CLONING AND CHARACTERIZATION OF PEARL MILLET REPETITIVE DNA SEQUENCES AND THEIR DISTRIBUTION AMONG RELATED SPECIES AND CEREALS

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

BY

GEORGE THOMAS

SCHOOL OF LIFE SCIENCES

UNIVERSITY OF HYDERABAD

HYDERABAD - 500 046

INDIA

APRIL, 1997

ENROLMENT NO. PL 5937

CONTENTS

STATEMENT

I hereby state that the work presented in this thesis entitled **"Cloning and characterization of pearl millet repetitive DNA sequences and their distribution among related species and cereals"** has been carried out by me under the supervision of Prof. N. C. Subrahmanyam, School of Life Sciences, University of Hyderabad, Hyderabad - 500 046, and that this work has not been submitted for any degree or diploma of any other University earlier.

GEORGE THOMAS

Prof. N. C. Subrahmanyam, FNA, FAAS, FASc.
School of Life Sciences
University of Hyderabad.

CERTIFICATE

This is to certify that Mr. George Thomas has carried out the research work embodied in the present thesis under my supervision and guidance for the full period prescribed under the Ph.D. ordinance of this University. I recommend his thesis entitled **"Cloning and characterization of pearl millet repetitive DNA sequences and their distribution among related species and cereals",** for submission for the degree of Doctor of Philosophy of this University.

Prof. N. C. Subrahmanyam FNA, FAAS, FASc.
Supervisor

Dean
School of Life Sciences
University of Hyderabad

Head
Dept. of Plant Sciences
School of Life Sciences
University of Hyderabad

ACKNOWLEDGEMENT

I wish to express my deep sense of gratitude to my supervisor Prof. N. C. Subrahmanyam for his encouragement, advice and for providing the laboratory facilities during the course of this investigation. I am indebted to him for the time and efforts spent on my behalf and for critically examining the manuscript.

I thank the Dean, School of Life Sciences and the Head, Department of Plant Sciences for providing the general facilities.

The help rendered by my colleagues in the laboratory and my friends is gratefully acknowledged.

I express my sincere appreciation to the field attendants Mr. Laxman and Mr. Anthaiah and Laboratory attendants Mr. Gopal and Mr. Sankar for their assistance, help and co-operation.

It is with much appreciation that I thank, Dr. Yashitola, CCMB, Hyderabad and Mr. Tony Jacob, CMC, Hyderabad for their help in computer analysis of the data. Support of the Technical staff of School of Life Sciences, University of Hyderabad is greatefully acknowledged. Special thanks are accorded to Mr. Anil K. Mishra for his help and co-operation.

I thank Mr. R. David Raju, for the timely and expert typing of this manuscript.
I am highly thankful to Mr. S.R. Gawali for the excellent photography.

The financial support of U.G.C in the form of Research Fellowship (JRF and SRF) is gratefully acknowledged.

Above all, I am deeply grateful to all my family members for their encouragement and support. Last but not the least I thank all my friends for their consistent moral support and encouragement to make this endeavour a success.

GEORGE THOMAS.

LIST OF TABLES

LIST FIGURES

Figure No. **Legend** **(page)**

Fig 4. Southern blot of restricted high copy DNA sequence clones hybridized to: **c** - labelled vector DNA and sheared ^{32}P labelled genomic DNA(s) from ***Pennisetum*** species: **g** - *P.* ***glaucum;*** **r** - *P.* ***ramosum;*** **m** - *P.* ***mezianum;*** **o** - *P.* ***orientale;*** **p** - *P.* ***purpureum.*** The clone numbers given at the top of each lane with a prefix **P**, **E** or **B** represent Pst I-, Eco RI- or Bam HI-based pearl millet genomic libraries respectively and the corresponding enzyme used for the digestion of the recombinant clones **(38)**

Fig 5. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; **o** - *P.* ***orientale***; **p** - *P.* ***purpureum***; **g** - *P.* ***glaucum*** hybridized to pearl millet high copy DNA sequence probes - PGP 005 (**a - e**); PGE 015 (**f - i**); PGE 123 (**j - l**) and the enzymes used are - Bam HI (●), Bgl I (○), Eco RI (✡), Eco RV (✶), Msp I (✸), Pst I (✪) or Sac I (✡) **(41)**

Fig 6. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; **o** - *P.* ***orientale***; **p** - *P.* ***purpureum***; **g** - *P.* ***glaucum*** hybridized to pearl millet high copy DNA sequence probes - PGE 123 (**a - c**); PGB 058 (**d**); PGB 074 (**e & f**); PGB 103 (**g & h**); PGB107 (**I & j**) PGB 582 (**k & l**) and the enzymes used are : Bam HI (●), Bgl I (○), Eco RI (✡), Hind III (✡), Msp I (✸) or Sac I (✡) **(44)**

Fig 7. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; o - *P.* ***orientale***; **p** - *P* ***purpureum***; **g** - *P.* ***glaucum*** hybridized to pearl millet high copy DNA sequence probes - PGB 582 (**a - d**); PGB 727 (**e - h**); PGB 788 (**i - l**) the enzymes used are - Bam HI (●), Bgl I(○), Eco RI (✡), Eco RV (✶), Hind III (✡), Msp I(✸) or Sac I (✡) **(48)**

Fig 8. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; o - *P* ***orientale***; **p** - *P.* ***purpureum***; **g** - *P.* ***glaucum*** hybridized to pearl millet high copy DNA sequence probes - PGB 625 (**a**); PGB 662 (**b & c**) and the enzymes used are : Bam HI (●) or Bgl I (○).

Bam HI-based pearl millet genomic libraries respectively, and the corresponding enzyme used for the digestion of the recombinant clones **(65)**

Fig 14. Southern blots of restricted genomic DNA from different cereals: **P - pearl millet (***P. glaucum)*; **S - sorghum** (*S. bicolor* var CSH - 11); **M - maize** (*Z. mays* var cm 119 x cm 120); **R - rice** (*O . sativa* var basmati); **W - wheat** (*T . aestivum* var chinese spring); **B - barley** (*H. vulgare* var betzes); **O - oats** (*A. sativa*); **RY - rye** (*S. cereale*) hybridized to pearl millet high copy DNA sequence probes: PGP 005 (**a** & **b**); PGE 015 (**c - f**); and the enzymes used are Bam HI (●), Eco RI (✡), Eco RV (✦), or Pst I (✪) **(68)**

Fig 15. Southern blots of restricted genomic DNA from different cereals: **P - pearl millet (***P. glaucum)*; **S - sorghum** (*S. bicolor* var CSH - 11); **M - maize** (*Z. mays* var cm 119 x cm 120); **R - rice** (*O . sativa* var basmati); **W - wheat** (*T . aestivum* var chinese spring); **B - barley** (*H. vulgare* var betzes); **O - oats** (*A. sativa*); **RY - rye** (*S. cereale*) hybridized to pearl millet high copy DNA sequence probes: PGE 123 (**a -d**); PGB 058 (**e**) ; PGB 074 (**f**) and the enzymes used are Bam HI (●), Eco RV (✦), Pst I (✪) or Xba I (□) **(70)**

Fig 16. Southern blots of restricted genomic DNA from different cereals: **P - pearl millet (***P. glaucum)*; **S - sorghum** (*S. bicolor* var CSH - 11); **M - maize** (*Z. mays* var cm 119 x cm 120); **R - rice** (*O . sativa* var basmati); **W - wheat** (*T . aestivum* var chinese spring); **B - barley** (*H. vulgare* var betzes); **O - oats** (*A. sativa*); **RY - rye** (*S. cereale*) hybridized to pearl millet high copy DNA sequence probes: PGB 103 (**a** & **b**) ; PGB 582 (**c - f**) and the enzymes used are Bam HI (●), Eco RI (✡), Eco RV (✦) or Xba I (□) **(72)**

Fig 17. Southern blots of restricted genomic DNA from different cereals: **P - pearl millet (***P. glaucum)*; **S - sorghum** (*S. bicolor* var CSH - 11); **M - maize** (*Z. mays* var cm 119 x cm 120); **R - rice** (*O . sativa* var basmati); **W - wheat** (*T . aestivum* var chinese spring); **B - barley** (*H. vulgare* var betzes); **O - oats** (*A. sativa*); **RY - rye** (*S. cereale*) hybridized to pearl millet high copy DNA sequence probes: PGB 727 (**a & b**); PGB 788 (**c & d**); PGB 625 (**e**); PGB 662 (**f**) and the enzymes used are Bam HI (●), Eco RI (✡), or Pst I (✪) **(74)**

Fig 18. Cereal-wise panels of DNA fragments generated by each enzyme in hybridizing to the set of probes were merged using a C-based

ABSTRACT

Members belonging to several families of repetitive sequences have been cloned and characterized from many crops over the past one and a half decades since the report of cloning the rye heterochromatic sequences in 1980. Value of such sequences as molecular markers in phylogenetic analysis, in identifying putative parents of naturally occurring polyploids, species/cultivar identification and in tracking chromatin in interspecific hybrids and chromosomes in addition lines have been well documented. In pearl millet monomers of one family of tandem repeat have been cloned and characterized so far. In spite of its considerable value in food, feed and forage production, pearl millet has not received the attention of researchers on par with other cereals. No attempt has been made so far to study the nature and extent of genomic homoeology/synteny between pearl millet and other cereals. Genome characterization and the development of molecular markers are vital for the effective design of breeding strategies of any species. The objectives of the present study were (i) to clone and characterize major families of pearl millet repetitive **DNA** sequences, (ii) to explore the utility of pearl millet repetitive **DNA** sequences as molecular markers, (iii) to examine the nature, conservation and modulation of repetitive sequences among related species and cereals and (iv) to explore the utility of pearl millet molecular markers in phylogenetic analysis of cereal genomes.

