38	0.99	158	31.35	-2.97
Chi square					0.08			1.81
					NS			**NS**
Grain Color								
Ivory	8474	53.5	827	51.78	1.72	290	57.54	-5.76
Cream	5664	35.76	603	37.76	-2	171	33.93	3.83
Yellow	954	6.02	86	5.39	0.63	23	4.56	0.83
Grey	285	1.8	29	1.82	-0.02	4	0.79	1.03
Deep Grey	280	1.77	33	2.07	-0.3	9	1.79	0.28
Grey Brown	95	0.6	7	0.44	0.16	2	0.4	0.04
Brown	60	0.38	7	0.44	-0.06	4	0.79	-0.35
Purple	14	0.09	3	0.19	-0.1	.	.	0.19
Purplish Black	13	0.08	2	0.13	-0.05	1	0.2	-0.07
Chi square					0.48			2.29
					NS			**NS**

NS = P > 0.05

Table 18. Proportion of accessions for major qualitative character classes in the world pearl millet collection and percentage deviation in the core sample and minimum core. The Chi square values are departure from proportions in the world collection

Character	World Collection		Core sample			Minimum Core		
	Count	Percent	Count	Percent	Deviation	Count	Percent	Deviation
Bristle Length								
Short	3267	20.34	317	19.81	0.53	94	18.65	1.16
Medium	10266	63.91	1052	65.75	-1.84	344	68.25	-2.50
Long	2111	13.14	222	13.88	-0.74	62	12.30	1.58
Chi square					0.11			0.34
					NS			**NS**
Spike Shape								
Cylindrical	4415	28.03	471	29.44	-1.41	137	27.18	2.26
Conical	1316	8.35	145	9.06	-0.71	41	8.13	0.93
Candle	7469	47.42	747	46.69	0.73	244	48.41	-1.72
Lanceolate	1821	11.56	163	10.19	1.37	53	10.52	-0.33
Chi square					0.31			0.34
					NS			**NS**
Spike Density								
Loose	9155	56.99	866	54.13	2.86	271	53.77	0.36
Medium	6277	39.08	661	41.31	-2.23	212	42.06	-0.75
Compact	626	3.90	73	4.56	-0.66	21	4.17	0.39
Chi square					0.38			0.05
					NS			**NS**
Grain Shape								
Obovate	3161	19.68	321	20.06	-0.38	81	16.07	3.99
Oblanceolat	3318	20.66	341	21.31	-0.65	100	19.84	1.47
Elliptical	1991	12.40	192	12.00	0.40	74	14.68	-2.68
Hexagonal	2872	17.88	292	18.25	-0.37	91	18.06	0.19
Globular	4717	29.37	454	28.38	0.99	158	31.35	-2.97
Chi square					0.08			1.81
					NS			**NS**
Grain Color								
Ivory	8474	53.5	827	51.78	1.72	290	57.54	-5.76
Cream	5664	35.76	603	37.76	-2.00	171	33.93	3.83
Yellow	954	6.02	86	5.39	0.63	23	4.56	0.83
Chi square					0.23			1.16
					NS			**NS**

NS = P > 0.05

Table 19. Proportion of accessions for rare qualitative character classes in the world pearl millet collection and percentage deviation in the core sample and minimum core. The Chi square values are departure from proportions in the world collection

Character	World Collection		Core sample			Minimum Core		
	Count	Percent	Count	Percent	Deviatio	Count	Percent	Deviation
Spike Shape								
Spindle	319	2.03	31	1.94	0.09	10	1.98	-0.04
Club	58	0.37	3	0.19	0.18	1	0.20	-0.01
Dumb-bell	254	1.61	26	1.63	-0.02	10	1.98	-0.35
Oblanceolate	53	0.34	8	0.50	-0.16	3	0.60	-0.10
Globose	47	0.30	6	0.38	-0.08	5	0.99	-0.61
Chi square					0.19			1.08
					NS			NS
Grain Color								
Grey	285	1.80	29	1.82	-0.02	4	0.79	1.03
Deep Grey	280	1.77	33	2.07	-0.30	9	1.79	0.28
Grey Brown	95	0.60	7	0.44	0.16	2	0.4	0.04
Brown	60	0.38	7	0.44	-0.06	4	0.79	-0.35
Purple	14	0.09	3	0.19	-0.10	.	.	0.19
Purplish Black	13	0.08	2	0.13	-0.05	1	0.20	-0.07
Chi square					0.25			1.13
					NS			NS

NS = P > 0.05

The Chi-square analysis over all the 69 subclasses, over 29 subclasses of qualitative characters and over 40 subclasses of quantitative characters showed that the core sample was not different from the expected frequency of the world collection. However, the minimum core derived from the core sample showed significantly different frequency distribution over quantitative characters and over all 69 characters when compared with the core (Table 20). The minimum core was more representative of the world collection for most of the characters studied.

Table 20. Comparison of frequency distribution over all qualitative and quantitative characters between world pearl millet collection, core sample and minimum core

Sample	Qualitative Characters Chi square (d.f. = 28)	Quantitative characters Chi square (d.f. = 39)	Over all characters Chi square (d.f. = 68)
Core	1.54 NS	5.96 NS	7.50 NS
Minimum core	5.91 NS	222.62 **	228.53 **

P < 0.05 and P < 0.01; NS = P > 0.05

4.3.4. Assessment of diversity

The range ratio for eleven quantitative traits was calculated comparing the range of each character in the world collection to that of core sample and minimum core (Table 21). More than eighty-percent range of the world collection was retained for almost all the characters tested both in the core sample and minimum core. The mean range retained was also more than 75% in both the cores. Some characters namely, plant height, number of productive tillers and spike thickness in the minimum core retained the full range of the core sample.

The diversity in eleven quantitative and five qualitative characters of the core sample and minimum core was estimated by Shannon-Weaver diversity index (Table 22). The mean diversity indices of the samples were compared with that of the world collection. No significant differences were observed for both the samples over qualitative and

Table 21. Range ratio for quantitative characters in the core sample and minimum core in comparison to the world collection

Characters	Core	Minimum Core
Days to flowering (*kharif*)	0.98	0.92
Days to flowering (*rabi*)	0.88	0.85
Plant Height (*kharif*)	0.93	1.00
Plant Height (*rabi*)	0.85	0.81
Number of productive tillers	0.77	1.00
Spike Exertion	0.73	0.81
Spike Length (*kharif*)	0.99	0.71
Spike Length (*rabi*)	0.91	0.65
Spike Thickness (*kharif*)	0.90	1.00
Spike Thickness (*rabi*)	0.79	0.93
1000-Grain Weight	0.83	0.99
Mean range ratio+S.E	0.78+0.039	0.88+0.041
Per cent of characters with R < 0.8	27.27	18.18

R = Range ratio.

Table 22. Shannon-Weaver diversity indices for eleven quantitative and five qualitative characters in world pearl millet collection, core sample and minimum core

Characters	World Collection	Core Sample	Minimum Core
Days to flowering (K)	0.59	0.60	0.57
Days to flowering (R)	0.60	0.61	0.61
Plant Height (K)	0.63	0.63	0.62
Plant Height (R)	0.63	0.62	0.62
Number of productive tillers	0.41	0.44	0.41
Spike Exertion	0.59	0.58	0.56
Spike Length (K)	0.56	0.55	0.56
Spike Length (R)	0.56	0.57	0.55
Spike Thickness (K)	0.61	0.62	0.61
Spike Thickness (R)	0.58	0.62	0.62
Thousand Grain Weight	0.62	0.62	0.61
Quantitative Mean+SE	0.58+0.097	0.59+0.056	0.58+0.058
Spike Shape	0.59	0.59	0.60
Spike Density	0.65	0.66	0.65
Bristle Length	0.47	0.47	0.44
Grain Shape	0.68	0.68	0.68
Grain Color	0.59	0.60	0.59
Qualitative Mean+SE	0.59+0.063	0.60+0.056	0.59+0.049
Total Mean+SE	0.58+0.086	0.59+0.056	0.58+0.055
Chi square (d.f = 15)		0.006	0.008
		NS	**NS**

K = *Kharif*; R = *Rabi*; NS = P > 0.05

quantitative characters independently, and over all the characters. Significantly less diversity was observed for number of productive tillers and bristle length in all the world collection as well as cores.

4.4. Influence of genotype x season interaction on characterization of core collection

The minimum core sample was tested in two different seasons to study the effect of season on characterization and determine the diversity present within and between accessions,. The analysis of variance for the data recorded on the check (HHB 67) over blocks was carried out for each of the seven quantitative traits separately in both seasons to test the block effects. Significant block effects were observed for most of the characters except number of productive tillers for 1998-99 *rabi* data. For 1999-2000 *kharif* data, significant block effects were recorded for days to flowering, plant height, spike length and spike thickness. No significant effects were recorded for checks, confirming the uniformity across the blocks.

Also, the data collected in two different seasons were subjected to a linear mixed model, using REML analysis to study the effect of environment on morphological characters. A random model was used to estimate variance components for the main season, main accession or genotype, and season x accession interaction effects. The effect of method for selecting the minimum core was estimated by Wald test for fixed effects. Mean values of 60-plant observation per accession for the seven quantitative characters of the minimum core over two different seasons were analyzed in the models described above. The estimated variance components along with the standard error for each random term for the characters are listed in Table 23. The main effect of season and variation due to accessions were not statistically significant for most of the characters except days to flowering and plant height. However, the variation due to interaction effect was significant for three characters, namely days to flowering, plant height and spike length.

The Wald statistic for fixed effect of the minimum core on variation of seven quantitative characters is presented in Table 24. The variation due to the core effect was

Table 23. REML (Residual Maximum Likelihood) variance component analysis for different quantitative characters of pearl millet core accessions

Sources of variation	Days to flowering			Plant Height		
	Variance Components		SE	Variance Components		SE
Season	166.17	**	235.86	4225.7	**	5979.6
Accessions	61.34	**	9.27	352.8	**	52.4
Season * Accessions	210.25	**	15.88	975.3	**	82.1

Sources of variation	Number of productive tillers			Spike Exertion		
	Variance Components		SE	Variance Components		SE
Season	0.0000	NS	0.0025	6.982	NS	9.914
Accessions	0.2942	NS	0.2127	7.933	NS	1.085
Season * Accessions	0.5641	NS	0.2852	5.921	NS	1.262

Sources of variation	Spike Length			Spike Thickness		
	Variance Components		SE	Variance Components		SE
Season	6.346	NS	9.081	0.795	NS	1.147
Accessions	3.421	NS	0.568	4.087	NS	0.569
Season * Accessions	34.588	**	1.855	4.179	NS	0.69

Sources of variation	1000-Grain Weight		
	Variance Components		SE
Season	0.0000	NS	0.0129
Accessions	0.2745	NS	0.7548
Season * Accessions	4.3124	NS	11.065

*,** $P < 0.05$ and $P < 0.01$; NS = $P > 0.05$

significant for all the characters studied except number of productive tillers and 1000-grain weight.

Table 24. REML variance component analysis on different quantitative characters to study the effect of method used to define a core

Character	Wald statistic for core (d.f. = 1)
Days to flowering	1353.6 **
Plant height	5329.0 **
Number of productive tillers	0.0 NS
Spike exsertion	478.2 **
Spike length	112.9 **
Spike thickness	925.0 **
1000-grain weight	5.0 NS

*,** $P < 0.05$ and $P < 0.01$; NS = $P > 0.05$

The Spearman rank order correlation was estimated on seven quantitative characters for the data collected in two different seasons (1998-99 *rabi* and 1999-2000 *kharif*). The correlation was also estimated between the existing characterization data and two season data separately. Inter season correlation was observed for all the characters studied. Significant correlation was observed for all quantitative characters except plant height and number of productive tillers between existing characterization data and core *kharif* data. The Spearman rank order correlation was not estimated for number of productive tillers and spike exsertion in the *rabi* season due to lack of data in the existing characterization database (Table 25).

4.5 Assessment of diversity

The mean, variance and range for seven quantitative characters were compared between the core sample developed from characterization data and the core data evaluated during two seasons. Systematic drifts (all downward) were observed for all the characters in the core evaluation data during both the seasons (Table 26). The student's t-test for homogeneity of means and F-test for variances resulted in significant differences for the

Table 25. Spearman rank order correlations for quantitative characters among two seasons (*kharif* and *rabi*) and the characterization data (data collected in different years from 1974-94)

Characters	Characterization data with core *kharif* data	Characterization data with core *rabi* data	Core kharif data with core rabi data
Days to flowering	0.55 **	0.46 **	0.24 **
Plant Height	0.06 NS	0.77 **	0.23 **
Number of productive tillers	0.15 NS	-	0.36 **
Spike Exertion	0.88 **	-	0.30 **
Spike Length	0.36 **	0.68 **	0.36 **
Spike Thickness	0.64 **	0.57 **	0.28 **
1000-Grain Weight	0.76 **	0.65 **	0.25 **

*,** $P < 0.05$ and $P < 0.01$, respectively; NS = $P > 0.05$

Table 26. Mean, range and variance for seven quantitative characters in the core accessions evaluated during two seasons and core established from available data

Characters	Core established from available characterization data			Core evaluated during *rabi* 1998-99			Core evaluated during *kharif 1999-2000*		
	Mean	Range	Variance	Mean	Range	Variance	Mean	Range	Variance
Days to flowering	76.0	38-152	531.03	91.0	66-150	103.19	72.0	43-138	435.90
Plant height	245.0	30-450	4966.94	145.0	35-200	607.46	235.0	45-365	2158.58
Number of productive tillers	3.0	1-15	3.17	2.0	1-4	0.20	2.0	1-5	0.21
Spike exsertion	3.4	-25-19	51.08	7.0	-6-17	12.92	3.2	-11-20	15.76
Spike length	28.5	9-90	139.41	19.8	5-45	34.74	23.4	5-53	44.72
Spike thickness	24.2	10-55	27.92	19.4	11-31	7.29	20.7	11-34	10.38
1000-grain weight	8.8	3-19	6.67	7.6	3-15	4.79	7.2	2-14	3.92

mean and variance for all the characters except for spike exsertion. Similar results were obtained when comparison of mean and variance was carried out with Newman-Keul's and Levene's test. Large variations were observed in the range of core evaluation data for all the characters accounting for large fraction of diversity for each of these characters.

The range ratio for seven characters for the minimum core data collected during two different seasons were compared with the minimum core sample established from available characterization data (Table 27). There was only 61% similarity among the *kharif* data and available data while 70% similarity was recorded with the *rabi* data over all characters studied. There were very few characters that retained more than 70% range ratio in both the seasons.

Shannon-Weaver diversity indices were used to estimate the diversity present in the core set after evaluation in two different seasons (Table 28). The means of Shannon-Weaver diversity indices of the core collection data evaluated in two seasons were compared to the diversity indices of the core sample derived from available characterization data. Chi-square analysis was done to estimate the significance of the differences. Data was recorded on seven quantitative and 14 qualitative characters during *rabi* and only seven quantitative characters during *kharif*. No significant differences were observed for diversity indices for the core evaluation data and available data. Very less diversity was recorded for grain shape during evaluation of core accessions.

During evaluation of the minimum core, data was recorded on some simply inherited qualitative characters namely, anther colour, early vigour, fodder yield potential, yield potential, node and blade pubescence, and node, blade and sheath pigmentation. No information was available on these characters in the characterization data of the available pearl millet database. Therefore, further comparison was not possible for these characters. The frequency distribution for different classes of these characters is presented in Table 29 along with the Shannon-Weaver diversity indices. Significant diversity was not observed for any of these characters except early vigor, yield potential and fodder yield potential of

Table 27. Range ratio for quantitative characters of the selected minimum core evaluated during two different seasons in comparison to the existing core data

Characters	Core evaluated during *kharif*	Core evaluated during *rabi*
Days to flowering	0.82	1.06
Plant Height	0.76	0.60
Number of productive tillers	0.21	-
Spike exertion	0.69	-
Spike Length	0.59	0.57
Spike Thickness	0.50	0.51
1000-Grain Weight	0.69	0.77
Mean Range Ratio+SE	0.61+0.057	0.70+0.063
Percent of ranges with R<0.70	42.86	60.0

R = Range ratio

Table 28. Shannon-Weaver diversity indices to compare core data collected in two different seasons and the existing evaluation data

Characters	Core developed from existing data	Core evaluated in *rabi* season	Core evaluated in *kharif* season
Qualitative			
Bristle length	0.44	0.56	.
Spike shape	0.60	0.32	.
Spike density	0.65	0.70	.
Grain shape	0.68	0.17	.
Grain color	0.59	0.44	.
Mean+SE	0.59+0.05	0.29+0.06	.
Quantitative			
Days to flowering	0.57	0.62	0.55
Plant height	0.62	0.61	0.61
Number of productive tillers	0.41	0.38	0.37
Spike exertion	0.56	0.64	0.63
Spike length	0.56	0.60	0.61
Spike thickness	0.61	0.63	0.63
1000-Grain weight	0.61	0.61	0.60
Mean+SE	0.58+0.06	0.58+0.05	0.57+0.07
Total Mean+SE	0.58+0.05	0.38+0.05	0.57+0.07
Chi square (d.f = 11 and 6, respectively)		0.61 NS	0.02 NS

NS = P > 0.05

Table 29. Frequency distribution of fourteen qualitative characters along with their diversity indices.

Characters	Descriptor States	Frequency	Per cent
Anther color	**Yellow**	447	88.69
	Purple	43	8.5
	Cream	14	2.8
		SDI	0.18
Node Pigmentation	**Non-Pigmented**	400	79.37
	Pigmented	104	20.63
		SDI	0.22
Blade Pigmentation	**Non-Pigmented**	497	98.61
	Pigmented	7	1.39
		SDI	0.03
Sheath Pigmentation	**Non-Pigmented**	498	98.8
	Pigmented	5	1.2
		SDI	0.03
Node Pubescence	**Non-Pubescent**	450	89.29
	Pubescent	54	10.71
		SDI	0.15
Sheath Pubescence	**Non-Pubescent**	411	81.55
	Pubescent	93	18.5
		SDI	0.21
Early Vigour	**Early**	53	10.5
	Medium	357	70.8
	Late	94	18.7
		SDI	0.62
Fodder Yield Potential	**Low**	8	1.6
	Medium	254	50.4
	High	242	48.1
		SDI	0.47

Table 29. Contd.

Characters	Descriptor States	Frequency	Per cent
Yield Potential	Low	23	4.6
	Medium	409	81.2
	High	71	14.1
		SDI	0.46
Grain Shape	Obovate	41	8.1
	Hexagonal	11	2.2
	Globular	452	89.7
		SDI	0.17
Grain Color	Cream-Yellow	91	18.1
	Grey-Deep Grey	400	79.4
	Brown	12	2.4
	Purplish Black	1	0.2
		SDI	0.44
Bristle Length	Short	475	94.2
	Medium	3	0.6
	Long	6	1.2
		SDI	0.56
Spike Shape	Cylindrical	55	10.9
	Candle	377	74.8
	Lanceolate	71	14.1
	Globose	1	0.2
		SDI	0.32
Spike Density	Loose	164	32.5
	Intermediate	266	52.8
	Compact	74	14.7
		SDI	0.32

Table 30. Correlation coefficients among quantitative characters observed in the core accessions in different environment

Character	Year	Plant height	Number of productive tillers	Spike exsertion	Spike length	Spike thickness	1000-grain weight
Days to flowering	1974-94	0.470**	-0.086	-0.172**	0.067	-0.099*	0.135**
	1998-99	0.563**	-0.058	-0.381**	0.109**	-0.123**	-0.099*
	1999-00	0.526**	-0.491**	-0.643**	0.447**	-0.028	-0.543**
Plant height	1974-94		-0.213**	-0.103**	0.385**	0.169**	0.224**
	1998-99		0.223**	-0.227**	0.413**	0.026	0.183**
	1999-00		-0.493**	-0.230**	0.553**	0.063	-0.214**
Number of productive tillers	1974-94			0.217**	-0.321**	-0.284**	-0.169**
	1998-99			0.223**	-0.316**	-0.358**	-0.145**
	1999-00			0.450**	-0.495**	-0.352**	0.125**
Spike exsertion	1974-94				-0.384**	-0.237**	-0.037
	1998-99				-0.565**	-0.205**	0.156**
	1999-00				-0.593**	-0.093**	0.470**
Spike length	1974-94					0.158**	0.109**
	1998-99					-0.056	0.006
	1999-00					-0.117**	-0.248**
Spike thickness	1974-94						0.237**
	1998-99						0.335**
	1999-00						0.317**

* $P < 0.05$ and 0.01

the crop. Further assessment of diversity indices for five other qualitative characters recorded during evaluation of core revealed very less diversity for grain shape.

Simple correlation coefficients were also calculated among seven quantitative characters for different years when the pearl millet germplasm was evaluated (1974-94) and the core evaluation during 1998-99 *rabi* and 1999-2000 *kharif*. Significant phenotypic associations were observed for most of the characters (Table 30).

Wide variation was observed for days to flowering, plant height, maturity, canopy spread of the crop, anther color, spike shape and thickness, presence and absence of bristles, grain shape and color; polymorphism through RFLP was also observed and are presented in different plates (Plate 1 to 24).

4.6. Application of RFLP markers for diversity assessment

4.6.1. Polymorphism detected by various probe-enzyme combinations

The genomic DNA isolated from different plants of ten accessions were digested with *Eco*RI, *Eco*RV, *Hind*III and *Dra*I and hybridized with the nine RFLP probes polymorphic for downy mildew resistance and five probes polymorphic for drought tolerance to study the diversity within and between the selected accessions.

Fourteen probe-enzyme combinations yielded 51 bands for each of the ten accessions (Table 31). Number of bands per probe-enzyme combination ranged from one (PSM 648-*Dra*I) to seven in PSM 858-*Dra*I combination. Among the different bands produced by RFLP probes, four (28.6%) were common among all the accessions, one (7.14%) was unique to a single accession (IP 3890), and rest were shared by two or more accessions (Table 32 and Fig 7). The banding patterns resolved in the autoradiographs for the 14 probe-enzyme combinations, between plants within an accession was represented by different shades (Fig 7). Within the yellow shades, polymorphism marked due to the presence of a band was represented with red shade while blue shade was used to represent the absence of a band resulting in polymorphism. The distribution of rare bands (specific to either plant within accession or among accessions) was not uniform. However, the

Plate 1: Variation for anther color observed in the pearl millet core accessions

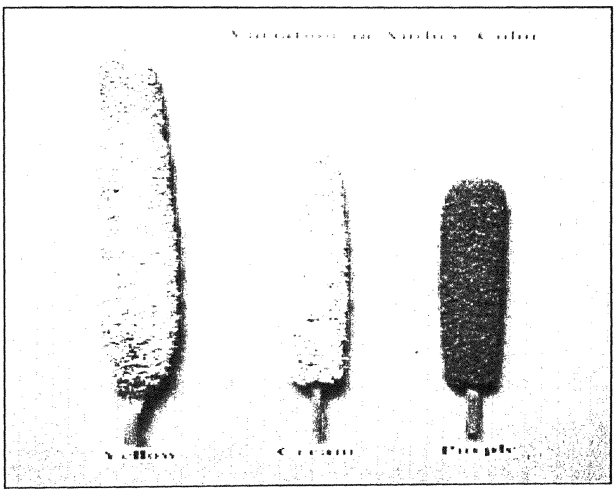


Plate 2: Variation observed for spike and grain characters within a core

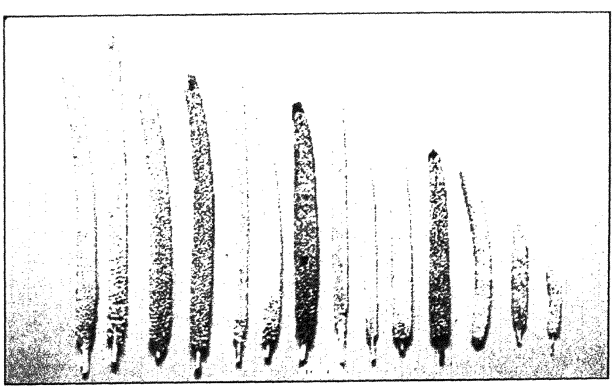


Plate 3: Variation observed for the presence and absence of bristles on the spikes and bristle length within an accession

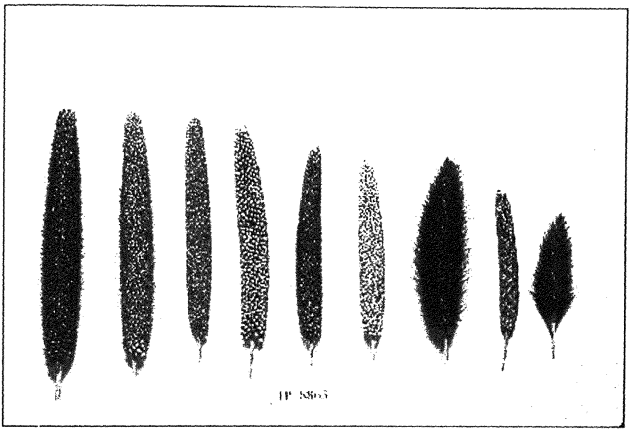


Plate 4: Variation observed for spike characters within a core accession

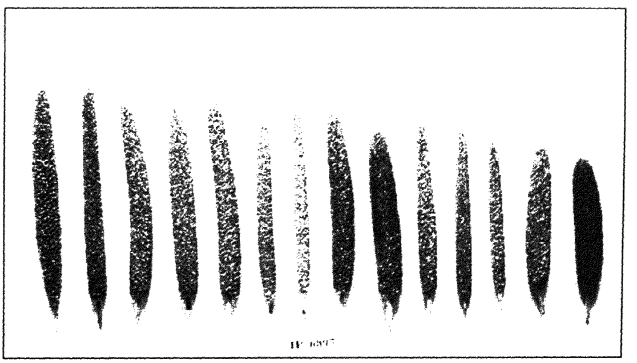


Plate 5: Variation for days to flowering, plant height, number of tillers, canopy spread in the pearl millet core

Plate 6: Pearl millet core accessions representing variability for days to flowering, plant height, maturity, spike length, presence or absence of bristles, and grain

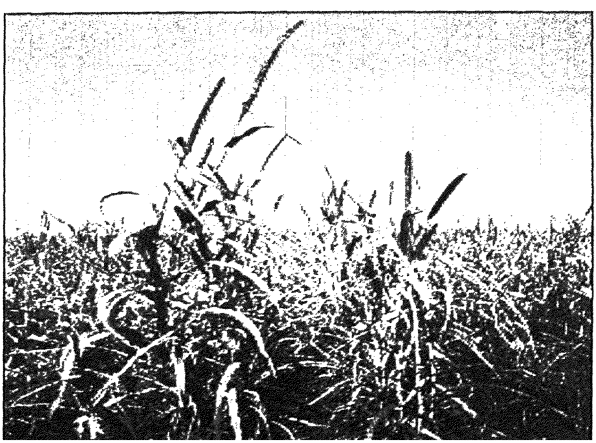


Plate 7: Variation observed for spike length, shape, thickness and grain color within an accession

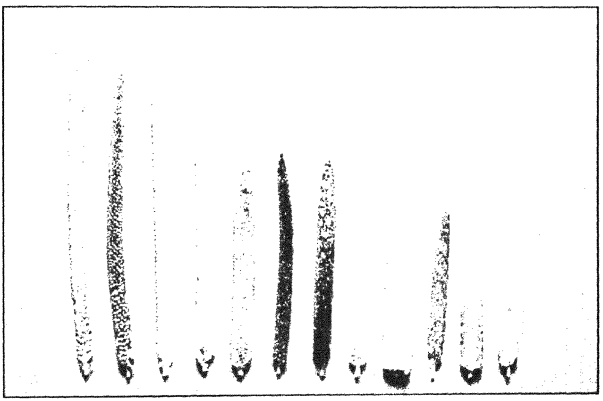


Plate 8: Variation observed for spike length, shape, thickness, density and grain color within an accession

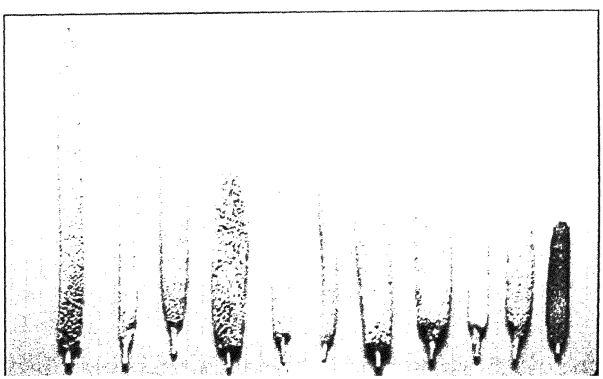


Plate 9: Variation for spike and grain characters in pearl millet entire collection

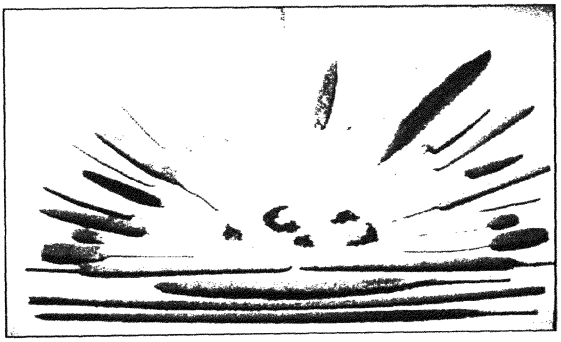


Plate 10: Variation for spike and grain characters in the pearl millet core collection

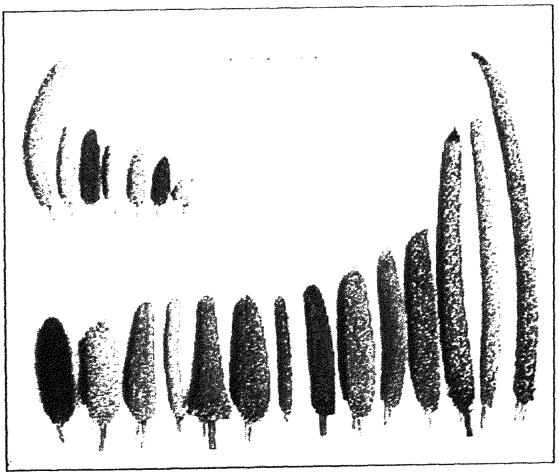


Plate 11: Pearl millet core accessions evaluated during two different seasons (*Rabi*, 1998-99 and *Kharif*, 1999-2000)

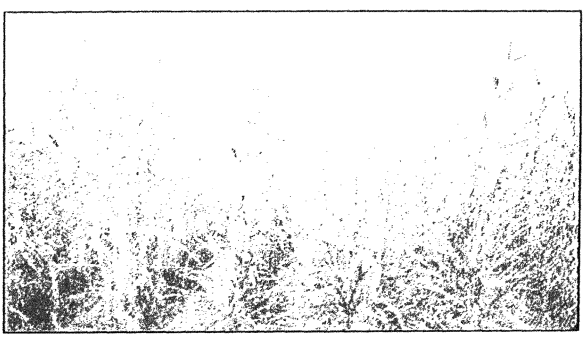


Plate 12: Polymorphism observed in accession (IP 3890) with RFLP probe (PSM 214) and enzyme *Dra* I