Bam HI- and Eco RI- based partial genomic libraries of pearl millet were constructed in pUC 18 vector using JM 109 strain of E.coli as host. Following colony hybridization with ^{32}P labelled total genomic **DNA** of pearl millet, colonies with high signal intensities were selected as putative high copy DNA sequence clones for further analysis (Clone designations comprises of three letters, P, G, B/E. The letters P.G. stands for pearl millet genomic while the third letter 'B' or 'E' stands for the enzyme Bam HI or Eco RI used for cloning and the number corresponds to the colony number in their respective genomic library. PGP 005, isolated from Pst I based genomic library of pearl millet, also was used). Cells from such colonies were amplified and plasmids

were extracted. Genomic DNA inserts were released by digesting the recombinant plasmids with corresponding enzyme and size fractionated on 0.8% agarose gels. Southern blots were hybridized with sheared radio-labelled total genomic DNA of pearl millet to confirm the relative abundance of the high copy DNA sequence inserts. Three of the Eco RI based and nine Bam HI based clones were further characterized and used as tools to study their distribution and extent of conservation among *Pennisetum* species differing in their basic chromosome number and major cereals

Copy number was estimated by quantitative dot blot analysis. Copy number ranged from 0.6 x 10^3 for PGB 727 to over 1 x 10^4 for PGB 625, confirming repetitive nature of these sequences in pearl millet genome. Pearl millet genomic DNA subjected to time course digestion with enzyme used for cloning was probed with the inserts from the corresponding enzyme based library to examine genomic organization of repetitive sequences. Time course digests of genomic DNA on probing with PGB 625 produced a smear in 5 min. The smear progressively turned into a ladder of fragments as digestion continued leading to a typical ladder of fragments, with an internal progression of 127bp sequence on complete digestion, indicating the tandem organization of PGB 625. Time course digested genomic DNA on probing with PGB 107 and PGB 662 produced a smear and discrete bands and the intensity of smear and fragments increased with prolonged digestion. The heterogeneity of the fragments homologous to PGB 107 and PGB 662 is indicative of a complex distribution at diverse sequence environment in the genome. Other than a thick smear from top to bottom in the lane, no distinct band was detected when Eco RI digested DNA was probed with PGE 096.

Restriction profiles of partially digested genomic DNA homologous to the probes PGE 015, PGB 074 and PGB 582 consisted of many bands and one of them had a size equal to the insert size of the probe used. The signal intensity of the fragment (equal to the insert size) increased with a concomitant decrease and final disappearance of large fragments on complete digestion implying that during the course of digestion, basic repeats belonging to each repetitive sequence family were being successively excised

from large stretches thereby increasing the intensity of the band corresponding to the size of cloned fragment. Results suggest that genomic sequences homologous to the inserts of PGE 015, PGB 074 and PGB 582 are mostly arranged in clusters of 2-10mer and such clusters are interspersed at different genomic locations. Different sized fragments in addition to the insert size found in the restriction profile hybridizing to PGB 727 is suggestive of its distribution in different configurations in the genome.

Thirteen repetitive sequences thus fall into two categories, tandem repeats and interspersed sequences. PGB 625 is the only tandem repeat family among the sequences. Among the interspersed sequences PGB 107, PGB 662 and PGE 096 have very complex distribution (complex interspersed sequences) and the remaining interspersed sequences are relatively less dispersed (simple interspersed) in pearl millet genome.

Southern blots of pearl millet genomic DNA following digestion with different enzymes were probed with PGE 123, PGB 582 and PGB 727 to examine the organization of the basic repeat unit. Some of the enzymes especially tetra-nucleotide sequence recognizing enzymes generated more than one fragment with sizes smaller than the insert and signal intensities of the fragments observed in such profiles were not proportional to their sizes suggesting that the basic unit of these clones are made up of different sub-repeats and their relative proportion vary.

Molecular analysis of *Pennisetum* species differing in their basic chromosome number (P. *ramosum* x=5; P. *mezianum* x=8; P. *orientate* x=9 and an allotetraploid P. *purpureum* x=7 with one genome common to P. *glaucum x=7)*) was carried out with pearl millet repetitive DNA sequences. Genomic DNAs from 19 accessions of pearl millet, including morphological marker stocks, male sterile lines, chlorophyll mutants and downy mildew resistant lines, were analyzed with the repetitive DNA probes.

Distribution and modulation of repetitive sequences among *Pennisetum* species were examined following Southern blot hybridization of cloned pearl millet sequences with sheared ^{32}P labelled genomic DNAs from five *Pennisetum* species. Signal intensity of each cloned fragment resulting from hybridization with genomic DNA from

each species as probe was estimated. The ratio of signal intensity of each insert to that of PGP 005 was calculated for the comparison of signal intensities within and between species. Simple interspersed sequences had almost the same level of abundance across the species. Differences in the relative abundance of some of the simple interspersed sequences is indicative of differential amplification of those sequences reflecting their chromosome / segment(s) specificity among the *Pennisetum* species which are secondary polyploids. Reiteration of sequences related to PGE 096, PGB 625 and PGB 662 vary over 10- fold across species which is a reflection of the modulation of these sequences during evolution. The abundance of PGB 625 is conserved while PGB 662 is deamplified in *P. purpureum* which is an allotetraploid.

Molecular analysis of genomic DNA of *Pennisetum* different species was carried out with repetitive DNA probes in combination with eight restriction enzymes (Eco RI, Pst I, Bam HI, Eco RV, Sac I, Bgl I, Hind III and Msp I) to examine the nature and extent of variation of sequences related to cloned pearl millet repetitive sequence. Restriction fragment length variation revealed by different probe - enzymes combinations differed markedly. Eco RI, Bam HI and Msp I revealed maximum polymorphism. Sequences related to PGE 015 and PGE 123 diverged maximally while sequences related to PGB 058 are highly conserved. The other sequences showed almost the same level of divergence. Thus different repetitive sequences differ in the extent of their sequence variation/divergence. Out of the total fragments detected by nine simple interspersed sequences with eight enzymes, 45% were conserved across the species, over 33% were common to two or more species and over 21% are species - specific. Species specific fragments and large proportion of conserved fragments between two or three species were faint indicating variation in a minor proportion of the sequences. Variation in the number of shared fragments between species may represent their phylogenetic relationship. Probe-enzyme combinations which revealed species - specific fragments were identified for use as diagnostic tools to identify chromosome(s) in alien genetic background.

The sequence divergence patterns revealed by the restriction polymorphisms were used to look at pair-wise similarity index. *P. purpureum* has 97% similarity with P. *glaucum*. Among the other species P. *orientate* is more related (87%) to P. *glaucum* than to P. *ramosum* and P. *mezianum* (~66%). Distribution, modulation and rearrangements of repetitive sequences among cereal (pearl millet, sorghum, maize, rice, wheat, barley, oats and rye) genomes were examined using cloned pearl millet repetitive sequence probes. Southern blots of cloned pearl millet repetitive DNA sequence inserts were hybridized with sheared ^{32}P labelled genomic DNA from cereals and the hybridization signal intensity of each cloned fragment in relation to that of PGP 005 was estimated. Two to three fold less abundance of simple interspersed sequences (PGP 005, PGE 015, PGE 123, PGB 058, PGB 074, PGB 103, PGB 582, PGB 727 and PGB 788) in barley, oats and rye is indicative of the dilution effect of these sequences with the increase in their genome size. High abundance of these sequences in rice, which has smallest genome size among cereals, signifies the conservation of these sequences. Although the abundance of the simple interspersed sequences varied, none of the sequences showed an order of magnitude variation. The high copy pearl millet sequences PGB 625 and PGB 662 are at very low level while sequences related to PGE 096 and PGB 107 are undetectable among cereal genomes. Restriction analysis also revealed the low copy nature of PGB 625 and PGB 662 in other cereals, suggesting modulation of such sequences during species divergence..