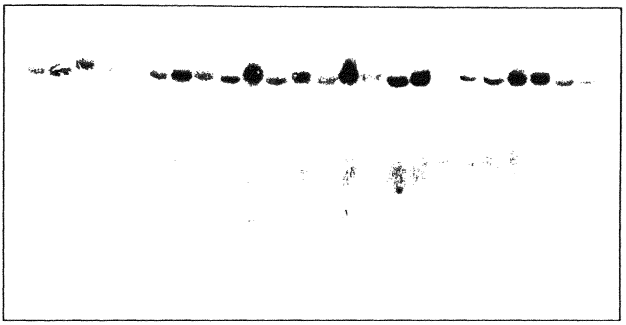


Plate 13: Polymorphism observed in accession (IP 4828) with RFLP probe (PSM 443) and enzyme *Dra*I

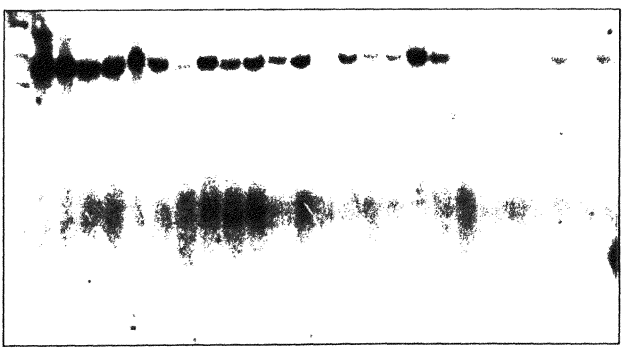
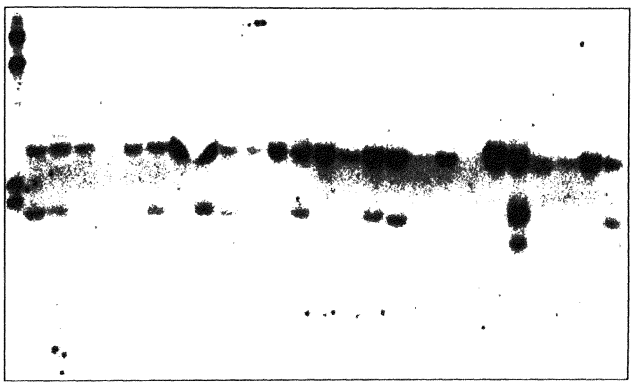


Plate 14: Polymorphism observed in accession (IP 4542) with RFLP probe (PSM 305) and enzyme *Dra* I



Polymorphism observed in accession (IP 3122) with RFLP probe (PSM 305) and enzyme *Dra* I

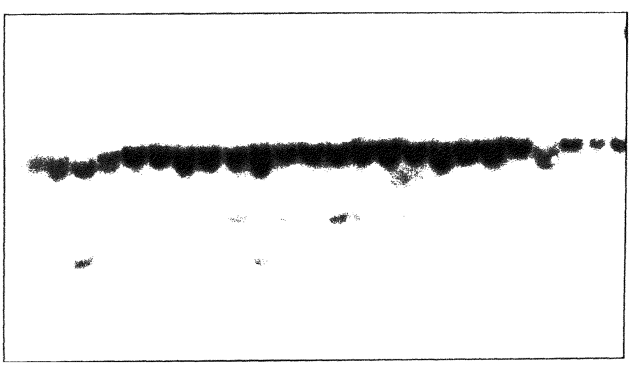
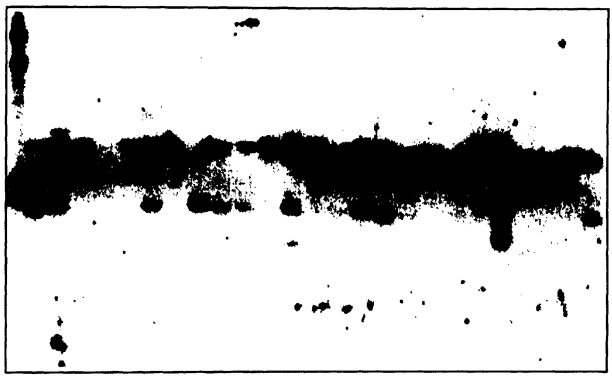


Plate 14: Polymorphism observed in accession (IP 4542) with RFLP probe (PSM 305) and enzyme *Dra* I



Polymorphism observed in accession (IP 3122) with RFLP probe (PSM 305) and enzyme *Dra* I

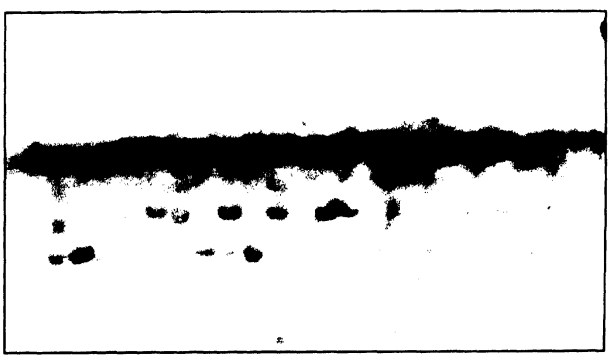
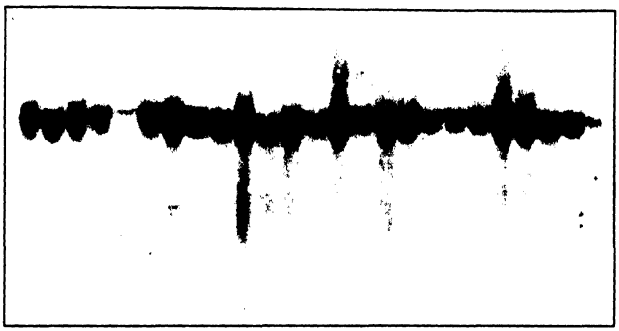
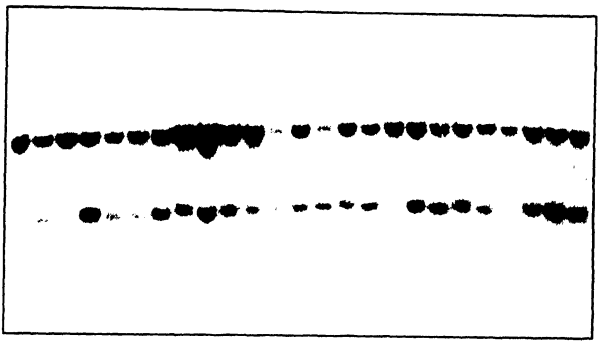


Plate 15: Polymorphism observed in accession (IP 3626) with RFLP probe (PSM 305) and enzyme *Dra*I



Polymorphism observed in accession (IP 3098) with RFLP probe (PSM 305) and enzyme *Dra*I

Plate 16: Polymorphism observed in accession (IP 3890) with RFLP probe (PSM 305) and enzyme *Dra*I



Polymorphism observed in accession (IP 10394) with RFLP probe (PSM 305) and enzyme *Dra*I

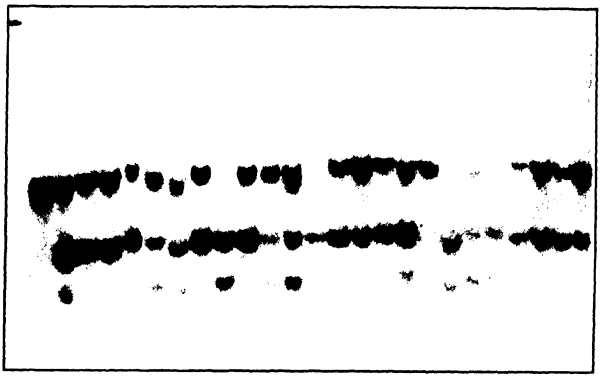
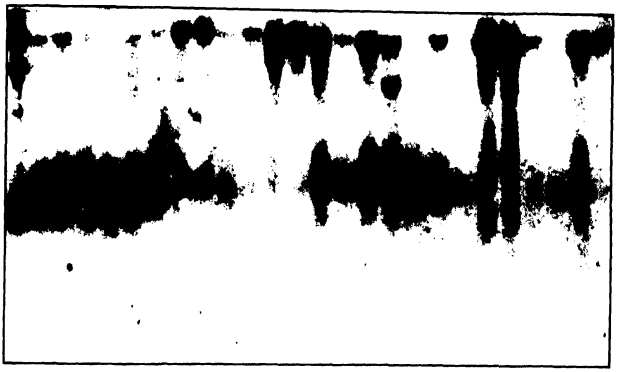


Plate 17: Polymorphism observed in accession (IP 4542) with RFLP probe (PSM 443) and enzyme *Dra*I



Polymorphism observed in accession (IP 8074) with RFLP probe (PSM 443) and enzyme *Dra*I

Plate 18: Polymorphism observed in accession (IP 3626) with RFLP probe (PSM 592) and enzyme *Hind* III

Polymorphism observed in accession (IP 4542) with RFLP probe (PSM 592) and enzyme *Hind* III

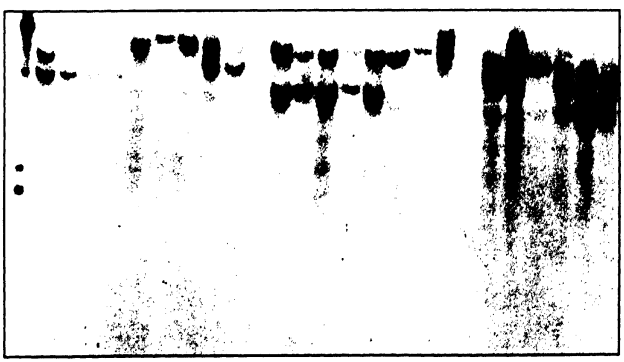


Plate 19: Polymorphism observed in accession (IP 3122) with RFLP probe (PSM 592) and enzyme *Hind* III

Polymorphism observed in accession (IP 10394) with RFLP probe (PSM 592) and enzyme *Hind*III

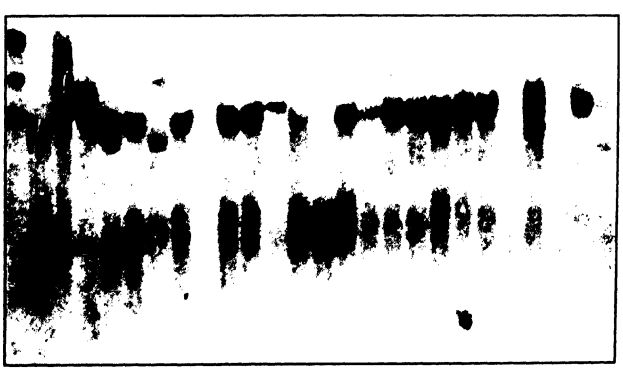


Plate 20: Polymorphism observed in accession (IP 3626) with RFLP probe (PSM 716) and enzyme *Hind* III

Polymorphism observed in accession (IP 8074) with RFLP probe (PSM 716) and enzyme *Hind*III

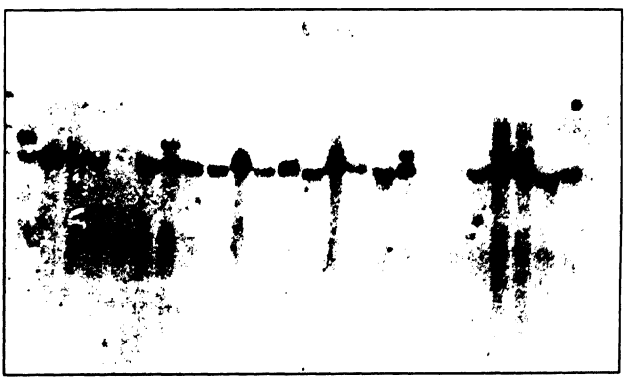
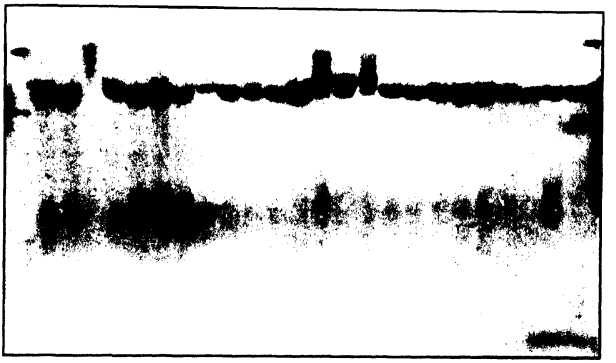


Plate 21: Polymorphism observed in accession (IP 8074) with RFLP probe (PSM 834) and enzyme *Hind* III



Polymorphism observed in accession (IP 4828) with RFLP probe (PSM 834) and enzyme *Hind*III

Plate 22: Polymorphism observed in accession (IP 8074) with RFLP probe (PSM 857) and enzyme *Hind* III

Polymorphism observed in accession (IP 4828) with RFLP probe (PSM 857) and enzyme *Hind* III

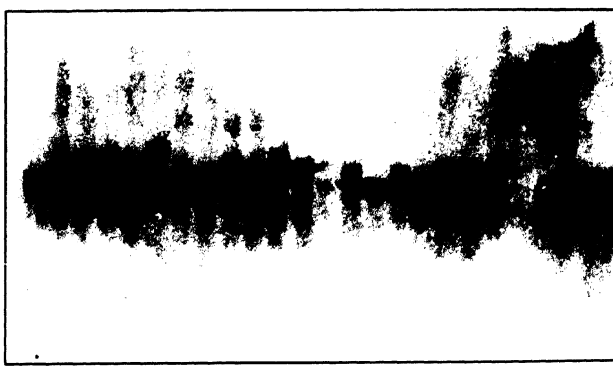
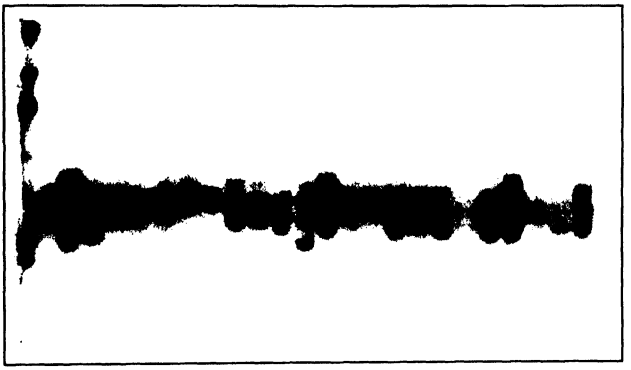


Plate 23: Polymorphism observed in accession (IP 4542) with RFLP probe (PSM 858) and enzyme *Dra*I



Polymorphism observed in accession (IP 3890) with RFLP probe (PSM 858) and enzyme *Dra*I

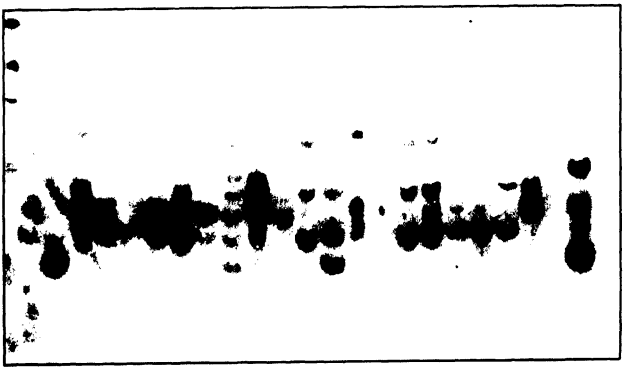


Plate 24: Polymorphism observed in accession (IP 10394) with RFLP probe (PSM 858) and enzyme *Dra*I