Molecular analysis of cereal genomes was carried out to investigate the nature and extent of variation of sequences related to pearl millet repetitive sequences among cereals. The relative levels of divergence/conservation of simple interspersed sequences were comparable to that of *Pennisetum* species. This substantiates the earlier observation that the simple interspersed sequences have different levels of conservation and tolerance to variation, and that they are under selection pressure. Out of the total number of fragments detected by nine interspersed sequences, 54% were shared by two or more cereals, 45% were species - specific and 0.7% were conserved across cereals. It is evident from the large proportion of intense shared and species-

specific fragments that the ancestral configurations are replaced with the variants from different lineages during evolution. Some of the variants are conserved in a specific group of cereals while other variants are of random nature among the cereals. The number of conserved variants in a group of cereals may represent their phylogenetic relationship to other such groups.

The percentage similarity of the sequences estimated from the restriction polymorphism among cereals is consistent with the taxonomical classification. Wheat and rye which come under sub-tribe Triticineae of the tribe Triticeae are closely related (86%) and these two cereals as a group have more similarity to barley (62%) within the tribe Triticeae, than to oats (46%) of the tribe Aveneae. These four cereals belonging to the sub-family Festucoideae formed a group and the other four tropical cereals formed another group and these two groups have ~20% common stem. Sorghum and maize of the tribe Andropogoneae sub-family Panicoideae are more related (80%) and this group in turn has 52% similarity with *Pennisetum* (tribe Paniceae) of the same sub-family. The present study showed that rice of the tribe Oryzeae sub-family Oryzoideae has more affinity to sub-family Panicoideae.

Hundred percent conservation of simple interspersed sequences among pearl millet accessions, 45% among *Pennisetum* species and nearly complete divergence among related cereal genomes, are indicative of their slow evolutionary turn over and the proportionality of divergence with increase in phylogenetic distance. These properties along with their comparable levels among the cereals make them ideal tools for the study of cereal genome evolution and the molecular phylogeny of cereals. The results obtained in the present study clearly demonstrated that repetitive sequences can be effectively used for the phylogenetic analysis of distantly related taxa irrespective of their genome size and repetitive DNA contents. Restriction analysis of other cereal genomes with simple interspersed sequences could reveal diagnostic probe - enzyme combinations to identify different species and to look at their phylogenetic relationship. The simple interspersed sequences developed in the present study thus have a blend of properties which make them `universal cereal molecular

markers' to use as species- / genome - specific markers in the development of saturated maps of cereal genomes.

INTRODUCTION

Studies employing DNA:DNA reassociation kinetics showed that eukaryotic genomes consist of various types of sequences which are present in many copies (Britten and Kohne 1968). Soon it was established that such repetitive sequences are common in eukaryotic genomes and constitute larger proportion of the genomes (Southern 1970; Chooi 1971; Wu et al. 1972; Flamm 1972; Miksche and Hotta 1973; Smith and Flavell 1974; Flavell et al. 1974). Genomes of eukaryotes are characterized by a tremendous range of nuclear DNA content even among closely related members of phylogenetic lineages (Bennett and Smith 1976). A substantial amount of the variation is attributable to the variation in the repetitive DNA of their genomes (Bennett et al. 1982; Bouchard 1982). The proportion of repetitive DNA sequences among major cereal genomes varies from 50% in rice (Despande and Ranjekar 1980) to 92% in rye (Flavell 1974).

1.1 Organization of Repetitive DNA Sequences:

Members belonging to several families of repetitive sequences have been cloned and characterized over the past one and a half decades since the report of cloning the rye heterochromatin sequences (Bedbrook et al. 1980a). Two methods, both exploiting the redundancy of repetitive sequences in the genome, have generally been employed to isolate the repetitive sequences. Repetitive sequences resolve as thick discrete bands on agarose gels when genomic DNA is digested with restriction enzymes. Genomic library highly enriched with specific repetitive sequences can be constructed by eluting DNA fragments from such bands and cloning (Jones and Flavell 1982b; Guidet et al. 1991; De Kochko et al. 1991; Wu and Wu 1992; Kamm et al. 1994; Grebenstein et al. 1995). In the other method, the genomic library is screened with labelled total genomic DNA or fractionated genomic DNA with a particular Cot value and the colonies which give the strongest hybridization signal intensities are selected as clones of bacterial cells harboring putative repetitive DNA (Perez-Vicente et al. 1992; Dobrzanska and Szurmak 1992; Szurmak and Dobrzanska 1993; Mao et al.

1994; Kiefer-Meyer et al. 1995; Mawal et al. 1995; Li et al. 1996)

Based on the copy number, the repetitive sequences are classified into highly repetitive and moderately repetitive sequences. It appears that this classification of repetitive sequences is not entirely straight forward and no hard and fast rule is available to demarcate the sequences into highly repetitive and moderately repetitive based on the reiteration frequency. Rather than any particular level of repetition frequency, certain features of genomic organization of repetitive sequences are generally employed to classify the repetitive sequences (Bouchard 1982). Two basic types of organization have been observed for repetitive DNA sequences. The repeating units can be arranged in tandem arrays or interspersed with unrelated repetitive or unique DNA sequences (Flavell 1980; Lapitan 1992).

Genome organization of a family of repetitive DNA sequences is investigated by Southern analysis of source DNA with cloned sequences belonging to the family of repetitive sequences. Ladder of fragments (the progression of a basic unit) is a diagnostic feature of tandem repeats (Bedbrook et al. 1980b; Belostotsky and Ananiev 1990; De Kochko et al. 1991; Wu and Wu 1992; Ingham et al. 1993; Kamm et al. 1994; Grebenstein et al. 1995). Restriction profiles generated by interspersed sequences usually consist of either one fragment equal to the size of the cloned insert or a set of fragments or a smear as per the interspersion and distribution of the family in the genome (Evans et al. 1983; Sonina et al. 1989; Guidet et al. 1991; Mao et al. 1994; Kiefer-Meyer et al. 1995; Li et al. 1996).

1.1.1 Tandem repeats:

Tandemly repeated DNA sequences usually represent the most highly repetitive sequence families in eukaryotic genomes (Lapitan 1992). Such repetitive sequences have been cloned and characterized from cereals such as rye (Bed brook et al. 1980a; Appels et al. 1986), *Oryza* species (Zhao et al. 1989; De Kochko et al. 1991; Wu et al. 1991; Wu and Wu 1992; Li et al. 1996), barley (Belostotsky and Ananiev 1990;

Vershinin et al. 1990), wheat (Metzlaff et al. 1986), maize (Peacock et al. 1981), pearl millet (Ingham et al. 1993; Kamm et al. 1994) and perennial oats (Grebenstein et al. 1995).

The four families of tandem repeat sequences reported by Bedbrook et al. (1980a) occupy 8-12% of the *Secale cereale* genome One of them with a 120bp repeat length constituting 2.4% of the genome has the highest copy number (1.5 x 10^6). The HvRT family of barley has 7 x 10^5 copies and comprises about 1.7% of the barley genome (Belostotsky and Ananiev 1990). The satellite DNA sequence of maize which is confined to knob heterochromatin has 1.2 x 10^6 copies per genome (Peacock et al. 1981).

Length of the basic repeat unit of repetitive sequence varies among different families. The HvRT family of barley consist of a 118 bp repeat length (Belostotsky and Ananiev 1990) and rye has a 120 bp repeat family (Bedbrook et al. 1980a). In rice, families with 352, 355, 498 and 756 bp repeat units have been reported (Wu et al. 1991; De kochko et al. 1991). Repetitive sequences located only at the knob heterochromatin of maize has 185 bp long repeats (Peacock et al 1981). Highly repetitive DNA component of perennial oats, *Helictotrichon,* comprises a 356 bp basic repeat unit (Grebenstein et al. 1995). The basic repeat lengths of tandem repetitive sequences are not completely identical within a species. However, they usually show more than 90% homology among them (Appels et al. 1986; Belostotsky and Ananiev 1990; Wu et al. 1991; Li et al. 1996). Differences among repeat units arise from base substitution, deletion and insertion during evolution (Appels et al. 1986; Lapitan 1992; Kamm et al. 1994).

The tandem repetitive sequences are characterized by their modulation among related species. In the genus *Oryza,* the distribution and modulation of sequences isolated from one genome type has been extensively studied in different genomes and accessions. Although the copy number varied from 290 (*O. rhizomatis* - CC genome) to 78,000(0. *alta* -CCDD genome), the sequences homologous to the tandem repetitive sequences isolated from *O. alta* were found tandemly arranged in genomes

AA, BB, BBCC, CC and EE (Li et al. 1996). The copy number of the CC-1 sequence cloned from 0. *officinalis* (CC-genome) showed a 1000 fold difference in copy number between two CCDD genomic species, 0. *alta* (15,000 copies) and 0. *latifolia* (15 copies). The same sequence was not detected in EE and FF genomes (Wu and Wu 1992). Genome specific tandem repeats have also been reported in rice (Zhao et al. 1989; De Kochko et al. 1991). The 'AA' genome specific repetitive sequence called pOs48 exhibits considerable variation in its copy number among different varieties of Asian cultivated rice (0. *sativa)* and related species within the 'AA' genome type (Zhao et al. 1989). The four repetitive sequences cloned and characterized from *Secale cereale* are not detected in S. *sylvestre* (Bedbrook et al. 1980a). The highly repetitive tandem arrays of perennial oats (*Helictotrichon convolutum)* are present in high abundance in DNA of species belonging to sub-genera of *Helictotrichon* and less abundant in sub-genera *Pubavenastrum* and *Pratavenastrum*. The same sequence is not detectable in species of *Arrhenatherum* and *Avena (* Grebenstein et al. 1995).

Among the cereals, there are few reports of tandem repeats which are detected beyond the boundary of a genus. The 120bp family of rye is present at high abundance in wheat and barley (Bedbrook et al. 1980a; Gupta et al. 1989). Similarly, the sequences related to the 350bp family of rye are detected in *Agropyron* species (Xin and Appels 1987). Grebenstein et al. (1995) reported that tandem repeat of perennial oats has 74% sequence similarity to an 'AA' genome specific repetitive DNA of *Oryza.*

1.1.2 Interspersed sequences:

A large proportion of higher plant genomes consists of repetitive DNA sequences interspersed with other repetitive sequences or low copy sequences (Flavell 1980; Lapitan 1992). Compared to tandem arrays, interspersed sequences have larger-repeating unit and less number of copies in the genome. In rye, the 630bp long Xba I family called RXX630 has 10^5 copies per genome (Mao et al. 1994) and the

monomer of dispersed sequence-R173 is 3.5Kb long and has 1.5 x 10^4 copies per genome (Guidet et al. 1991; Rogowsky et al.1991). In rice, two families of dispersed sequences (11-7-1 and 11-3-1) have 1900 and 800 copies per genome (Mawal et al. 1995) while another family of dispersed sequence specific to chromosome 5 has 900 copies (Wang et al. 1995). The 1.8 Kb long dispersed sequence called 'dialect-I' cloned from *Hordeum vulgare* has 5000 copies per genome (Sonina et al. 1989). Four families of dispersed sequences pHch 1, pHch 2, pHch 4 and pHch 5 from *Hordeum chilense* have 2.6, 2.1, 2.6 and 2.0 Kb repeat lengths respectively (Hueros et al. 1990). Two dispersed sequences with 1.3 and 1.4Kb repeat unit lengths have been partially characterized in *Sorghum vulgare* (Kumar et al. 1990). A highly repetitive interspersed sequence present in several *Avena* species has been reported from *Avena sativa* (Fabijanski et al. 1990). The 371 bp and 1241bp long dispersed sequences cloned and characterized from wheat have 7600 and 30,000 copies respectively (Dobrzanska and Szurmak 1992; Szurmak and Dobrzanska 1993). In maize, 2.5Kb and 2.8Kb sequences, representing a family repeated 11,000 times and a 5.5Kb sequence of another family repeated 28,000 times per genome have been cloned and characterized (Berlani et al. 1988a and 1988b).

The structure and organization of interspersed repetitive sequences are more complex and variable than those of tandem repetitive sequences. Interspersed sequences are distributed throughout the genome or are found only at a few chromosomal regions. Members of a family can be arranged as clusters of repeating units distributed at different genomic locations. Such arrangements are called 'clustered and scrambled' (Evans et al. 1983). The rye specific R173 family is distributed on all seven rye chromosomes and is interspersed all along the chromosome arms (Guidet et al. 1991). Members of R173 family occur generally as monomers and flanking regions of the repetitive motifs are different from each other comprising both repetitive and low copy sequences. (Rogowsky et al. 1991). In *Hordeum chilense,* members of the pHch 1 family are distributed on all the chromosomes, and both single units and small clusters are scattered throughout the

genome (Hueros et al. 1993). Wang et al. (1995) reported a moderately repetitive sequence from rice localized only at the centromeric heterochromatin of the long arm of chromosome 5.

Unlike the monomers of tandem arrays which are derivatives of a cardinal core sequence, the monomers of interspersed sequences are heterogeneous and have very complex organization. Direct repeats, inverted repeats, dyad symmetry and unrelated sequences are often seen within the repeating motif (McIntyre et al 1987; Hueros et al. 1993; Szurmak and Dobrzanska 1993; Kiefer-Meyer et al. 1995; Wang et al. 1995). In *H. chilense,* the 2.6Kb repeat of pHch 1 family is constituted by at least three classes of unrelated sequences which do not cross-hybridize (Hueros et al. 1993). The 1.4Kb long dispersed sequence cloned and characterized from *Thinopyrum elongatum* has two each of 186bp and 176bp direct duplications (McIntyre et al. 1987). The 1.7Kb long CCDD genome specific dispersed sequence cloned from *Oryza latifolia* has several direct and inverted repeats of varying lengths within the basic repeat unit (Kiefer-Meyer et al. 1995). The chromosome 5 specific sequence of rice contains two sub-repeats of 37bp and 19bp long sequences connected by 30 to 90 bp long sequences with high degree of similarity (Wang et al. 1995).

Compared to tandem repeats, the interspersed sequences have wider distribution extending beyond a genus. The 1.4 Kb family of sequences isolated from *Thinopyrum elongatum* has almost same abundance in *Agropyron cristatum, Pseudoroegneria spicata* and *Thinopyrum junceiforme* whereas it is in much smaller proportion in *Dasypyrum villosum, Psathyrostachys juncea, Secale cereale* and *Triticum aestivum* (McIntyre et al. 1987). Sequence homologous to the RS-I family of interspersed sequence of *Avena sativa* is distributed in all genome types of the genus *Avena* but with more proportion in C-genome containing species (Fabijanski et al. 1990). The 1241bp long interspersed sequence identified in wheat has same dispersion pattern in rye and also occurs in barley genome but with a different configuration. Sequence related to this family of sequences is absent in maize and oats (Szurmak and Dobrzanska 1993). Of the four interspersed sequences reported

by Hueros et al. (1990) from *Hordeum chilense,* three are present at very low proportion in wheat. The rye RXX630 family of interspersed sequences are also present in wheat genome, but is not detected in barley (Mao et al. 1994) 'Dialect-I' of barley (Sonina et al. 1989), 1.3Kb and 1.4Kb Xba I families of *Sorghum vulgare* (Kumar et al. 1990), pHch 1 family of *Hordeum chilense* (Hueros et al. 1990), R-173 family of rye (Rogowsky et al. 1991) and 1.7 Kb family of *Oryza latifolia* (Kiefer-Meyer et al. 1995) are genome specific from which they are cloned.

1.2 Applications of Repetitive Sequences in Basic and Applied Research:

Several lines of evidences suggest that the repetitive sequences are vital tools in phylogenetic analysis and tracing the origin of genomes in naturally occurring polyploids. Before the advent of molecular cloning techniques, repetitive DNA homology between related species estimated by re-association kinetics were used to assess the phylogenetic relationships. Flavell et al. (1977) constructed a phylogenetic tree of cereal (oats, barley, wheat and rye) evolution based on the proportions of repetitive sequences of each of the cereal genomes that are homologous with repeated sequences in the other species. It was proposed that wheat and rye diverged after their common ancestor had diverged from the ancestor of barley. This was preceded by the divergence of the common ancestor of wheat, rye and barley and the ancestor of oats. Molecular cloning techniques provide us with additional tools to characterize members of a repetitive DNA family. The distribution of such sequences among related taxa and the extent of sequence divergence can be examined by restriction analysis of genomic DNA with cloned DNA probes. Dvorak et al. (1988) pointed out the potential advantages of using variations in repetitive sequences in drawing phylogenetic inferences than chromosome pairing in interspecific hybrids or polymorphism in single-copy sequences. Repetitive sequences are at a number of loci or interspersed throughout the genome and thus mark simultaneously all or a number of chromosomes. Each sequence family comprises many individual members with

tendencies to homogenize; polymorphism generated slowly at such loci may, therefore, be more meaningful in interspecific comparisons than polymorphism at single copy loci. DNA probes are potential diagnostic indicators of species relatedness especially in instances in which traditional genome analysis by a cytogenetic approach cannot be done owing to lack of viable hybrids (Hoang-Tang et al. 1991). Phylogenetic inferences drawn from the study of the organization of repetitive DNA sequence families in related genomes are consistent with the existing taxonomic grouping (Gupta et al. 1989; Dvorak and Zhang 1992; Ferrer et al. 1995). The 120bp family of highly repetitive sequence from rye was used to study the species relatedness among 61 accessions representing 25 *Hordeum* species. Sequence homologous to 120 bp family was not detected in three closely related species *H. agriocrithon, H. spontaneum* and *H. vulgare* (Gupta et al. 1989). Similar conclusion was drawn by Ferrer et al. (1995) on the basis of the distribution of two repetitive sequences of *H. chilense.* The T genome represented by *H. vulgare* was least related to the remaining *Hordeum* genomes. Phylogenetic tree of 13 diploid species of *Triticum* based on the variations of 31 repetitive sequences was consistent with the cytotaxonomical and morphological grouping of these species (Dvorak and Zhang 1992).