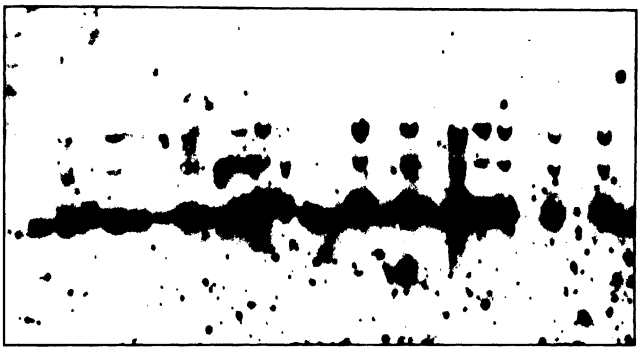


Plate 25: Polymorphism observed in accession (IP 8276) with RFLP probe (PSM 857) and enzyme *Hind*III

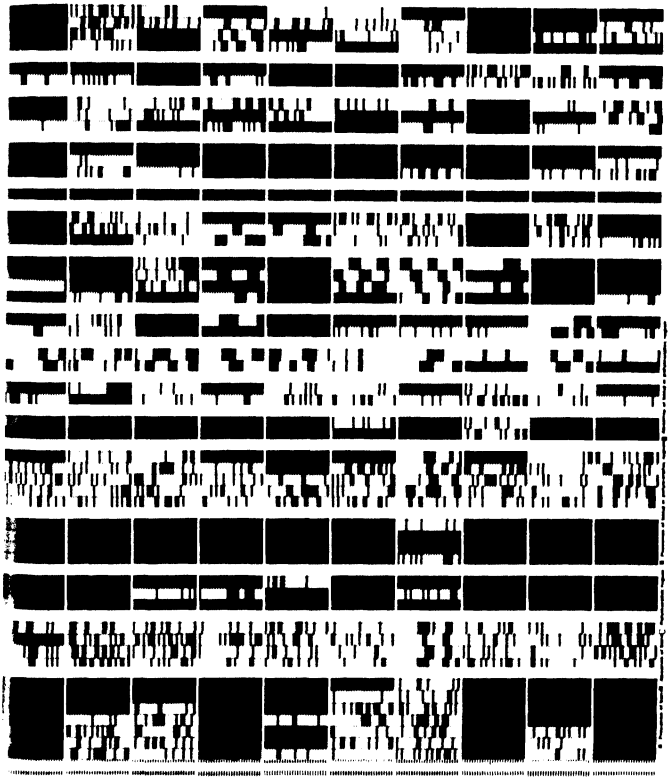


Table 31. Frequency of RFLP bands in the pearl millet core accessions with different probe-enzyme combinations used in the study

Probe-Enzyme	Total No. of Bands	No. of bands present in accessions									
		IP 3098	IP 3122	IP 4542	IP 13608	IP 8074	IP 4828	IP 3890	IP 8276	IP 10394	IP 3626
PSM 858 / *Dra*I	7	2	3	5	2	3	6	7	2	4	3
PSM 565 / *Hind*III	4	4	4	4	4	4	4	4	4	4	4
PSM 25 / *Eco*RV	3	1	1	2	2	3	1	2	1	1	1
PSM 321 / *Eco*RI	4	4	4	3	2	1	3	4	3	3	3
PSM 592 / *Hind*III	5	4	5	5	4	3	4	5	4	5	5
PSM 214 / *Dra*I	2	1	1	1	1	1	2	1	2	1	1
PSM 443 / *Dra*I	2	2	2	2	2	2	2	2	2	2	2
PSM 716 / *Hind*III	2	2	2	2	2	2	2	2	2	2	2
PSM 716 / *Dra*I	2	1	2	1	2	1	2	2	2	2	2
PSM 305 / *Hind*III	4	2	3	4	3	2	4	4	3	2	3
PSM 305 / *Dra*I	3	2	3	3	3	3	3	2	2	3	3
PSM 648 / *Dra*I	1	1	1	1	1	1	1	1	1	1	1
PSM 713 / *Eco*RI	3	2	3	3	3	2	2	3	2	3	3
PSM 857 / *Hind*III	3	2	3	3	3	3	2	3	2	3	3
PSM 618 / *Dra*I	2	2	2	1	1	1	1	2	2	2	2
PSM 834 / *Hind*III	4	2	4	4	4	4	4	3	1	2	3

Table 32. Frequency of common and rare bands in different plants within an accession

RFLP Bands	Total	IP 3098	IP 3122	IP 4542	IP 13608	IP 8074	IP 4828	IP 3890	IP 8276	IP 10394	IP 3626
					Core Accessions evaluated for molecular study						
PSM 858											
A	105	.	12	4	25	4	1	12	25	21	.
B	121	25	17	10	.	25	9	6	.	4	25
C	134	25	11	17	.	25	5	9	.	17	25
D	81	.	11	11	25	4	18	12	.	.	.
E	110	.	.	6	.	.	13	16	25	25	25
F	20	.	.	.	.	.	2	18	.	.	.
G	6	.	.	.	.	.	.	6	.	.	.
PSM 565											
A	88	6	14	8	5	18	4	10	11	5	7
B	95	11	13	10	3	7	21	7	6	3	14
C	161	25	11	16	12	17	9	20	19	19	13
D	81	8	9	9	4	11	3	16	7	4	10
PSM 25											
A	250	25	25	25	25	25	25	25	25	25	25
B	47	.	.	4	10	25	.	8	.	.	.
C	6	.	.	.	.	6	.	.	.	.	.
PSM 321											
A	235	25	25	25	25	25	25	10	25	25	25
B	225	25	25	25	25	.	25	25	25	25	25
C	200	25	25	25	.	.	25	25	25	25	25
D	54	25	25	.	.	.	.	4	.	.	.
PSM 592											
A	101	8	9	14	5	10	15	18	6	8	8
B	115	18	8	18	18	12	7	12	9	3	10
C	129	13	18	10	7	5	10	17	14	20	15
D	109	16	6	10	5	.	13	17	13	21	8
E	58	25	6	5	.	.	.	8	.	3	11
PSM 214											
A	40	.	.	.	.	.	25	.	15	.	.
B	226	25	25	25	25	25	7	25	19	25	25
PSM 443											
A	121	16	25	21	4	17	17	4	11	4	2
B	200	25	12	2	25	21	23	25	20	22	25
PSM 716											
A	181	14	20	11	19	15	21	15	25	16	25
B	110	18	12	12	14	12	3	7	4	7	21

Table 32. Contd.

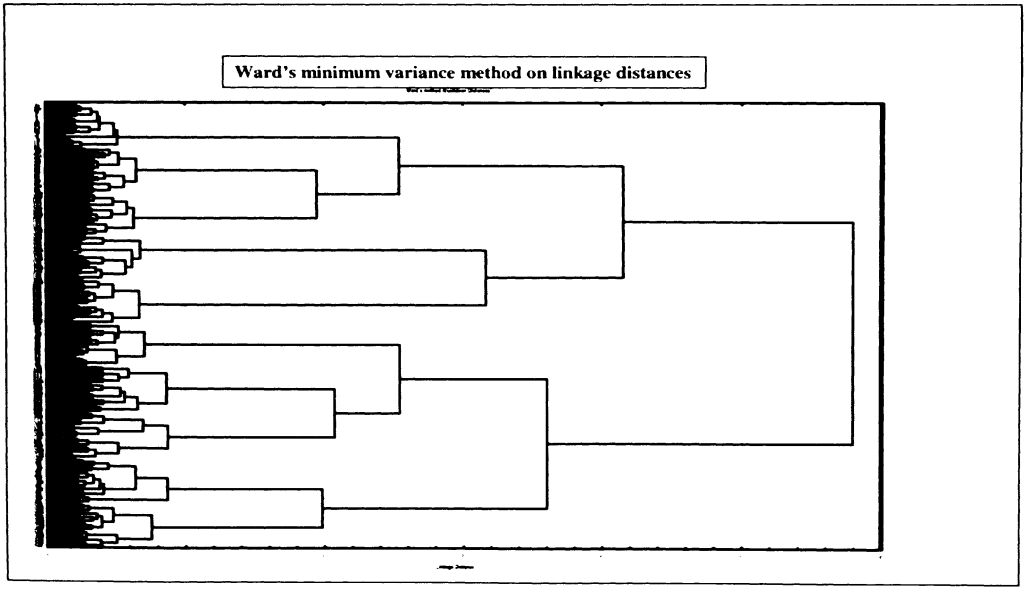
RFLP Bands	Total	IP 3098	IP 3122	IP 4542	IP 13608	IP 8074	IP 4828	IP 3890	IP 8276	IP 10394	IP 3626
					Core Accessions Evaluated for Molecular Study						
PSM 716											
A	154	21	20	25	25	25	7	5	4	12	10
B	141	.	12	.	11	.	25	25	25	18	25
PSM 305 / HindIII											
A	95	.	8	25	17	.	25	15	.	.	4
B	213	24	25	15	25	25	13	19	17	25	25
C	204	25	25	11	14	25	14	15	25	25	25
D	43	.	.	14	.	.	10	12	7	.	.
PSM 305 / DraI											
A	192	25	25	23	12	16	22	18	25	19	7
B	188	25	12	9	21	17	23	16	25	15	25
C	122	.	9	3	25	25	12	10	.	13	25
PSM 648 / DraI											
A	250	25	25	25	25	25	25	25	25	25	25
PSM 713											
A	53	.	17	2	.	.	.	14	.	5	15
B	201	25	23	25	.	25	25	25	25	25	3
C	250	25	25	25	25	25	25	25	25	25	25
PSM 857											
A	86	1	9	25	13	25	.	5	.	3	5
B	221	25	21	17	25	18	25	25	25	25	15
C	107	.	4	8	11	7	18	20	25	2	12
PSM 618											
A	65	5	9	.	7	.	.	12	8	13	11
B	229	25	25	25	25	25	25	25	10	19	25
PSM 834											
A	153	25	6	25	11	8	25	3	.	25	25
B	131	25	13	25	10	25	6	3	.	16	8
C	115	.	14	11	19	15	6	19	25	.	6
D	21	.	13	6	.	1	1	.	.	.	.

frequency of rare bands among plants within each accession was not able to detect as much variation as it was detected by a large variation among accessions (Fig 7).

4.6.2. Diversity among and within pearl millet accessions at molecular level

Distance matrix between and within accessions were calculated by percent similarity for each pair wise comparison, based on the presence or absence of bands across all the probe-enzyme combinations and across all accessions. Associations among 10 accessions as revealed by cluster analysis using Ward's minimum variance method for all the combinations were obtained based on the similarity indices generated from the distance matrix (Fig 8). It revealed 10 groups at 0.10% similarity level and 25 plants in each accession formed distinct clusters indicating within plant variation in each cluster. However between cluster variation separated each of these accessions into different groups. At 0.40% similarity, two distinct cluster were formed grouping five accessions in each cluster. IP 3098, IP 3122, IP 4542, IP 13608 and IP 8074 were grouped together in one cluster, and IP 4828, IP 3890, IP 8276, IP 10394 and IP 3626 in another cluster. The probe-enzyme combinations used in the study could not reveal high variation within accessions. In general, the accessions were different from each other at a high precision level of clustering.

Fig 8. Dendrogram based on Ward's minimum variance method and linkage distances for molecular data



5. Discussion

5. DISCUSSION

The need for reduction in the size of the germplasm collections was recognized for better utilization of germplasm. Several core collections that represent the complete genetic spectrum of the world collection have been established in several crops (Brown, *et al.*, 1987; Hamon and van Sloten, 1989; Loos and van Duin, 1989; von Bothmer *et al.*, 1990; Holbrook *et al.*, 1993; Diwan *et al.*, 1994; Basigulp *et al.*, 1995; Crossa *et al.*, 1995; Prasada Rao and Ramanath Rao, 1995; Tohme *et al.*, 1995; Mahajan *et al.*, 1996; Balfourier *et al.*, 1998; Bisht *et al.*, 1998 a,b; Ortiz *et al.*, 1998). Looking into the need for efficient handling of large pearl millet germplasm collection, it may be worthwhile to establish a core collection that may represent the world diversity for further improvement of the crop.

5.1. Studies on the base collection

The geographical distribution of the world collection of pearl millet at ICRISAT showed a disparity between the regions and the countries within each region. However, this represents a wide distribution of pearl millet growing areas throughout the world. The adaptability of pearl millet to dry conditions makes it a preferential crops system for the semi-arid tropics, explaining its distribution to the largest growing areas in Africa and in the Indian subcontinent. The diffused belt of origin, stretching from Senegal to Western Sudan (see Appendix) where cultivated pearl millet is widely distributed represents the greatest morphological diversity (Harlan and de Wet, 1971). The collection at ICRISAT represents the accessions largely collected from these regions. The unequal representation of landraces from different regions may be because of its domestication pathways and the use of improved varieties instead of landraces in the past. The poor representation of some countries in the entire collection could also be attributed to the difficulties in collection from the fields, the political situations in the prospected areas, the socio-economic conditions of the farmers, or even the countries undergoing civil wars. The world collection,

however, fits the evolutionary patterns, nearly representing the geographical distribution of the crop.

The global phenotypic diversity as measured by Shannon-Weaver diversity indices for different regions supports the hypothesis of Murty *et al.* (1967), Upadhyay and Murty (1970), and Tostain *et al.* (1987). Further, the diversity indices can be used for future sampling keeping a region as indicator. There is a wide range of diversity in the database maintained at ICRISAT for all the morpho-agronomic characters. The study revealed poor representation of good tillering plants with longer and thicker spikes bearing bold grains. The possible reason for such rare representation could be due to the inherent morphological characteristics relevant to adaptation to introduced climatic conditions.

In general, the present study highlighted the great potential of cultivated pearl millet for further research by the breeders mostly for hybrid development, as recently highlighted by Khairwal and Singh (1999); Rai and Virk (1999) and Talukdar *et al* (1999). The agro-morphological diversity present within and between regions supports the above statement. However, due to the large collection size (16063 accessions analyzed) the results might have been skewed only to certain tendencies that could not be well represented in the world collection (e.g. West African region). However, the level of phenotypic diversity obtained in the study provides a better knowledge of the germplasm maintained in the collection that would be of great value to breeding programs (Hanna and Rai, (1999) and Witcombe (1999).

The prospects for effectively utilizing geographic stratification could have been enhanced further with more precise geographic and ecological information. Passport data often lacks precise geographic and ecological information within regions or countries. With longitude and latitude information, there could have been an opportunity to use geographic information systems for ecogeographic mapping. There are cases when geographic or other forms of stratification are essential. Often the main source of data available for developing a core is the geographic and taxonomic information contained in the passport data. Stratification is also needed to simply reduce the voluminous data for

analysis to form a core collection. The size of a distance matrix is equal to (N-1) N / 2, where N is the number of accessions. Thus as N increases, distance matrix becomes very large and unwieldy. In this study, similar type of passport data that lacked precise geographic and ecological information was used to develop and evaluate the core.

5.2. Establishment of core collection

Several methods have been used by researchers to develop core collections, namely, in perennial relatives of *Glycine* (Brown *et al.* 1987), *Arachis hypogea* (Holbrook *et al.*, 1993), Tulipa (Loos and van Duin, 1989), annual medics (Diwan *et al.*, 1994, 1995) and perennial *Medicago* (Basigulp, *et al.*, 1995). There are many studies that suggest hierarchical stratification is the most effective strategy for developing a core collection by stratifying groups of accessions that are likely to be genetically similar (Spagnoletti-Zeuli and Qualset, 1987; Brown, 1989a). Some studies that suggests grouping based on country of origin of accessions as simple and effective way to partition variation in germplasm collections are also there (Peeters and Martinelli, 1989; Holbrook *et al.*, 1993). Brown (1989a) suggested that every subspecies or geographic region should be represented in the core, which should then be developed through stratified sampling methods.