Based on the variation in repetitive nucleotide sequences Dvorak and Zhang (1990) proposed a scheme for the identification of diploid species related to a specific genome of a polyploid species. This involves the identification of diagnostic bands (DB) and diagnostic intensities (DI) by analyzing the restriction profiles of a number of accessions of each relevant taxon and screening for such diagnostic bands or intensities in the polyploids under investigation. The 'repetitive nucleotide sequences identity' between the diploid and polyploid is then estimated by dividing the number of DBs plus DIs of a specific diploid that are observed in a polyploid with the total number of DBs and DIs in diploid. This scheme was shown to be promising in elucidating the origin of genomes in naturally occurring polyploids (Dvorak and Zhang 1990; Zhang and Dvorak 1991, 1992; Zhang et al. 1992).

Molecular analysis using repetitive sequences has been employed to

complement information gathered by traditional cytogenetic analysis (Talbert et al. 1993; Rayburn and Gill 1986). The 'D' genome - specific repetitive sequence called Dgas44 from *Triticum tauschii* is evenly distributed on all chromosomes of D genome, and is a powerful tool in detecting 'D' genome in polyploid *Triticum* species (McNeil et al. 1994). Repetitive sequence probes are vital tools in hybrid genome analysis (Mclntyre et al. 1987; Perez-Vicente et al. 1992; Ferrer et al. 1995; Svitashev et al. 1995; Francis 1995) in tracing the chromosome segments carrying valuable traits in alien translocation lines (Rogowsky et al. 1991; Kim et al. 1992; Bournival et al. 1994) and in species/cultivar identification (Harcourt and Gale 1991; Liu et al. 1992; Pecchioni et al. 1993).

1.3 Scope of the Study

Many families of repetitive sequences (both tandem and interspersed) have been cloned and characterized from several Gramineae species. Uses of such sequences as molecular markers in phylogenetic analysis, in identifying putative parents of naturally occurring polyploids, species/cultivar identification, and in tracking chromatin in interspecific hybrids and chromosomes in addition lines have been well documented. In pearl millet, only a monomer of a family of tandem repeats has been cloned and characterized so far (Ingham et al. 1993; Kamm et al. 1994). No interspersed sequences have been reported in pearl millet. Ingham et al. (1993) identified a 140bp long Kpn I family of satellite sequences in pearl millet. The 137 bp long Hae III family of sattellite DNA reported by Kamm et al. (1994) is homologous to Kpn I family. The Hae III family is present as tandem arrays in the wild species P. *mollissimum* and P. *violaceum*. *In situ* hybridization experiments revealed centromeric localization of this sequence on all seven chromosomes and indicated chromosome specific difference in copy numbers. Sequences related to Hae III variants was not detected in foxtail millet, wheat, rye and barley. One family of Ac/Ds like transposable element was found in pearl millet genome and its relationship to Ac/Ds element in

related cereals was analyzed (Huttley et al. 1995). An RFLP based genetic map of pearl millet covering a map distance of 303cM was constructed (Liu et al. 1994) and quantitative trait loci for downey mildew resistance are mapped (Jones et al. 1995).

Several research groups have been working on comparative mapping of cereal genomes and such maps of sorghum and maize (Whitkus et al. 1992), wheat, barley and rye (Devos et al. 1993), rice and maize (Ahn and Tanksely 1993), rice and wheat (Kurata et al. 1994b), wheat and barley in comparison to rice, maize and oats (Van Deynze et al. 1995a), oats in comparison to wheat, rice and maize (Van Deynze et al. 1995b), comparisons of linkage segments of rice with syntenic blocks (rice blocks) of Triticeae, maize, sorghum, sugar cane and foxtail millet (Moore et al. 1995) and rice and barley (Saghai-Maroof et al. 1996) have been reported. Comparative maps provide information on gene order conservation (synteny) in a group of taxa, the fate of duplicated genes and chromosomal segments in aneuploids and polyploids, and the genomic contribution of parental taxa to presumptive hybrid species and introgressed populations. Based on the information from the comparative maps of cereal genomes, it has now become possible to examine the genomic rearrangements occurring during the course of evolution and the detection synteny among cereal genomes. Such studies, as proposed by Helentjaris (1993), would be helpful in designing strategies in the development of saturated maps. Such comparative analysis of cereal genomes would result in a 'generic' Gramineae map and would add valuable insights into chromosome evolution and grass speciation.

In spite of its considerable value in food, feed and forage production, pearl millet has not received the attention of researchers on par with other cereals. Genetic map distance of 303 cM (Liu et al. 1994) is far below the expectation when one compares the pearl millet genome size (2.5 pg) to that of rice (0.5 pg) showing 1575 cM (Kurata et al. 1994a). More saturated map is desirable for efficient use in pearl millet genetic studies and breeding programmes. No attempt has been made so far to study the nature and extent of genomic homoeology/synteny between pearl millet and other cereals. Hulbert et al. (1990) reported that molecular markers from

maize showed weak or no on hybridization signal with pearl millet DNA. An understanding of phylogenetic relationships between pearl millet and other cereals would be helpful in choosing a common set of molecular markers from other cereals for the construction of saturated maps in pearl millet.

Objectives of the present study were:

- To clone and characterize major families of pearl millet repetitive DNA sequences.
- To explore the utility of pearl millet repetitive DNA sequences as molecular markers.
- To examine the nature, conservation and modulation of repetitive sequences among related species and cereals.
- To explore the utility of pearl millet molecular markers in phylogenetic analysis of cereal genomes.

MATERIALS AND METHODS

2.1 Plant Materials:

Plant materials used in the present study are listed in Table I. Seedlings of pearl millet accessions and other cereals were raised in glass house. For DNA extractions, shoots from 3 weeks old seedlings were collected and shoots from other *Pennisetum* species, which are perennials, were collected 3 weeks after pruning.

2.2 Construction and Screening of Genomic Library

2.2.1 Extraction and purification of genomic DNA:

Total genomic DNA was extracted from leaf tissue according to Sharp et al. (1988). Shoots were frozen in liquid nitrogen and ground to a fine powder using pestle and mortar. The powder was transferred into a conical flask and suspended in 5ml extraction buffer (100mM Tris - HCl [pH 8.5], 100mM NaCl, 50mM EDTA, 2% SDS) per gram of leaf tissue and incubated at 37 °C for 2.5h . An equal vol of 1:1 ratio of Tris buffered phenol : CHISAM (24 chloroform : 1 isoamyl alcohol) was added into the lysate, stirred for 10 min and centrifuged at 12,000 x g for 15 min. The supernatant was collected and DNA was precipitated with 0.6 volume of iso-propanol. DNA was spooled, washed in 70% ethanol, dried under vacuum and dissolved in 1 x TE (10mM Tris-HCl [pH 8.0], 1mM EDTA [pH 8.0]. DNA was treated with RNase A (50µg/ml final concentration) followed by proteinase K (100 µg/ml final concentration) at 37 °C for 2h. DNA was extracted twice with equal volume of 1:1 ratio of Tris - buffered phenol : CHISAM and once with an equal vol of CHISAM alone. DNA was reprecipitated by adding 1/10 volume of 3M sodium acetate (pH 5.2) and 2 volume of chilled (-20°C) absolute ethanol. DNA was spooled, washed with 70% ethanol, vacuum dried for 5 min, and dissolved in 1 x TE.

The purity of the DNA samples was determined spectrophotometrically by measuring the absorbance (O.D) at 280nm, 260nm and 230nm. DNA was quantified considering one O.D.U at 260 nm as 50 µg of double stranded DNA. The A_{260}/A_{280} ratio

Table I: Plant material used

Species / accession	Character	Source
P. glaucum (2n = 2x = 14)		
Green sibs of albino		Prof. N. C. Subrahmanyam Univ. of Hyderabad, INDIA.
GWS-14	Green/white striping	Prof. J.L. Minocha PunjabAgri.University Ludhiana, INDIA.
VCM-36	Green/white striping	GRU, ICRISAT, Patancheru, A.P., INDIA
IP 5009	Yellow stripe variegation	"
IP 9712	Yellow stripe variegation	"
IP 8166	Purple plant	"
IP 8275	Glossy leaves	"
IP 10401	dwarf	"
IP 12598	Branched ear base	"
IP 12617	Branched ear base	"
PMC 23A	Cytoplasmic male sterile	"
DSA 59-1A	Cytoplasmic male sterile	"
DSA 118A	Cytoplasmic male sterile	"
DSA 134A	Cytoplasmic male sterile	"
DSA 144-1A	Cytoplasmic male sterile	"
ICMA 88004	Downy mildew resistant	"
IP 18292	Downy mildew resistant	"
IP 18293	Downy mildew resistant	"
IP 18294	Downy mildew resistant	"
P. ramosum (2n=2x=10)	-	"
P. mezianum (2n=4x=32)	-	"
P. orientale (2n=2x-18)	-	"
P. purpureum (2n=4x=28)	-	"
Sorghum bicolor cv. CSH 11 (2n=2x=20)	-	"
Zea mays SCH. CM 119xCM 120 (2n=2x=20)		Dr. R. Sai Kumar, Maize Research Station A.P.A.U., India.
Oryza sativa cv. basmati (2n=2x=24)		Directorate of Rice Research ICAR, A.P., INDIA..
Hordeum vulgare cv. betzes (2n=2x=14)		Prof. A. Hagberg Swedish Univ. of Agric. Sciences, Svalov, Sweden
Triticum aestivum cv. Chinese spring (2n=6x=42)		IARI, NEW DELHI, INDIA
*Avena sativa (2*n=6x=42)		Dr. N.N. Singh, IARI, NEW DELHI, INDIA.
*Secale cereale (2*n=2x=14)		"

was found to be > 1.8. The quality of DNA was checked on 0.8% agarose gel electrophoresis.

2.2.2 Preparation of genomic DNA fragments for ligation:

Genomic DNA of *P. glaucum* IP10401 was separately digested with 5U/μg Bam HI and Eco RI in 20 μl assay vol at 37 °C for 3 h. After adding 205fil TE (pH 8.0) the digested samples were purified by extracting once with 1:1 ratio of phenol and CHISAM, and once with an equal vol of CHISAM. After the final extraction, aqueous phase was collected and 25 μl of 4M LiCI was added. The DNA was precipitated by adding 750 μl chilled (-20 °C) ethanol at -20 °C for 30 min and pelleted at 15,000 rpm (JA 18.1 Beckman rotor) at 4 °C for 15 min. The pellet was washed in 70% ethanol, dried under vacuum and dissolved in TE to a final concentration of 250ng/μl. The quality of the digested DNA was checked on 0.8% agarose gel.

2.2.3 Preparation of vector DNA for cloning:

Aliquots of pUC 18 plasmid (Bohringer Mannheim) were digested separately with 5U/μg each of Bam HI and Eco Rl and purified as given in 2.2.2.

2.2.4 Dephosphorylation of vector DNA:

The linearized vector (pUC18) DNA was treated with 0.25U/μg of calf intestinal phosphatase (Pharmacia) at 37 °C for 2h. The enzyme was heat inactivated at 85 °C for 10min. Plasmids were purified as described in 2.2.12. and quality of the linearised dephosphorylated vector was checked on 0.8% agarose gel

2.2.5 Ligation of genomic DNA to pUC18 DNA:

The ligation mixture contained 300μg of insert DNA, 100ng vector DNA and 0.2U of T4 DNA ligase in a final vol of 10 μl. Ligation was carried out at 15 °C for 12h.

2.2.6 Preparation of culture media and plates:

LB liquid medium : 1% bactotryptone, 0.5% bactoyeast extract and 0.5% NaCl were dissolved in double distilled water, suitable aliquots were dispensed into conical flasks, plugged with cotton and autoclaved.

LB agar plates: To LB liquid medium 1.5% bactoagar was added and autoclaved. After cooling to 50 °C filter sterilized stock solution (25 mg/ml) of ampicillin was added to the medium (100μg/ml final concentration) and 25ml was dispensed into each petriplate. The plates were kept for 20min, sealed with parafilm and stored at 4 °C. Screening plates contained IPTG (40ng/ml) and X-gal (32μg./ml) in addition to ampicillin.

2.2.7 Preparation of competent *E. coli* cells.

E. coli JM109 cells were made competent by $CaCl_2$ treatment following Cohen's Method (described in Sambrook et al. 1989). Single colony of *E. coli* JM109 was inoculated in 5ml LB medium and incubated overnight at 37 °C on a platform shaker at 200 rpm. 40 ml of LB medium was inoculated with 400 μl of the fresh overnight culture and incubated at 37 °C for 2 h on a platform shaker at 200 rpm. The cells were harvested by centrifugation at 5000 rpm (Sorvall SS-34 rotor) at 4 °C for 10 min and the bacterial pellet was suspended in 8ml of ice cold 0.1M $CaCl_2$ solution. The resultant cells were stored at 4 °C.

2.2.8 Transformation of competent *E. coli* JM109 with the ligated mixture:

In a thin walled test tube kept on ice, 2μl ligated mixture was made upto 10μl with the addition of 50mM Tris Cl (pH 7.12), mixed with 200 μl competent cells and kept on ice for 40 min. The mixture was subjected to heat shock at 42 °C for 3 min. After 10 min at room temperature (~ 30 °C) 1ml of LB medium was added to the transformed cells and were allowed to recover at 37 °C for 40 min on a shaker at 200 rpm. Onto each plate containing LB/Amp/X-gal/IPTG, 400μl of the transformed *E. coli* were plated

and incubated at 37 °C for 17 h. Colonies were selected on the basis of their colour. Cells with recombinant plasmids give white colonies whereas cells with self ligated pUC 18 give blue colonies on IPTG/X-gal agar plates. The transformation efficiency was found to be 4.9 x 10^5 colonies/jig of circular pUC 18 control DNA.

2.2.9 Replica plating:

Gridded nylon membrane was placed over a pre-made LB Agar/Amp plate and the white colonies (cells with recombinant DNA constructs) were transferred onto it with sterile tooth-picks. The petriplates were incubated at 37 °C for 18 h to enable colonies to grow on the master nylon discs. The colonies on the gridded master disc were replica plated onto fresh nylon membrane discs which were transferred to LB Agar/Amp plates and incubated at 37 °C for 18 h.

2.2.10 *In Situ* lysis of bacterial cells on replica filters:

In situ lysis was carried out according to the procedure of Grustein and Hogness (described in Sambrook et al. 1989). Whatman (3M) papers in four plastic trays were saturated with one of the following solutions I) 10% SDS ii) denaturing solution (0.5 N NaOH, 1.5M NaCl) iii) neutralizing solution (1.5M NaCI, 0.5M Tris Cl [pH 7.4] iv) 2x SSC. Each replica membrane was peeled off from the petriplate and placed onto Whatman papers prewetted with 10% SDS for 3 min and shifted to denaturing solution for 5 min followed by neutralizing solution for 5min and 2x SSC for 5 min. The filters were air dried for 30 min and the DNA was immobilized by baking at 80 °C in a vacuum oven for 2 h.

2.2.11 Colony hybridization:

The replica membranes were hybridized with raido-labelled genomic DNA of P. *glaucum.* Membrane discs with immobilized DNA were placed in 2x SSC for 5min and transferred to prewashing solution (5x SSC, 0.5% SDS and 1mM EDTA [pH 8.0]) and

incubated at 50 °C for 30 min. Bacterial debris on the nylon membrane discs was scraped off with clean gloved hand. Prehybridization was carried out in a solution containing 10% Denhart's solution, 2x SSC , 1% SDS, 5mM phosphate buffer (pH 8.0) and 100 jag/ml of sheared denatured salmon sperm DNA at 65 °C for 2 h.

Approximately 200ng of pearl millet genomic DNA was sheared to give DNA fragments of 2-6 Kb in size and radio-labelled using the random prime labelling kit according to the manufacturer's (BRIT-INDIA) instructions. The labelled genomic DNA was added to the prehybridization solution and hybridization was carried out at 65 °C in a shaker water bath for 24 h.

Following hybridization, the nylon discs were washed twice in solution I (1xSSC, 1% SDS) for 15min each and twice in solution II (0.1 x SSC, 1% SDS) at 65 °C for 1 h each with constant agitation.

The membranes were dripped dry and wrapped in Saran wrap. X-ray film (INDU) was exposed to hybridized membrane in a suitable X-ray cassette and kept at -70 °C. Autoradiograms were developed after 7 days of exposure.