5.2.1. Geographical stratification and cluster analysis

The world collection was first grouped based on country of origin into 25 groups to study the effect of stratification in developing a core collection of pearl millet. The geographic origin of the accessions, either the primary origin for natural populations, or secondary area of adaptation as in case of landraces, appears to be an efficient criterion for organizing the collection before sampling a core. However, the usual stratification by country of origin has a ideal theoretical basis in terms of natural selection pressure that rarely stops at administrative boundaries. Hence, the geographical groups should be further subjected to hierarchical cluster analysis. With the Ward's minimum variance method, clusters are merged at each step by minimizing the variance within clusters and thus maximizing variance among clusters (Romesburg, 1984). The similarity among the cluster is measured in terms of the Euclidean distance that is the ANOVA sum of squares between two clusters summed over all the variables. This resulted in 25 different clusters. All those accessions

that are predominantly late-maturing that grow over 300 cm tall and producing very long spikes have been grouped together. Similarly, the early maturing accessions that are ecologically and geographically similar have been grouped together. The accessions from Mali, Sudan, Senegal, Niger and Nigeria are grouped together except Namibia though it is a neighboring country. There must have been limited movement of people and material from this country. All the accessions from Sierra Leone with distinct day-length and temperature sensitive accessions with small, elliptical grains are grouped together. This confirmed the distinct forms of Sierra Leone which was classified into a distinct race, *Leonis* by Brunken *et al*, 1977. Considerable variation for time to flowering was observed among accessions within a country; it is as high as 36-150days in Ghana, 36-149 days in Togo, 35 to 146 days in Burkina Faso, 44 to 152 days in Benin and it is as low as 140 to 159 days in Sierra Leone, 36 to 73 days in Pakistan, and 47 to 87 days in Namibia (Table 4). The accessions from Mali, India, Central African Republic, and Burkina Faso are highly diverse for plant height. Diversity for productive tillers is high in germplasm from India, Sudan and Cameroon. The accessions from Ghana, Pakistan, India, Yemen and Togo have very high positive exsertion indicating its usefulness in crop improvement. Accessions with very long spikes are found from Niger, Nigeria and Namibia. The germplasm from Togo, Namibia and Ghana produces very thick spikes but most of the accessions from India have thin spikes. Thus, looking into the high variability between accessions, the stratification into different clusters was done using multivariate approach. The use of hierarchical approach to split the world collection into smaller and smaller groups was also suggested by Hintum (1995). Further stratification using multivariate approach was suggested by Spagnoletti-Zeuli and Qualset (1993), Holbrook *et al.* (1993), Hintum (1995), Hamon *et al.* (1995), Skinner *et al.* (1999) and Holbrook (1999).

5.2.2. Evaluation of selection method for the development of core collection

The proportional sampling strategy suggested by Brown (1989b) was used to select ten-percent accessions randomly from each cluster to form the core sample of 1600 accessions. Further stratification of these 1600 accessions was carried out after validating the representability to get a minimum core of 504 accessions (3.14 percent of the entire collection). The core sample developed by proportional strategy showed significant

changes in the frequency distribution for countries of origin. The core represented significant similarity in frequencies from world collection that was evident for the most and least represented countries.

The mean, range and variance of quantitative characters of core sample were compared to world collection. It was observed that the core sample developed following proportional strategy, showed highly significant similarity to the world collection in mean, range and variance for almost all the characters. The minimum core was then compared with the core sample and world collection. Significant similarity was observed in the mean, range and variance of the minimum core with core sample as well as world collection for all the characters.

Significant effects on frequency distributions of ten quantitative characters were not observed in the core sample while the minimum core derived from the core sample produced significant deviant distributions for most of the characters. The deviation in frequencies of phenotypic classes was not in any particular direction. For some characters, the number of accessions were reduced in most frequent classes and increased in the least frequent classes, while for some other characters, showing significant deviations, opposite trend was observed. The deviations observed over all the classes of quantitative characters were also non-significant. Furthermore, the comparison of frequent and rare classes separately resulted in significant deviations in major classes in the minimum core.

The phenotypic frequencies of core sample and minimum core were compared with the world collection frequencies for five qualitative characters. Since only quantitative characters were used in cluster analysis for stratification, the qualitative characters will give the measure of sampling strategy on these characters. There was no significant deviation for any of the five qualitative characters in the core sample as well as the minimum core. The major frequency classes of qualitative character were compared to world collection frequency and significant differences were not observed for any of the characters in both core sample and minimum core. The rare variant frequencies when compared to world collection frequencies also showed non-significant deviations. The reason for non-

significant results through Mann-Whitney test may be attributed to the uniformity in pattern of distribution of different frequent classes of the characters in world collection, core and minimum core.

The deviation in frequency over all the 69 subclasses of qualitative and quantitative characters, over 29 subclasses of qualitative characters alone and over 40 subclasses of quantitative characters alone in the core sample was not different from the expected frequency of the world collection. However, the minimum core derived from the core sample showed significant differences in frequency distribution over quantitative characters and over all 69 characters.

More than eighty- percent of range for the characters of world collection was retained both in core sample and minimum core. The mean range ratio was also more than 75 percent in both core and minimum core. The percentage of characters with less than 0.80 range ratio was to the extent of 27 and 18 percent, respectively in core and minimum core. Same level of diversity was observed in the core sample and minimum core when diversity estimates of world collection were compared with that of the core samples using Shannon-Weaver diversity indices.

Thus, proportional sampling strategy used in the present study to develop a core sample was found effective in representing the world variation of the entire pearl millet collection. The sampling strategy was better as there was no significant differences in the mean, range and variance for any of the characters studied. The frequency distributions of both qualitative and quantitative characters were not affected by proportional sampling except some characters in the minimum core which may be due to over representation of some of the classes. The core sample also retained more than 80 percent range of the characters of the world collection by maintaining overall diversity of the world collection. Hence, the minimum core developed in the present study could be designated as the core sample for the entire pearl millet collection. Among different sampling strategies adopted in earlier studies to develop a core sample from the entire collection, proportional strategy gave better representation (Erskine and Muehlbeur, 1991; Schoen and Brown, 1993 and

1995; Galwey, 1995; Yonezava *et al.*, 1995; Bataillon *et al.*, 1996). The use of proportional strategy alone in developing core collections is also reported (Holbrook *et al.*, 1993; Basigulp *et al.*, 1995; Prasada Rao and Ramanatha Rao, 1995; Ortiz *et al.*, 1999).

While discussing the effect of sampling strategies used in the development of core collection in *Glycine tomentella* and *Hordeum vulgare*, Brown (1989b) reported that the number of accessions selected in proportion to the size of each group represented the original collection better if redundancy does not occur among the groups. Again, if rare variants occur in the largest groups, the proportional strategy identifies the best core. Similarly, proportional strategy adopted in the present study could identify a representable sample to form the core for world pearl millet collection.

5.3. Influence of genotype x season interaction on characterization of the established core

Accessions were grown in two different seasons during 1998-99 *rabi* and 1999-2000 *kharif* to study the effect of environment on phenotypic diversity of the selected minimum core developed by using proportional sampling strategy from pearl millet world collection. These two environmental conditions are typical of the semi-arid tropics. *Kharif* is the rainy season with longer day duration (June-October) while *rabi* is the dry season with short day lengths (November-April). During *kharif*, the day-length gradually decreases from 13.93 h in June to 12.47 h in October; and temperature varies from 34.1C in June to 20.2C in October. Similarly, during *rabi*, day-length variation varies from 12.1 h in November to 12.7 h in April. The mean temperature range between 10.6C to 36.3C. Two characters, days to flowering and plant height are greatly influenced due to seasonal differences that may be attributed to photoperiod and/or temperature sensitivity of the accessions. Thus, most of the accessions from West Africa (center of origin for pearl millet) flowered earlier during *rabi* than *kharif*, confirming the short-day nature of the crop. Similarly, increase in plant height was found to be associated with delayed flowering. Similar results have been reported in the study of Burton, 1965. Generally, the plant flowers early under 12 h photoperiod, however longer photoperiod (14-16 h) delays flowering affecting phenology of the crop (Mangat *et al* 1999).

Genetic diversity within accessions is widespread, particularly in outbreeding crops like pearl millet. The strategies for selecting a core collection shall be evaluated in terms of their success in capturing rare but widespread alleles that occur at low frequencies in a large proportion of accessions (Brown, 1989b). However, when the selection of a core collection is based on evaluation data, maintaining one state of each character for each accession it may be impossible to take account of within-accession variation. Taking this in account, records were taken on 60 plant observations for each core accession during evaluation of the established core, to represent the maximum variability within accession and to describe the effect due allogamous nature of the crop. Although high variability was observed within accessions for most of the characters studied but they also showed significant variability between accessions. Most of the characters could have been affected by the regeneration of accessions wherein systematic drifts among the characters might have occurred due to seasonal effects over the period (1974-94) on evaluation. The augmented design analysis revealed block effects for very few characters in both the seasons.

To study the effect of environment on stability of accessions of the core sample, the morphological data collected over two seasons, was analysed by estimating the variances due to effect of seasons, accessions and accession x season interaction in a linear mixed model using REML analysis. The main effect of season and variation due to accessions were not statistically significant for most of the characters except days to flowering and plant height. However, the variation due to interaction effect was significant for three characters, namely days to flowering, plant height and spike length. Due to phtoperiod and temperature sensitivity of the crop, these two characters are most affected by seasonal differences. The variation due to the core effect was significant for all the characters studied except number of productive tillers and 1000-grain weight.

The present study indicated that, the method used to select the core represented significant variability for most of the characters and there was significant influence of season on the accessions as expected. As mentioned earlier, pearl millet is photo and

thermo-sensitive crop so any fluctuation in day length and daily temperature has direct effect on the growth and development of the crop, this was well depicted in the performance of accessions during these two different seasons. However, inter season Spearman rank order correlation for the quantitative characters showed highly significant correlations for most of the characters in both seasons when compared with the core sample from the available data. Thus, it can be concluded that in general, there has been stability among most of the characters for the accessions grown over different seasons or environments at ICRISAT, Patancheru, India. Therefore, these characters can be further used in diversity studies. The diversity calculated for qualitative and quantitative characters also did not deviate significantly from the respective diversity calculated for the characters on existing data.

5.4. Application of molecular markers (RFLP) for diversity assessment

There are many advantages and disadvantages in assessment of diversity from agronomic data and those developed from molecular data. Agronomic characters may be immediately useful for selection of desirable genes or genotypes but, if quantitative, these will be subjected to strong genotype by environment interactions. This would not be expected with RFLP-based evaluations. Refinement of core may be simplified using molecular markers. The genetic distance of new accessions could also be compared with those of existing accessions without conducting extensive field evaluations. But unless linked to specific traits, the random genetic variation is not immediately useful for selection and breeding purpose. Regardless of the approach, agronomic or molecular, any effort to develop and test a core collection will reveal new and important information about the nature and needs of the world germplasm collection. Development of core collection based on passport data and then assessment of genetic diversity in various subgroups generated from passport data with isozymes and DNA markers is reported by Mumm and Dudley (1994). Thus, for the present study, RFLP markers were used to characterize a small sample of core accessions. The wide range of polymorphism displayed by RFLP markers, depending on species has been earlier reported by Nodari *et al.* (1992). Additional advantages of RFLPs include their better genome coverage and environmental stability. However, RFLP technology is cumbersome and costly, which effectively limits the sample size.

Fourteen different RFLP probes generated 51 polymorphic bands in 250 plants from 10 different accessions. There was a high within accession variability for most of the probes studied. A high percentage of rare alleles were also observed. Distance matrix was formed by calculating distances among accessions by each pair wise comparison based on absence and presence of bands. Dendrogram was constructed using Ward's minimum variance method, which differentiated the accessions into 10 distinct groups representing the between accession variation.

The clusters formed by RFLP analysis did not reveal any particular pattern of variation according to known morphological variability, however all the accessions were separated from each other with minimum of 0.20% dissimilarity. The plants within each accession and within each cluster were spread over revealing within variation. The molecular characterization in this study revealed, not surprisingly, very high within accession variability owing to the nature of the crop. However, the higher variability between accessions suggests that each of these accessions are distinct. The high degree of polymorphism in pearl millet detected through RFLP markers has been reported (Liu *et al.*, 1994).

This preliminary study demonstrated the utility of RFLP markers to assess the variation in pearl millet accessions as reported earlier in different crops such as maize (Messmer *et al.*, 1992), rice (Zhang *et al.*, 1992), barley (Zhang *et al.*, 1993), European barley (Melchinger *et al.*, 1994), oats (Moser and Lee, 1994), maize (Dubreuil and Charcosset, 1998), barley (Casas *et al.*, 1998) and *Paspalum* species (Jarret *et al.*, 1998).

Thus from the present study, it can be concluded that proportional sampling strategy adopted to identify a core collection from the pearl millet world collection was appropriate. The difference between the multivariate approach for stratification and hierarchical approach of classification by country of origin was not significant however, multivariate approach was a better option as it could capture many rare variants. Hierarchical stratification cannot deal with missing data, at least as implemented in

standard statistical packages, which require a complete, rectangular data matrix. Furthermore, exclusion of accessions with incomplete data is likely to reduce the diversity of the core collection considerably. This is taken care in multivariate approach wherein core collection is selected on the basis of several characters, using them to define clusters of accessions and then applying selection strategies to the clusters. High stability among most characters is observed at ICRISAT location revealing systematic drift, though accessions were grown over years in different environments. Molecular analysis also proved that RFLP markers were able to detect polymorphism present among and within pearl millet accessions.

The established core sample and the minimum core can be further explored for developing trait-based gene pools for easy identification of economically useful traits. The molecular study can be further extended to assess diversity among and within African core accessions as well as among accessions from all geographical regions. Multivariate statistical methods could be used to assign individual accessions on the basis of phenotypic data that have been previously recognized using a marker-based analysis of diversity. This approach would involve a dynamic analysis consisting of multiple rounds in each of which diversity will be analyzed at the molecular and morphological levels until a satisfactory grouping of accessions is achieved.

6. Summary

6. SUMMARY

The importance of crop genetic resources was realized with the expansion of breeding programs, as the success of these programs largely dependent on availability of wide genetic diversity. But extensive cultivation of high yielding varieties, crop substitution, recurrent drought and urbanization, the risk of erosion of traditional landrace diversity from farmer's fields increased. This became a concern for the collection and conservation of genetic resources and great emphasis was given for *ex situ* conservation. Collection and conservation of many germplasm though became invaluable assets for crop improvement, but the size of these collections became so large that major international germplasm banks are already facing the burden of properly conserving and characterizing such large collections. This restrained the effective evaluation of accessions and impeded better utilization of germplasm. The conservation of germplasm resources has little practical importance unless it is effectively used for crop improvement. Frankel and Brown (1984) have recognized the need for a reappraisal of plant genetic resources. They suggested the use of core collection, a representation of genetic diversity of a crop species and its relatives with minimum repetitiveness, as a solution to the hindrance in handling large collections. The concept of representative core collection received substantial attention of the scientists all over the world dealing with large germplasm collections. As a result of this realization, many projects are currently involved in developing core collections for cereals, legumes, fruits, vegetables and forage crops. Several differences are observed in many aspects of core collection development i.e. the criteria used to differentiate groups of accessions from which accessions are to be selected for the development of the core collection and also sampling strategies used. The application of selection methods vary with the genetic structure of crop species, reproductive pattern of the crop, pattern of acquisition and priorities of germplasm curators.