2.2.12 Amplification and extraction of plasmid DNA.

Plasmid was extracted according to the alkali lysis minipreparation method of Birnboin and Dolly (1979) (described in Sambrook et al. 1989). 10ml of LB medium containing 30μg/ml of ampicillin was inoculated with an isolated colony of transformed bacteria and incubated at 37 °C for 12-16h on a platform shaker at 200 rpm. The cells were harvested at 5,000 rpm (in a Sorvall SS-34 rotor) at 4 °C for 10 min and the bacterial pellet was suspended in 200 μl of ice cold solution I (50mM glucose, 25mM Tris HCl [pH 8.0], 10mM EDTA [pH 8.0] and transferred to a 1.5 ml microfuge tube, vortexed well and 400μl of solution II (0.2N NaoH, 1% SDS) was added and mixed by inverting the tube rapidly for 5 times. The suspension was further vortexed after the addition of 300 μl of ice-cold solution III (5M potassium acetate, glacial acetic acid; and centrifuged at 12,000 rpm for 5 min in the Sorvall 24 S microfuge. 600 μl of the

supernatant was collected and extracted with 1:1 ratio of phenol and CHISAM. It was centrifuged in the microfuge as described before and 500 µl of the aqueous phase was collected. DNA was precipitated by adding 1ml of chilled (-20 °C) ethanol and pelleted at 12,000 rpm for 5 min in the microfuge. The DNA pellet was washed with 70% ethanol, dried under vacuum and dissolved in 25µl of TE. The plasmid DNA was treated with 100 µg/ml (final concentration) of RNase A at 37 °C for 30 min and its quality and quantity were checked on 0.8% agarose gel.

Recombinant plasmids were also extracted following the same procedure. The recombinant plasmids were restricted to excise the insert, the resulting fragments were size fractionated on agarose gels and stained in 0.5 µg/ml ethidium bromide solution for 30 min. The insert sizes were determined by the gel documentation system (UVP INC San Gabrriel) with Hind III digest of λDNA and Taq I digest of pUC 18 plasmid as size markers.

2.3 Blot Hybridizations:

2.3.1 Restriction and Southern analysis of cloned fragments:

Aliquots of recombinant plasmids restricted with the respective enzyme were fractionated on 1% agarose gels along with known quantities of linearised pUC18 (STD-DNA) the gel was stained with ethidium bromide and the area under the pUC18 band of each clone and STD-DNA was measured in a UVP-gel documentation system. The area of pUC18 band of each clone was compared with STD-DNA to ensure loading of equimolar concentration of each recombinant plasmid. Restricted recombinant plasmids each containing 50ng of vector (pUC18) were fractionated on agarose gel and vacuum blotted (LKB vacu gene vacuum bloting unit-Pharmacia) onto nylon membrane (NEN-Du Pont) according to the manufacturer's instructions. The blots were baked at 80 °C in a vacuum oven for 2 h.

Prehybridization was carried out in a polythene bag containing prehybridization solution (500 mM phosphate buffer [pH 7.4], 7% SDS, 1mM EDTA, 1% BSA, 100 jig/ml

denatured salmon sperm DNA) at 65 °C for 3-4 h. Approximately 300ng of genomic DNA was sheared to give DNA fragments of 2-6 *Kb* in size and radio-labelled using the random prime labelling kit according to the manufacturer's (BRIT-INDIA) instructions. Heat denatured labelled genomic DNA probe was added into the same prehybridization solution and hybridization was carried out at 65 °C in Haake's shaker water bath (100 rpm) for 16-18 h.

Following hybridization, the membrane blots were washed twice in solution I (0.2% SDS, 2x SSC) for 15 min each and twice in Solution II (0.2% SDS, 1 x SSC) for 30 min each at 65 °C in Haake's Shaker water bath. The blots were air dried, wrapped in Saran wrap and X-ray film was exposed to hybridized blots in X-ray cassette (Sigma) at ~70 °C. Exposure time was judged by monitoring the counts on membrane with GM Counter. Hybridization signal intensities of the clones were determined by the gel documentation system (UVP INC. San Gabrriel).

2.3.2 Restriction and Southern blot hybridization of genomic DNA:

DNA samples were digested with 5U of restriction enzymes (Bangalore Genei, Bangalore) per µg of genomic DNA at 37 °C for 3 h and size fractionated on 1% agarose gels. Blotting and hyrbridizations were followed as described in 2.3.1. Purified inserts of recombinant plasmids (purified using Gene Clean Kit - Bangalore Genei - Bangalore) were raido-labelled following BRIT random prime labelling protocol and used as probes.

2.3.3 Removal of the probe from the hybridized blot:

The probe was stripped off the hybridized blot by washing it in 500 ml solution A (0.25 N NaOH, 0.2% SDS) at 37 °C for 1h and in 500 ml of solution B (0.1x SSC, 0.1% SDS, 0.2M Tris Cl [pH 7.5] at 37 °C for 30 min.

2.3.4 Dot blot analysis:

Serial dilutions of genomic DNA (4μg, 2μg, 1μg, 0.5μg, 0.25μg, 0.125μg, 0.0625μg, 0.0312μg) and aliquots of linearised recombinant plasmids that can give 2ng, 1ng, 0.5ng, 0.25ng, 0.125ng, 0.0625ng, 0.03125ng and 0.01563ng pUC 18 (See 2.3.1) were blotted onto nylon membrane (NEN-Du Pont) using Bio-Dot Microfilteration Apparatus (Bio-Rad) following manufacturer's instructions. Strips carrying a set of serial dilutions of pearl millet genomic DNA and genomic clones were hybridized with corresponding probes as described in 2.3.1 and autoradiograms were developed. Density scans of autoradiograms were obtained using a UVP gel documentation system. Copy number of the cloned sequences in pearl millet genome were estimated from the number of inserts present in the aliquot of each linearised recombinant plasmid dot blotted which is equal to the number of pUC18 molecules present in the quantity of the pUC18 contained in the aliquot dot blotted (1.77×10^{13} /50 jig/ml as per Sambrook et al. 1989) since each recombinant plasmid releases one molecule each of vector and insert.

2.4 Computer Analysis

2.4.1 Comparison of restriction profiles:

Present investigation involved analysis of restriction profiles generated by different probe-enzyme combinations in different species and genera. A programme was written in C language. Probe-specific fragments are colour coded and hybridization profiles of fragments generated by each enzyme (used) are merged for each species. Species-wise lanes for each enzyme panel and enzyme-wise lanes for each species are used to facilitate comparative analysis.

2.4.2 Phylogenetic analysis:

The profiles of bands among the species for each probe - enzyme combination were visually compared. Each band was considered as an independent character for the phylogenetic analysis. Bands which were too faint or obscure were ignored.

Number of bands shared by two species revealed by a probe-enzyme combination reflects the relatedness of the species with respect to genome organization of the sequence homologous to the probe. Percentage relatedness (r) is calculated by using the formula r = 2xy/x+y X 100 where ' x ' represent total number of fragments in one species, ' y ' represent total number of fragments in the other species and ' xy ' represent number of bands shared by the two species. Percentage divergence of the species with respect to the sequence analyzed is '100 - r'. Divergence was computed in pair-wise combinations for all probe-enzyme combinations and used for the construction of phylogenetic trees using PHYLIP package.