Pearl millet [*Pennisetum glaucum* (L) R Br.] is the sixth most important cereal, primarily grown for grain production. Presently, the genetic resources in pearl millet at ICRISAT consisted of 20,642 cultivated accessions from 50 countries representing the largest collection of pearl millet assembled and conserved any where in the world. A major portion of the germplasm came from the collection missions collaboratively launched by ICRISAT and other institutions in Africa and Asia. Redundancies were expected in such a huge collection. The present study was undertaken to define a pearl millet core collection for the world collection to reduce the number of accessions to a manageable extent for extensive evaluation and selection of well adapted accessions for further breeding programs and to fulfill the following objectives:

- To compare multivariate vs. hierarchical classification for determining core collection that maximizes diversity.
- To evaluate the adequacy of mean characterization data to determine core collection and do diversity assessment in a cross-pollinated crop.
- To assess the diversity within and between the selected core sample using morpho-agronomic traits and molecular markers.

We, began to review many issues related to outline the steps and questions involved in selecting core and minimum core, with a special focus on cross-pollinated species.

The pearl millet world collection was classified into four geographical regions. These geographical regions were compared with the entire collection for diversity that behaved similarly for most of the characters. The accessions were then classified into 25 different groups based on country of origin which were further subjected to hierarchical cluster analysis using Ward's minimum variance method of grouping accessions based on degree of similarity for characters. The similarity among the clusters was measured in terms of the Euclidean distance. The cluster analysis divided the entire collection into 25 distinct clusters with differences in the diversity and size among the clusters formed.

Using proportional strategy, a core sample of 1600 accessions was selected representing 10 percent sample of the entire collection.

The core sample thus formed was evaluated for representation of diversity by various statistical tests. Means and variances of quantitative characters of the core sample were compared with that of the entire collection. Further, a minimum core of 504 accessions was drawn from the core sample, representing 3.14 percent sampling of the entire collection. The minimum core was further evaluated for representation to the core sample and entire collection by several statistical tests. Levene's test for homogeneity of variances and Newman-Keuls test for post-hoc comparison of means indicated that none of the characters compared had significant different means and variances except for a few characters. The deviations in frequency distributions from the world collection frequency were also not significant for most of the quantitative and qualitative characters. Furthermore, the deviations in frequency on the basis of over all subclasses (69) of both qualitative and quantitative characters together, over subclasses of qualitative (40) and quantitative (29) characters separately was also not significant. The diversity indices were also almost of the same level as in the entire collection when compared by Shannon-Weaver diversity indices.

All the statistical tests used to evaluate the representation of the core sample and the minimum core proved that the proportional sampling method used in the selection of core and minimum core represented well the world collection. Thus, the core sample of 1600 accessions can be designated as the global core collection for entire pearl millet germplasm at ICRISAT and the minimum core can be quickly used for identifying some economic traits by looking into a manageable sample of accessions.

The credit for better performance of proportional strategy in identifying the core sample can be attributed to the factors, namely presence of redundancy in all groups, no variation for the presence of redundancy among the clusters, and occurrence of rare variants in larger groups.

The effect of environment on phenotypic diversity was also studied by growing the selected minimum core (developed by proportional strategy after stratifying through cluster analysis and geographical origin) in two different seasons during 1998-99 *rabi* and 1999-2000 *kharif*. Pearl millet being a thermo and photosensitive crop greatly influenced by these two seasons and can be attributed for G X E interaction. Most of the accessions from Africa were found both thermo and photosensitive. The morphological data collected over these two seasons were analysed by estimating the variation due to season, accession, method of classification and season x accession interaction in a linear model using Residual Maximum Likelihood (REML) analysis.

Variation due to accession and season was significant only for days to flowering and plant height and the interaction effect of season x accession was effective for days to flowering, plant height and spike length. The variation due to core, that is, selection method was significant for most of the characters studied. The results indicated that the different environmental conditions has not significantly affected the quantitative characters studied except days to flowering and plant height which may be attributed to the seasonal effects which greatly influence the number of days to flowering and growth of the crop. However, most of the characters remained unchanged during these two different seasons, which indicates a relatively higher stability of most of the characters over seasons in different years at ICRISAT, Patancheru location. Inter season Spearman rank order correlations were also highly significant for all the characters. The pattern of correlations among the characters within each season was also similar. The conclusions that can be drawn from the present study are,

(i) The sampling strategy (proportional) used in the present study for the development of a core sample resulted in the selection of a representative core.

(ii) Significant differences were not observed between the multivariate and hierarchical approach for stratification, however multivariate approach performed better for capturing many rare variants.

(iii) Though pearl millet is a cross-pollinated crop wherein the use of means for characterization could give skewed results but it proved well in the selection of a representative core.

(iv) The stability of accessions for most of the characters in different environments over years, except days to flowering and plant height was relatively high.

(v) Though wide within variations were observed for most of the accessions both at morphological and molecular level, but higher variation between accession proved that each of these accessions are distinct.

(vi) The accessions selected for diversity studies at molecular level revealed high within and between variations with the use of RFLP markers.

This global core established from the pearl millet world collection further could be used for evaluating accessions in breeding programs for many useful characters such as photoperiod sensitivity, resistance to various diseases (mostly for downy mildew), tolerance to drought, saline and alkaline conditions.

7. References

7. REFERENCES

Anand Kumar, K. and Andrews, D.J. 1993. Genetics of qualitative traits in pearl millet: A Review. **Crop Sci.**, 33: 1-20.

Anand Kumar, K. and Appa Rao, S. 1987. Diversity and utilization of pearl millet germplasm. Pp. 69-82. In: Proceedings of the International Pearl Millet Workshop, 7-11 April, 1986, ICRISAT Center, India (J.R. Witcombe and S.R. Beckerman, eds.). ICRISAT, Patancheru, India.

Anderson, W.F., Holbrook, C.C. and Culbreath, A.K. 1996. Screening the peanut core collection for resistance to tomato spotted wilt virus. **Peanut Sci.**, 23: 57-61.

Appa rao, S., 1999. Pearl Millet: Genetic resources. In: Pearl Millet Breeding. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 49-81.

Appa Rao, S. and de Wet, J.M.J., 1999. Pearl Millet: Taxonomy and Evolution. In: Pearl Millet Breeding. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 30-47.

*Balfourier, F and Charmet, G. 1994. Etude methodologique de la conservation de resources genetiques de ray-grass anglais (*Graminee fourragere*) par multiplication en pools de populations naturelles. **Genet. Sel. Evol.**, 26: 203-218.

Balfourier, F., Charmet, G., Prosperi, J.M., Goulard, M. and Monestiez, P. 1998. Comparison of different spatial strategies for sampling a core collection of natural populations of fodder crops. **Genet. Sel. Evol.**, 30: 215-235.

Balfourier, F., Prosperi, J.M., Charmet, G., Goulard, M and Monestiez, P. 1999. Using spatial patterns of diversity to develop core collections. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genet. Res. Instit., Rome, Italy. pp. 37-49.

Basigulp, D.H., Barnes, D.K. and Stuckev, R.E., 1995. Development of a core collection for perennial *Medicago* plant introductions. **Crop Sci.**, 35: 1163-1168.

Bataillon, T.M., David, J.L. and Schoen, D.J., 1996. Neutral Genetic Markers and Conservation Genetics: Simulated germplasm collections. **Genetics**, 144: 409-417.

Beyermann, B., Nurnberg, P., Weihe, A., Meixer, M., Epplen, J.T. and Borner, T. 1992. Fingerprinting plant genomes with oligonucleotide probes specific for single repetitive DNA sequences. **Theor. Appl. Genet.**, 83: 691-694.

Bisht, I.S., Mahajan, R.K., Loknathan, T.R. and Agrawal, R.C., 1998a. Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. **Gen. Res. and Crop Evol.**, 45 (4): 325-335.

Bisht, I.S. Mahajan, R.K. and Patel, D.P., 1998b. The use of characterization data to establish the Indian mungbean core collection and assessment of genetic diversity. **Gen. Res. and Crop Evol.**, 45 (2): 127-133.

Bisht, I.S., Mahajan, R.K., Lokanathan, T.R., Gautam, P.L., Mathur, P.N. and Hodgkin, T., 1999. Assessment of genetic diversity, stratification of germplasm accessions in diversity groups and sampling strategies for establishing a core collection of Indian sesame (*Sesamum indicum* L.). **Gen. Res. and Crop Evol.**, 119: 35-46.

Brown, A.H.D., 1989a. Core collections: a practical approach to genetic resources management. **Genome**, 31: 818-824.

Brown, A.H.D., 1989b. The case for core collections. In: The use of plant genetic resources. Eds. A.H.D. Brown, O.H. Frankel, D.R. Marshall and J.T. Williams. Cambridge University Press, Cambridge, UK. pp. 136-156.

Brown, A.H.D., 1990. The role of isozyme studies in molecular systematics. **Australian Systematic Bot.**, 3: 39-46.

Brown, A.H.D., 1995. The core collection at the crossroads. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK, pp. 3-20.

Brown, A.H.D, Grace, J.P. and Speer, S.S., 1987. Designation of a core collection of perennial *Glycine*. **Soybean Genet. Newsletter**, 14: 59-70.

Brown, A.H.D., Brubaker, C.L. and Grace, J.P., 1997. Regeneration of germplasm samples: wild versus cultivated plant species. **Crop Sci.**, 37: 7-13.

Brown, A.H.D. and Spillane, C. 1999. Implementing core collections- principles, procedures, progress, problems and promise. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genet. Res. Instit., Rome, Italy. pp. 1-9.

Brunken, J.N., de Wet, J.M.J. and Harlan, J.R., 1977. The morphology and domestication of pearl millet. **Econ. Bot.**, 31: 163-174.

Burton, G.W. 1985. Collection, evaluation and storage of pearl millet germplasm. **Field Crops Res.**, 11: 123-129.

Casas, A.M., Igartua, E., Valles, M.P. and Molina-Cano, J.L., 1998. Genetic diversity of barley cultivars grown in Spain, estimated by RFLP, similarity and coancestry coefficients. **Plant Breeding**, 117: 429-435.

Casler, M., 1995. Patterns of variation in a collection of perennial ryegrass accessions. **Crop Sci.**, 35: 1169-1177.

Charmet, G. Balfourier, F. and Monestiez, P., 1994. Hierarchical clustering of perennial ryegrass populations with geographic continuity constraint. **Theor. Appl. Genet.**, 88: 42-48.

Charmet, G., Balzouzzian, F., Ravel, C. and Denis, J.B. 1993. Genotype x environment interaction in a core collection of French perennial ryegrass populations. **Theor. Appl. Genet.**, 86: 731-736.

*Chavarriaga-Aguirre, P., Maya, M.M., Tohme, J., Duque,M.C., Carlos,I., Bonierbale, M.W., Kresovich,S. and Kochert,G., 1999. Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. **Mol. Breeding**, 5: 263-273.

Cordeiro, C.M.T, Morales, E., Ferreira, P., Rocha, D.M.S, Costa, I, Valois, A. and Silva, S. 1995. Towards a Brazilian core collections of cassava. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK, pp. 155-168.

Crossa, J., 1989. Methodologies for estimating the sample size required for genetic conservation of outbreeding crops. **Theor. Appl. Genet.**, 77: 153-161.

Crossa, J., Hernandez, C.M., Bretting, P., Eberhart, S.A. and Taba, S., 1993. Statistical genetic consideration for maintaining germplasm collections. **Theor. Appl. Genet.**, 86: 673-678.

Crossa, J., Taba, S., Eberhart, S.A. and Bretting, P., 1994. Practical considerations for maintaining germplasm in maize. **Theor. Appl. Genet.**, 89: 89-95.

Crossa, J., Delacy, I.H. and Taba, S., 1995. The use of multivariate methods in developing a core collection. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK, pp. 77-92.

Deu, M., Gonzales-de-Leon, D., Glaszman, J.C., Degremont, I., Chantereau, J., Lanaud, C. and Hamon, P., 1994. RFLP diversity in cultivated sorghum in relation to racial differentiation. **Theor. Appl. Genet.**, 88: 834-844.

Diwan, N., Bauchan, G.R. and McIntosh, M.S., 1994. A core collection for the United States annual *Medicago* germplasm collections. **Crop Sci.**, 34: 279-285.

Diwan, N., McIntosh, M.S. and Bauchan, G.R., 1995. Methods of developing a core collection of annual *Medicago* species. **Theor. Appl. Genet.**, 90: 755-761.

Dodds, J.H. and Watanbe, K., 1990. Biotechnical tools for plant genetic resources management. **Diversity**, 6: 26-28.

Dubreuil, P. and Charcosset, A., 1999. Relationships among maize inbred lines and populations from European and North-American origins as estimated using RFLP markers. **Theor. Appl. Genet.**, 99: 473-480.

Dussert, S., Chabrilange, N., Anthony, F., Engelmann, F., Recalt, C. and Hamon, S., 1997. Variability in storage response within a coffee (*Coffea* ssp.) core collection under slow growth conditions. **Plant Cell Reports**, 16: 344-348.

Eberhart, S.A., Bramel-Cox, P.J. and Prasada Rao, K.E., 1998. Preserving genetic resources. In: Proceedings of the International Conference on Genetic Improvement of Sorghum and Pearl Millet., Lobbock, Texas, 23-27 Sept. 1996, pp. 25-41.

Ellis, P.R., Pink, D.A.C., Phelps, K., Jukes, P.L., Breeds, S.E. and Pinnegar, A.E., 1998. Evaluation of a core collection of *Brassica oleracea* accessions for resistance to *Brevicoryne brassicae*, the cabbage aphid. **Euphytica**, 103: 149-160.

Erskine, W.A. and Muehlbauer, F.J., 1991. Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. **Theor. Appl. Genet.**, 83: 119-125.

FAO Quarterly Bulletin of statistics. 1999.

Federer, W.T., 1961. Augmented designs with one way elimination of heterogeneity. **Biometrics**, 17: 447-473.

Feinberg, A.P. and Vogelstein, B., 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. **Anal. Biochem.**, 132: 6-13.

Franco, J., Crossa, J., Villasenov, J., Taba, S. and Eberhart, S.A., 1997. Classifying Mexican Maize Accessions using Hierarchical and Density search methods. **Crop Sci.**, 37: 972-980.

Franke, M.D., Brenneman, T.B. and Holbrook, C.C., 1999. Use of a core collection to identify resistance to *Rhizoctonia* limb rot in peanut and the correlation of resistance with hypocotyle infections of seedlings. **Plant Dis.**, 83: (In Press).

Frankel, O.H., 1984. Genetic perspectives of germplasm conservation. In: Genetic Manipulation: Impact on Man & Society. Eds. W. Arber, K. Llimensee, W.J. Peacock and P. Starlinger, Cambridge University Press. Cambridge. pp. 161-170.

Frankel, O.H. and Brown, A.H.D., 1984. Plant genetic resources today: a critical appraisal. In: Crop Genetic Resources: Conservation and Evaluation. Eds. J.H.W.Holden and J.T. Williams, George Allen and Unwin, London, UK, pp. 249-257.

Galwey, N.W., 1995. Verifying and validating the representativeness of a core collection. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 187-198.

Genstat 5 Reference Manual, 1993. Oxford University Press Inc., New York.

Gepts, P., 1995. Genetic markers and core collection. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 127-146.

Grauke, L.J. and Thompson, T.E., 1995. Evaluation of Pecan [*Carya illinoinensis* (Wangenh.) K. Koch] germplasm collections and designation of a core subset. **HortScience**, 30 (5): 50-54.

Hamon, S. and van Sloten, D.H., 1989. Characterization and evaluation of okra. In: The use of plant genetic resources. Eds. A.H.D. Brown, O.H. Frankel, D.R. Marshall and J.T. Williams, Cambridge University Press, Cambridge, UK., pp. 173-196.

Hamon, S. and Noirot, M., 1990. Some proposed procedures for obtaining a core collection using quantitative plant characterization data. Paper presented at International workshop on okra held in NBPGR, New Delhi, India, October., pp. 8-12.

Hamon, S., Noirot, M. and Anthony, F., 1995. Developing a coffee core collection using the principal components score strategy with quantitative data. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 117-126.

Hanna, W.W. and Rai, K.N., 1999. Inbred Line Development. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 257-267.

Harinarayana, G., Ananad Kumar, K. and Andrews, D. J., 1999. Pearl Millet in Global Agriculture. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 479-506.

Harlan, J.R. and de Wet, J.M.J., 1971. Towards a relational classification of cultivated plants. **Taxon**, 20 (4): 509-517.

Hillis, D.N., 1987. Molecular versus morphological approaches to systematics. **Ann. Rev. Ecology and Systematics**, 18: 23-42.

Hintum, Th. J.L. van and Haalman, D., 1994. Pedigree analysis for composing a core collection of modern cultivars, with examples from barley (*Hordeum vulgare* s. lat.). **Theor. Appl. Genet.**, 88: 70-74.

Hintum, Th. J.L. van., 1995. Hierarchical approaches to the analysis of genetic diversity in crop plants. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayce publications. Chichester, UK., pp. 23-24.

Hintum, Th. J.L. van, Von Bothmer, R. and Visser, D.L., 1995. Sampling strategies for composing a core collection of cultivated barley (*Hordeum vulgare* s. lat.) collected in China. **Hereditas,** 122: 7-17.

Hodgkin, T., Quingyuan, G., Xiurong, Z., Yingzhong, Z., Xiangyun, F., Gautam, P.L., Mahajan, R.K., Bisht, I.S., Lokanathan, T.R., Mathur, P.N. and Mingde, Z., 1999. Developing sesame core collections in China and India. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genetic Resources Institute, Rome, Italy. pp. 74-81.

Holbrook, C.C., Anderson, W.F. and Pittman, R.N., 1993. Selection of a core collection from the U.S. germplasm collection of peanut. **Crop Sci.**, 33: 859-861.

Holbrook, C.C. and Anderson, W.F., 1995. Evaluation of a core collection to identify resistance to late leaf spot in peanut. **Crop Sci.**, 35: 1700-1702.

Holbrook, C.C., Stephenson, M.G. and Johnson, A.W., 1997. Level and geographical distribution of resistance to *Meloidogyne arenaria* in the U.S. germplasm collection of peanut. **Agron. Abstr.**, 157.

Holbrook, C.C., 1999. Testing and utilization of a core collection for the US germplasm collection of peanut. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genetic Resources Institute, Rome, Italy. pp. 68-73.

Huaman, Z., Aguilar, C. and Ortiz, R., 1999. Selecting a Peruvian sweetpotato core collection on the basis of morphological, eco-geographical, disease and pest reaction data. **Theor. Appl. Genet.**, 98: 840-844.

Huaman, Z., Ortiz, R., Zhang, D. and Rodriguez, F., 2000. Isozyme analysis of entire and core collections of *Solanum tuberosum* subsp. *andigena* potato cultivars. **Crop Sci.**, 40: 273-276.

Hutcheson, K., 1970. A test for comparing diversities based on the Shannon formula. **J. Theor. Biol.**, 29: 151-154.

IBPGR AND ICRISAT, 1993. Descriptors for pearl millet [*Pennisetum glaucum* (L.) R. Br.]. IBPGR, Rome, Italy; ICRISAT, Patancheru, India.

Igartua, E., Gracia, M.P., Lasa, J.M., Medina, B., Molina-Cano, J.L. and Romagosa, I., 1998. The Spanish barley core collection. **Gen. Res. and Crop Evol.**, 45: 475-481.

Islieb, T.G., Beute, M.K., Rice, P.W. and Hollowell, J.E., 1995. Screening of the peanut core collection for resistance to *Cylindrocladium* black rot and early leaf spot. **Proc. Am. Peanut Res. and Educ. Soc.**, 27: 25.

Jarret, R.L., Liu, Z.W and Webster, R.W., 1998. Genetic diversity among *Paspalum* spp. as determined by RFLPs. **Euphytica**, 104: 119-125.

Johnson,R.C., Johnston, W.J., Nelson, M.C., Simon, C.J. and Golob, C.T., 1999. Core utilization and development - an example with *Poa pratensis* L. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genetic Resources Institute, Rome, Italy. pp. 49-60.

Jung, H.G., Sheaffer, C.C, Barnes, D.K. and Halgeron, J.L., 1997. Forage quality variation in the U.S. alfalfa core collection. **Crop Sci.**, 37: 1361-1366.

Khairwal, I.S., Chandgi Ram. and Chhabra, A.K., 1990. Pearl Millet: Seed Production and Technology. Manohar Publications, New Delhi, India.

Khairwal, I.S. and Sigh, S., 1999. Quantitative and qualitative traits. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 119-155.

Kimura, M. and Crow, J.F., 1964. The number of alleles that can be maintained in a finite population. **Genetics,** 49: 725-738.

Knupffer, H. and Hintum Th. J.L. van, 1995. The barley core collection- an international effort. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 171-178.

Liu, C.J., Witcombe, J.R., Busso, C.S., Pittaway, T.S., Nash, M. and Gale, M.D., 1994. Construction and application of RFLP-based genetic maps in pearl millet. **Theor. Appl. Genet.**, 89: 481-487.

Loos, B.P. and van Duin, P.J.W., 1989. Establishing a core collection representing genetic variation in tulip. **Plant Gen. Res. Newsletter**, June-September. pp. 11-12.

Lux, H. and Hammer, K., 1994. Molecular markers and genetic diversity-some experience from the gene bank proceedings of the Eucarpia meeting: Evaluation and exploitation of genetic resources. Pre-breeding. 15-18 March, 1994, Clermont Ferrand, France. pp. 49-53.

Mackay, M.C., 1990. Strategic planning for effective evaluation of plant germplasm. In: Wheat genetic resoueces: meeting diverse needs. Eds. I.P. Srivastava and A.B. Damania. John Wiley & Sons publication, England., pp. 49-53.

Mackay, M.C., 1995. One or many core collections? In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 199-210.

Mahajan, R.K., Bisht, I.S., Agarwal, R.C. and Rana, R.S., 1996. Studies on South Asian okra collection: Methodology for establishing a representative core set using characterization data. **Genet. Res. and Crop Evol.**, 43: 249-255.

Mangat, B.K., Maiti, R.K and Khairwal, I.S., 1999. Pearl millet biology. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 1-27.

Maniatis, T., Fritsch, E.R. and Sambrook, J., 1982. Quantification of DNA and RNA. In: Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory Press. pp. 448-469.

Marshall, D.R., 1989. Limitations to the use of germplasm collections. In: The use of plant genetic resources. Eds. A.H.D. Brown, O.H. Frankel, D.R. Marshall and J.T. Williams, Cambridge University Press, Cambridge, UK., pp. 105-122.

Mathews, P. and Ambrose, M.J., 1994. Development and use of a core collection for the John Innes *Pisum* collection. Proceedings of the Eucarpia meeting: Evaluation and exploitation of genetic resources. Pre-breeding. 15-18 March, 1994, Clermont Ferrand, France. pp. 99-102.

Melchinger, A.E., Graner, A., Singh, M. and Messmer, M., 1994. Relationship among European barley germplasm: Genetic diversity among winter and spring cultivars revealed by RFLPs. **Crop Sci.**, 34: 1191-1199.

Messmer, M.M., Melchinger, A.E., Boppenmaier, J., Brunklaus-Jung, E. and Herrmann,R.G., 1992. Relationships among early European maize inbreds: I. Genetic diversity among flint and dent lines revealed by RFLPs. **Crop Sci.**, 32: 1301-1309.

Miklas, P.N., Delorme, R., Hannan, R. and Dickson, M.H., 1999. Using a subsample of the core collection to identify new sources of resistance to white mold in common bean. **Crop Sci.**, 39: 569-573.

Moody, D.B., Rathjen, A.J., Cartwright, B., Paull, J.F. and Lewis, J., 1988. Genetic diversity and geographical distribution of tolerance to high levels of soil boron. In: Proceedings of the 7th International Wheat genetic symposium. Eds. T.E. Miller and R.M.D. Koebner. Cambridge University Press, Cambridge pp. 859-865.

Morales, E.A.V., Valois, A.C.C. and Costa, I.R.S., 1995. Core collections for gene banks with limited resources. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 241-249.

Moser, H. and Lee, M., 1994. RFLP variation and genealogical distance, multivariate distance, heterosis, and genetic variance in oats. **Theor. Appl. Genet.**, 87: 947-956.

Mumm, R.H. and Dudley, J.W., 1994. A classification of 148 U.S. maize inbreds: I. Cluster analysis based on RFLPs. **Crop Sci.**, 34: 842-851.

Mumm, R.H., Hubert, L.J. and Dudley, J.W., 1994. A classification of 148 U.S. maize inbreds: II. Validation of cluster analysis based on RFLPs. **Crop Sci.**, 34: 852-865.

Murthy, B.R., Upadhyay, M.K. and Manchanda, P.L., 1967. Classification and cataloguing of a world collection of genetic stocks of *Pennisetum*. **Indian J. Genet. and Pl. Breeding**, 27: 313-394.

Nodari, R.O., Koinange, E.M.K., Kelly, J.D. and Gepts, P., 1992. Towards an integrated linkage map of common bean. I. Development of genomic DNA probes and levels of restriction fragment length polymorphism. **Theor. Appl. Genet.**, 84: 186-192.

Noirot, M., Hamon, S. and Anthony, F. 1995. The principal component scoring: a new method of constituting a core collection using quantitative data. **Genet. Res. and Crop Evol.**, 43: 1-6.

Ortiz, R., Ruia-Tapia, E.N. and Mijica-Sanchez., 1998. Sampling strategy for a core collection of Peruvian Quinoa germplasm. **Theor. Appl. Genet.**, 96: 475-483.

Ortiz, R., Madsen, S., Ruiz-Tapia, E.N., Jacobsen, S.E., Mujica-Sanchez, A., Christiansen, J.L. and Stolen, O., 1999. Validating a core collection of Peruvian Quinoa germplasm. **Genet. Res. and Crop Evol.**, 46: 285-290.

Ouendeba, B., Ejeta, G., Hanna, W.W. and Kumar, A.K.,1995. Diversity among African pearl millet landrace populations. **Crop Sci.**, 35: 919-924.

Pederson, G.A., Fairbrother, T.E. and Greene, S.L., 1996. Cyanogenesis and climatic relationships in U.S. white clover germplasm collection and core subset. **Crop Sci.**, 36: 427-433.

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

Peeters, J.P., Wilkes, H.G. and Galwey, N.W., 1993. The use of eco-geographical data in the exploitation of variation from gene bank. **Theor. Appl. Genet.**, 80: 110-112.

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

Prasada Rao, K.E. and Ramanath Rao, V., 1995. The use of characterization data in developing a core collection of sorghum. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 109-116.

Radovic, G. and Jelovac, D., 1994. The possible approach in maize core collection development. Proceedings of the Eucarpia meeting: Evaluation and exploitation of genetic resources. Pre-breeding. 15-18 March, 1994, Clermont Ferrand, France. pp. 109-115.

Rai, K.N., Appa Rao, S. and Reddy, K.N., 1995. Pearl Millet. In: Biodiversity in Trust: Conservation and use of Plant Genetic Resources in CGIAR Centers. Eds. D.Fuccillo, L. Sears and P. Stapleton. Cambridge University Press, Cambridge. pp. 243-258.

Rai, K.N. and Virk, D.S., 1999. Breeding Method. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 185-211.

Reddy, K.N., Appa Rao, S. and Mangesha, M.H., 1996. Diversity in pearl millet germplasm from Sudan. **Indian J. Pl. Genet. Resources**, 9 (1): 43-51.

Reddy, K.N., Appa Rao, S. and Mangesha, M.H., 1996. Diversity in pearl millet germplasm from Central African Republic. **Genet. Resources and Crop Evol.**, 43: 303-308.

Ribu, H.K.M and Hilu, K.W., 1994. Detection of interspecific and intraspecific variation in *Panicum* millets through random amplified polymorphic DNA. **Theor. Appl. Genet.**, 88: 412-416.

Robinson, D.L., 1987. Estimation and use of variance components. **The Statistician,** 36: 3-14.

Romesburg, H.C. 1984. Cluster analysis for researchers. Lifetime Learning Publ., Belmont, CA.

Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A. and Allard, R.W., 1984. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. **Proc. Natl. Acad. Sci., USA,** 81: 8014-8018.

SAS Institute. 1989a. SAS User's Guide: Statistics. State University Press, Cary, NC, USA.

Schoen, D.J. and Brown, A.H.D., 1993. Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. **Proc. Natl. Acad. Sci., USA,** 90: 10623-10627.

Schoen, D.J. and Brown, A.H.D., 1995. Maximizing allelic diversity in core collections of wild crop relatives: the role of genetic markers. In: Spagnoletti-Zeuli, P.L. and Qualset, C.O., pp. 55-76.

Sharp, P.J., Kries, M., Shewry, P.R. and Gale, M.D., 1988. Location of -amylase sequences in wheat and its relatives. **Theor. Appl. Genet.**, 75: 286-290.

Skinner, D.Z., Bauchan, G.R., Auricht, G. and Hughes, S., 1999. A method for the efficient management and utilization of large germplasm collections. **Crop Sci.**, 39: 1237-1242.

Skinner, D.Z., Bauchan, G.R., Auricht, G. and Hughes, S., 1999. Developing core collection from a large annual *Medicago* germplasm collection. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genetic Resources Insitute, Rome, Italy. pp. 61-67.

Skroch, P.W., Nienhuis, J., Beebe, S., Tohme, J. and Pedraza, F., 1998. Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. **Crop Sci.**, 38: 488-496.

Spagnoletti-Zeuli, P.L. and Qualset, C.O., 1987. Geographical diversity for quantitative spike characters in a world collection of *durum* wheat. **Crop Sci.**, 27: 235-241.

Spagnoletti-Zeuli, P.L. and Qualset, C.O., 1993. Evaluation of five strategies for obtaining a core subset from a large genetic resource collection of *durum* wheat. **Theor. Appl. Genet.**, 87: 295-304.

Spagnoletti-Zeuli, P.L. and Qualset, C.O., 1995. The durum wheat core collection and the plant breeder. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 213-227.

Sneath, P.H.A. and Sokal, R.R., 1973. Numerical Taxonomy. W.H. Freeman and Company, San Francisco.

Statsoft, 1997. Statistica for windows. Statsoft, USA.

Steel, R.G.D. and Torrie, J.H., 1980. Principles and procedures of statistics, A biometrical approach. Second edition. McGraw-Hill Book Co , New York.

Taba, S., Diaz, J., Franco, J. and Crossa, J., 1998. Evaluation of Caribbean maize accessions to develop a core subset. **Crop Sci.**, 38: 1378-1386.

Talukdar, B.S., Khairwal, I.S. and Singh, R., 1999. Hybrid Breeding. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 269-301.

Tohme, J., Jones, P., Beebe, S. and Iwanga, M., 1995. The combined use of agro-ecological and characterization data to establish the CIAT *Phaseolus vulgaris* core collection. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 95-107.

Tohme,J., Gonzalez, D.O., Beebe, S and Duque, M.C., 1996. AFLP analysis of gene pools of a wild bean core collection. **Crop Sci.**, 36: 1375-1384.

Tohme, J., Beebe, S. and Iglesias, C., 1999. Molecular characterization of the CIAT bean and cassava core collections. In: Core collections for today and tomorrow. Eds. R.C. Johnson and T. Hodgkin. International Plant Genetic Resources Insitute, Rome, Italy. pp. 28-36.

Tolbert, D.M., Qualset, C.O., Jain, S.K., Craddock, J.C., 1978. A diversity analysis of a world collection of barley. **Crop Sci.**, 19: 789-794.

Tostain, S., Riandey, M.F. and Marchais, L., 1987. Enzyme diversity in pearl millet (*Pennisetum glaucum*). 1. West Africa. **Theor. Appl. Genet.**, 77: 634-640.

Vaughan, D.A., 1991. Choosing rice germplasm for evaluation. **Euphytica,** 54: 147-154.

Vaughan, D.A. and Jackson, M.T., 1995. The core as a guide to the whole collection. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 229-239.

Vierling, R.A., Xiang, Z., Joshi,C.P., Gilbert,M.L. and Nguyen, H.T., 1994. Genetic diversity among elite sorghum lines revealed by restriction fragment length polymorphism and random amplified polymorphic DNAs. **Theor. Appl. Genet.**, 87: 816-820.

von Bothmer, R., Fishbeek, G., van Hintum, T., Hodgkin, T. and Knupffer, H. 1990. The barley core collection. Final report of the BCC working group. Weihenstaphan, pp. 14.

Ward., J.H., 1963. Hierarchical grouping to optimize an objective function. J. American **Statistical Association,** 58: 236.

Weltzien, E., 1989. Differentiation among barley landrace populations from the Near East. **Euphytica**, 43: 29-39.

Wilson, J.P., Burton, G.W., Zongo, J.D. and Dicko, I.O., 1990. Diversity among pearl millet landraces collected in Central Burkina Faso. **Crop Sci.**, 30: 40-43.

Witcombe, J.R. 1999. Population Improvement. Eds. I.S. Khairwal, K.N. Rai, D.J. Andrews and G. Harinarayana. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, India. pp. 213-256.

Xiurong, Z., Yingzhong, Z., Xiangyun, F., Yong, C., Qingyuan, G., Yurong,L. and Yongning, W., 1999. Establishment and development of sesame germplasm core collection in China. **Pl. Genet. Resources Newsletter**, 119: 47-50.

Yang, X. and Quiros, C., 1993. Identification and classification of celery cultivars with RAPD markers. **Theor. Appl. Genet.**, 86: 205-212.

Yang, W., Antonio, C de Oliveira, Ian Godwin, Keith Schertz and Jeffrey L Bennetzen, 1996. Comparison of DNA marker technologies in characterizing plant genome diversity: variability in Chinese Sorghums. **Crop Sci.**, 36: 1669-1676.

Yonezawa, K., Nimura, T. and Moroshima, I., 1995. Sampling strategies for use in stratified germplasm collections. In: Core collections of plant genetic resources. Eds. T. Hodgkin, A.H.D. Brown, M.J.L van Hintum and E.A.V. Morales. A Wiley-Sayee publications. Chichester, UK., pp. 35-53.

Zeven, A.C., Waninge, J., van Hintum, T. and Singh, S.P., 1999. Phenotypic variation in a core collection of common bean (*Phaseolus vulagaris* L.) in the Netherlands. **Euphytica**, 109: 93-106.

Zhang, Q., Saghai-Maroof, M.A., Lu, T.Y. and Shen, B.Z. 1992. Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. **Theor. Appl. Genet.**, 83: 495-499.

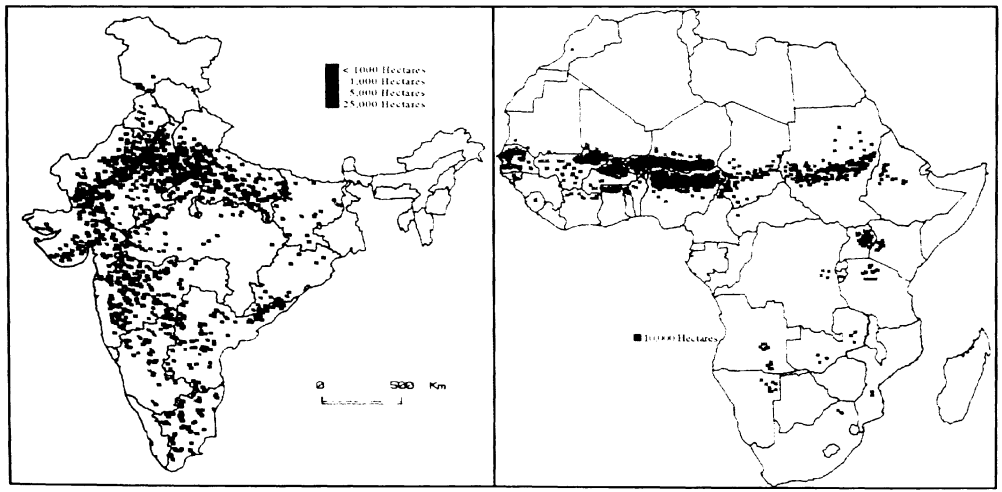
Zhang, Q., Saghai-Maroof, M.A. and Kleinhofs, A., 1993. Comparative diversity analysis of RFLPs and Isozymes within and among populations of *Hordeum vulgare* spp *spontaneum*. **Genetics**, 134: 909-916.

---

* Originals not seen

8. Appendix

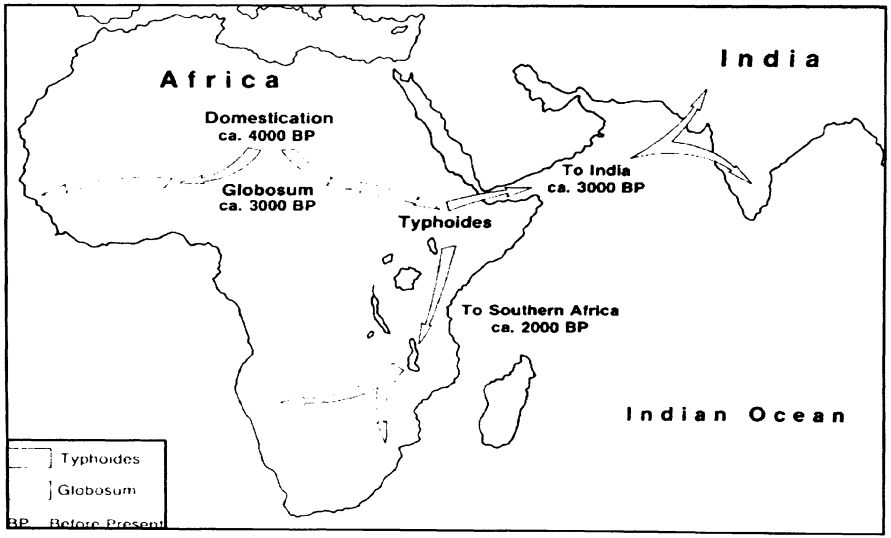
Fig 1. Distribution of pearl millet in India and Africa.



Source: Agricultural Situation in India, 1989

Source : FAO Bulletin of Statistics, 1990

Fig 2. Domestication and introduction of pearl millet into Africa and Indian subcontinent



Preparation of buffers and other chemicals for RFLP procedure

1M Tris.Cl. pH 8.5 (1000ml)

121.1g of Tris + 800ml of sterile dH_2O in 2000ml beaker + little Hydrogen chloride ⇒ stir – adjust the pH TO 8.5 and make the volume upto1000ml.

5M NaCl (1000ml)

292.2g of NaCl + 750ml of distilled water and autoclave

0.5M EDTA pH 8.0 (500ml)

93.05g EDTA + 300ml sterile dH2O in 1000ml beaker + little sodium hydroxide ⇒ Stir while heating ⇒ add little amount of sodium hydroxide in between – it takes time to dissolve ⇒ continue till it gets dissolved and turns transparent ⇒ adjust pH to 8.0 with sodium hydroxide ⇒ make the volume upto 500ml with SDW.

Proteinase K (2mg/ml)
20mg proteinase K + 10ml SDW ⇒ shake gently and divide into 10 aliquots of 1ml each to get conc. ⇒ store at -20°C.

2x Extraction buffer pH 8.0

Constituent	g/500ml	Final conc.
0.5M Tris-Cl pH 8.0	150ml	0.15M
0.5M EDTA pH 8.0	8ml	0.008M
NaCl	2.92g	0.1M
SDS (kept at 65°C)	7.5g	1.5%

Store at room temperature

70% Ethanol

Constituent	ml/litre	Final conc.
Absolute Ethanol	700ml	70%
SDW	300ml	-

$T_{50}E_{10}$ Buffer

Constituent	100 ml^{-1}
0.5M Tris-Cl pH 8.0	10ml
0.5M EDTA pH 8.0	2.0ml

Denaturing solution

Constituent	litre^{-1}
1.5M NaCl	87.66g
0.5M NaOH	20.00g

Neutralizing Buffer

Constituent	litre^{-1}
1.5M NaCl	87.66g
1M Tris	121.1g
Adjust pH TO 8.0	

Chloroform (24 : 1)

Constituent	100 ml
Isoamyl alcohol	4.0ml
Chloroform	96.0ml

0.25N HCl

Constituent	litre^{-1}
Conc. HCl (11.6N)	21.6ml
SDW	78.4ml

Ammonium acetate
57.75g ammonium acetate in 100ml SDW

50x TAE Buffer

Constituent	litre^{-1}
Trizma base	242.0g
Glacial acetic acid	57.1ml
0.5M EDTA pH 8.0	100ml

20x SSC

Constituent	litre^{-1}
3M NaCl	175.32g
0.3M Sodium citrate	88.23g

S Buffer

Constituent	litre^{-1}
100mM Tris. Cl.	100ml
100mM NaCl	20ml
50mM EDTA pH 8.0	100ml
20% SDS	100ml
Store warm	

20% SDS

200g SDS in 1litre of warm SDW ⇒ stir until dissolved

RNase (10mg/ml)

100mg RNase + 10ml SDW ⇒ Place in boiling water for 20min. ⇒ Cool slowly ⇒ Dispense into 10 aliquots of 1ml each ⇒ Store at -20°C.

Kesara's Loading Buffer

Constituent	$50ml^{-1}$
Bromophenol blue	0.25g
Xylene cyanol	0.25g
Glycerol	25ml
$Na_2EDTA.2H_20$	0.93g

Stir until dissolves ⇒ Dispense into 2ml screw-cap tubes ⇒ Store at 4°C.

5x HSB

Constituent	$litre^{-1}$
NaCl	175.3g
PIPES	30.3g
$Na_2EDTA.2H_20$	7.45g

Adjust pH 6.8 with 4M NaOH and autoclave.

Denhardt's III

Constituent	$100\ ml^{-1}$
Gelatin	2g
Ficoll-400	2g
PVP-360	2g
SDS	10g
Sodium pyrophosphate	5g

Store at 65°C.

Carrier DNA
5g Salmon testes DNA + 1litre SDW ⇒ Autoclave ⇒ Dispense in 50ml aliquots ⇒ Store at -20°C.

^{32}P Blots Wash Solution I

Constituent	$2litre^{-1}$
20x SSC	200ml
20% SDS	100ml
SDW	1700ml

^{32}P Blots Wash Solution II

Constituent	$2litre^{-1}$
20X SSC	20ml
20%SDS	100ml
SDW	1880ml

Probe Stripping Solution

Constituent	2litre^{-1}
20% SDS	50ml
10x SSC	10ml
SDW	1940ml

Developer
1000ml SDW + 157 g D-19 ⇒ Add slowly

Stop bath (3% HAC)
30ml HAC (Acetic acid)

Rapid Fixer

Constituent	litre^{-1}
Solution A	250ml
SDW	750ml
Solution B	28ml

Southern Transfer based on Reed and Mann, 1985 (NAR 13 [2] 7207-7221)

1. Nylon membranes are cut according to the size of the gel and pre-washed in sterile distilled water.
2. Take a large square petri-dish and pour 500ml of 0.4M NaOH.
3. Place a piece of glass on top, soak three sheets of Whatman 3mm paper wicks in 0.4M NaOH and place on the glass.
4. Starting with one of the gel edges, gradually slide the gel from the gel tank onto the petry-dish. Air-bubbles trapped in between the gel and Whatman sheets are removed.
5. Place the nylon membrane (Amersham Hybond-N^+) on top of the gel. Remove the trapped air-bubbles between the gel and the membrane.
6. Wet a piece of Whatman 3mm paper cut to the size of the gel and place on top of the nylon membrane. Remove the trapped air-bubbles.
7. Place two dry Whatman paper sheets and 500g weight on top.
8. Leave overnight.

Purification of DNA inserts from plasmid DNA (Maniatis *et al.*, 1982)

According to this procedure, the gene inserts of the clones were cleaved from their vectors using the appropriate restriction endonuclease (s) and fractionated by electrophoresis on a minigel of 0.8% agarose in TAE buffer containing ethidium bromide (0.5μg/ml). The electrophoresis was carried out with TAE buffer for 3 h at 20v/cm. The gels were observed on a UV-transilluminator and the desired fragment was transferred on to NA 45 membrane (Schleicher and Schull, Inc., Keene, NH) by placing the membrane in a slit just behind the band o interest and allowing the electrophoresis to resume for further 30 min. The DNA was eluted from the membrane by addition of sufficient high salt buffer (1M NaCl, 0.1mM EDTA, 20mM Tris-HCl, pH 8.0) to cover the membrane followed by incubation at 65°C for 45 min. Ethidium bromide was removed by extraction with TE saturated n-butanol and DNA was precipitated with 0.5 vol of isopropanol at -80°C for 30 min and pelleted in a Sorvall microfuge at 10,000 rpm for 10 min. The pellet was washed in 70% ethanol, dried under vacuum and dissolved in $T_{10}E_1$.

Studies on the establishment of a core collection of pearl millet (*Pennisetum glaucum*) germplasm

By
Ranjana Bhattacharjee

Major Advisor: Dr. I. S. Khairwal
Professor, Department of Plant Breeding
CCS Haryana Agricultural University, Hisar 125004, Haryana, India

ABSTRACT

The pearl millet germplasm collection has grown fairly large in size impeding proper evaluation and utilization of accessions. Development of procedures for reducing the collection size to a manageable and accessible level is an important issue in management of germplasm collections. The objective of this study was to develop a core collection of pearl millet (Pennisetum glaucum), representing a sample of the whole collection with minimum repetitiveness and maximum diversity. The pearl millet collection at ICRISAT consisted of 20,642 cultivated accessions from 50 different countries. The uncharacterized accessions were identified initially and eliminated retaining 16,063 accessions for further study. The accessions were stratified into four geographical regions and twenty-five countries of origin. Cluster analysis by Ward'' method further stratified the accessions into twenty-five groups based on eleven quantitative characters. Proportional sampling strategy was used to select ten percent of the whole collection to form the core sample. Further clustering of the core sample resulted in a working collection, representing 3.14 percent of the whole collection and termed as minimum core. Different statistical tests were used to evaluate the proportional method of selecting the core samples. The strategy was useful in selecting the core samples that retained more than 70 percent of the total variation in the whole collection. No significant difference was observed in hierarchical and multivariate approach in developing the core, however multivariate approach was useful in capturing more rare variants. The selected minimum core sample was evaluated in two different seasons. The data collected was analyzed for the effect of environment on characterization of accessions in different seasons. No significant effect of season was observed for variation among and within accessions. Correlation studies revealed considerable stability of characters over environments. Assessment of diversity at molecular level using RFLP markers revealed polymorphism within accessions of a subsample of core collection. The dendrogram grouped these 10 different accessions into 10 different clusters revealing the between accession variation. The minimum core, as a representative sample of the whole pearl millet collection has many potential uses and would increase the utilization of pearl millet genetic resources in future.

5/11/2001
Major Advisor

Head of the Department