RESULTS AND DISCUSSION

3.1 SCREENING AND CHARACTERIZATION OF GENOMIC LIBRARIES

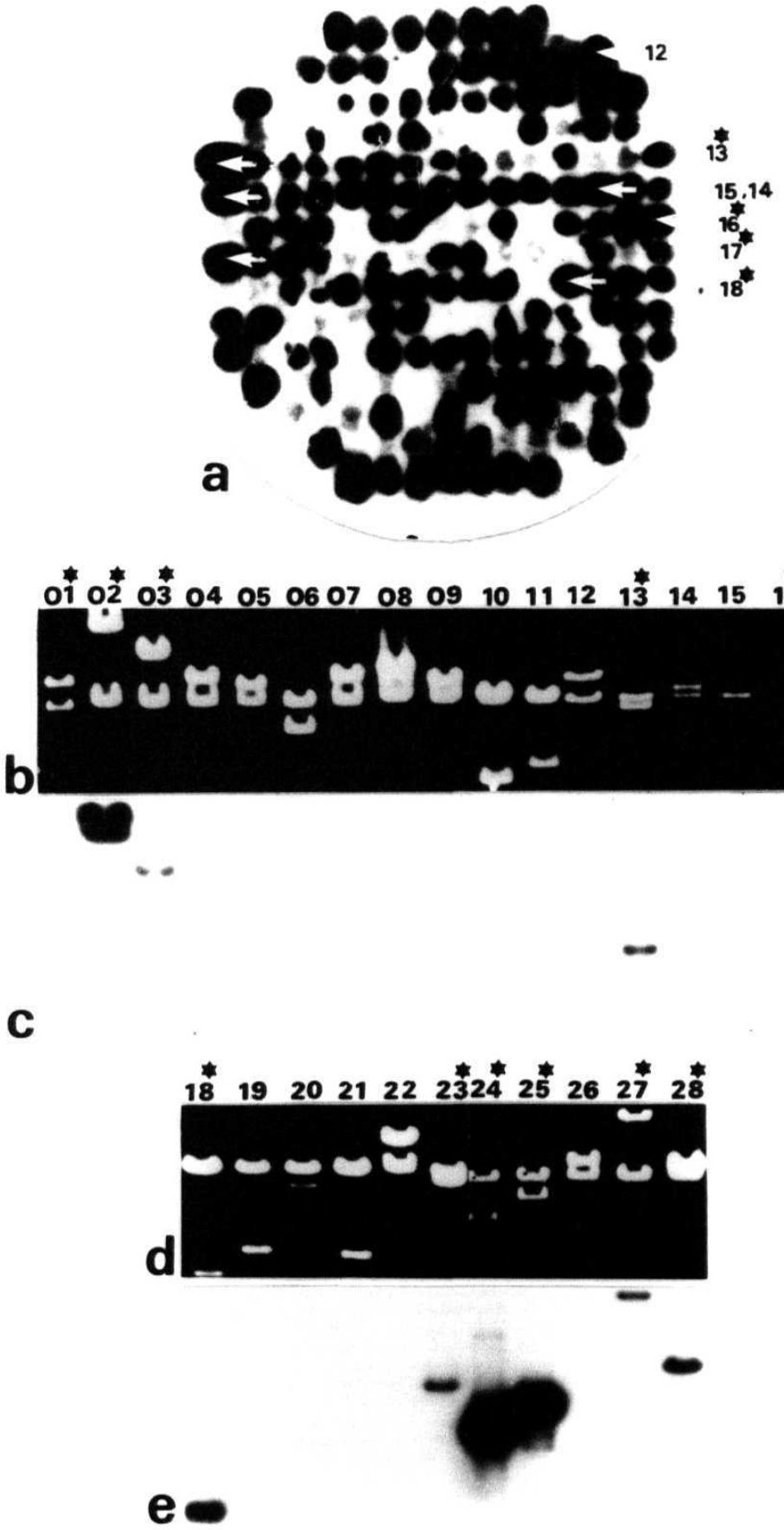
Construction of Genomic Libraries

Bam HI and Eco RI based partial genomic libraries of pearl millet were constructed (as presented in section 2.2) A total of 955 Bam HI and 576 Eco RI derived clones were screened. *In situ* lysed replicas of the genomic libraries were hybridized with ^{32}P-labelled sheared pearl millet genomic DNA. A representative autoradiogram of such hybridized replica membrane is given in Fig. 1a. Replica membranes carrying *E.coli* JM109 cells transformed with pUC18 was used as control and found to be blank (data not shown). Differences in signal intensities of hybridization could be used in choosing clones carrying repetitive sequences inserts. Nearly 150 colonies (~ 10% of the colonies screened) had very high signal intensities.

Variation in signal intensities observed among colonies following colony hybridization with pearl millet genomic DNA could be primarily due to the variation in the copy number of the sequence (insert) represented in the total genomic DNA used as the probe. Differences in the insert sizes and/or differences in the colony sizes emanating from the number of cells in the inoculum could also contribute to the overall signal intensity differences.

Seventy five colonies (5% of the total) which gave very high signal were selected for further screening. Recombinant plasmids were extracted following the alkali lysis method (see section 2.2.12) and their inserts were excised using corresponding enzymes. Resulting fragments were size fractionated on 1% agarose gels, stained with ethidium bromide and viewed under UV using a transilluminator (data not shown). Insert sizes ranging from 0.7 Kb to 9.1 Kb were found in 32 clones while 43 clones had no discernible inserts. Following Southern blot hybridizations of size fractionated

Fig 1. a. A representative autoradiogram of pearl millet partial genomic library following hybridization with sheared ^{32}P labelled pearl millet genomic DNA. Colonies that showed highest hybridization signal intensities were selected for further analysis (arrow heads indicate selected colonies from the master plate of the replica filter, lane numbers given on the right correspond to the clones in gel photograph b). Cells of the selected colonies were amplified, plasmid extracted, size fractionated on 0.8% agarose gels following digestion with corresponding enzyme and ethidium bromide stained (b & d). Lane numbers, clone designations as described in the text and fragment sizes are given in Table- 1. Southern blots from these gels were hybridized to sheared ^{32}P labelled total genomic **DNA** of ***P. glaucum*** **(c** & **e).** Asterisks indicate the clones with high copy DNA sequence inserts selected for further analysis.


12
13
15,14
16
17
18
a
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17
b
c
18 19 20 21 22 23 24 25 26 27 28
d
e

inserts from recombinant plasmids with ^{32}P -labelled sheared total pearl millet genomic DNA, inserts from 20 clones showed differences in their signal intensities in autoradiograms following 24 h exposure. While 16 inserts were detectable with ethidium bromide staining, inserts of four clones which were not detectable by staining also lighted up (data not shown). The probe was stripped off from the blots and rehybridized with probes (one at a time) prepared from the clones which lighted up early on Southern hybridization to check their cross homologies. Twelve out of 20 clones had distinct sequences which were detected even with ethidium bromide staining. Photographs of ethidium bromide stained size fractionated inserts following restriction of putative repetitive sequence clones and the low copy sequence clones with sizable inserts are presented in Fig. 1b & d. Insert sizes of these clones are given in Table II. Aliquots (100ng) of restricted clones, were size fractionated, Southern blotted and hybridized with 300ng of sheared labelled pearl millet genomic DNA. Autoradiograms are presented in Fig. 1c & e. Nine Bam HI - and three Eco RI- based clones carrying repetitive sequence inserts were selected for further analysis. Clone designations comprise three letters P,G, B/E followed by a number. The letter 'P' denotes *Pennisetum* and 'G' denotes genomic and B or E stands for the restriction enzyme (i.e., Bam HI/Eco Rl) used for the construction of libraries. The number corresponds to the colony number of the master genomic library. One of the clones (PGP005) with a 3.1 Kb insert from a Pst I-based genomic library of pearl millet constructed in our laboratory was found to be a high copy sequence. Southern hybridizations of labelled PGP005 with the 12 putative repeated sequence inserts revealed partial homology with PGB 727. Southern analysis following reciprocal probing of Bam HI restricted PGP 005 and Pst I restricted PGB 727 with probes prepared from linearised clones revealed a 2.0Kb common fragment between those two clones. PGP 005 was also included for further analysis.

Table II. Pearl millet genomic clones and insert sizes

Lane	Clone	Insert size (Kb)
01	PGE 015*	3.7
02	PGE 096*	9.1
03	PGE 123*	6 8
04	PGE 990	3.8
05	PGE 1004	3 2
06	PGE 1031	1.7
07	PGE 1032	3.5
08	PGE 1047	3.5
09	PGE 1106	3 3
10	PGE 006	0 7
11	PGE 119	0.9
12	PGB 008	3 4
13	PGB 058*	2 4
14	PGB 061	2.8
15	PGB 073	1.0
16	PGB 074*	1.7
17	PGB 103*	1.0
18	PGB 107*	0.8
19	PGB 113	1.1
20	PGB 132	2.2
21	PGB 143	1.0
22	PGB 161	3 4
23	PGB 582*	2.3
24	PGB 625*	1.4
25	PGB 662*	1.9
26	PGB 686	3 3
27	PGB 727*	7.1
28	PGB 788*	3 0

"High copy DNA sequence clones selected for analysis.

Characterization of the Clones

Estimation of copy numbers

Copy number was estimated by quantitative dot blot analysis. Dot blot strips each carrying serial dilutions of one of the 13 putative repetitive sequence clones restricted with the corresponding enzyme and pearl millet DNA was probed with the respective clone. From the density scans of the autoradiograms, area under three dots corresponding to 3 consecutive dilutions showing saturation levels of hybridization were used and the copy number in the pearl millet genome was calculated considering the 1C nuclear DNA content of pearl millet as 2.5pg (Bennett and Smith 1976). Copy number ranged from 600 for PGB 727 to over 10,000 for PGB 625 (Table III) confirming repetitive nature of these sequences in pearl millet genome.

Genomic organization

Two approaches were followed to investigate the genomic organization of these repetitive sequences. In one set of experiments, the pearl millet DNA was subjected to time course digestions with the enzyme used for cloning and probed with the inserts from the corresponding enzyme-based library. Hybridization profiles were compared to determine the distribution of the sequence(s) in the genome. In another set of experiments, source DNA was restricted with 21 hexamer / tetramer recognizing enzymes including isoschizomers and Southern hybridized with labelled cloned sequences. The hybridization profiles of restricted fragments were scrutinized to deduce the configuration of repeats in each family of repetitive sequences.

Southern analysis following time course digestion of genomic DNA

Representative autoradiograms following Southern blot hybridizations of time course digested genomic DNA probed with different repetitive sequences are given in Fig. 2 a, b, d to g.

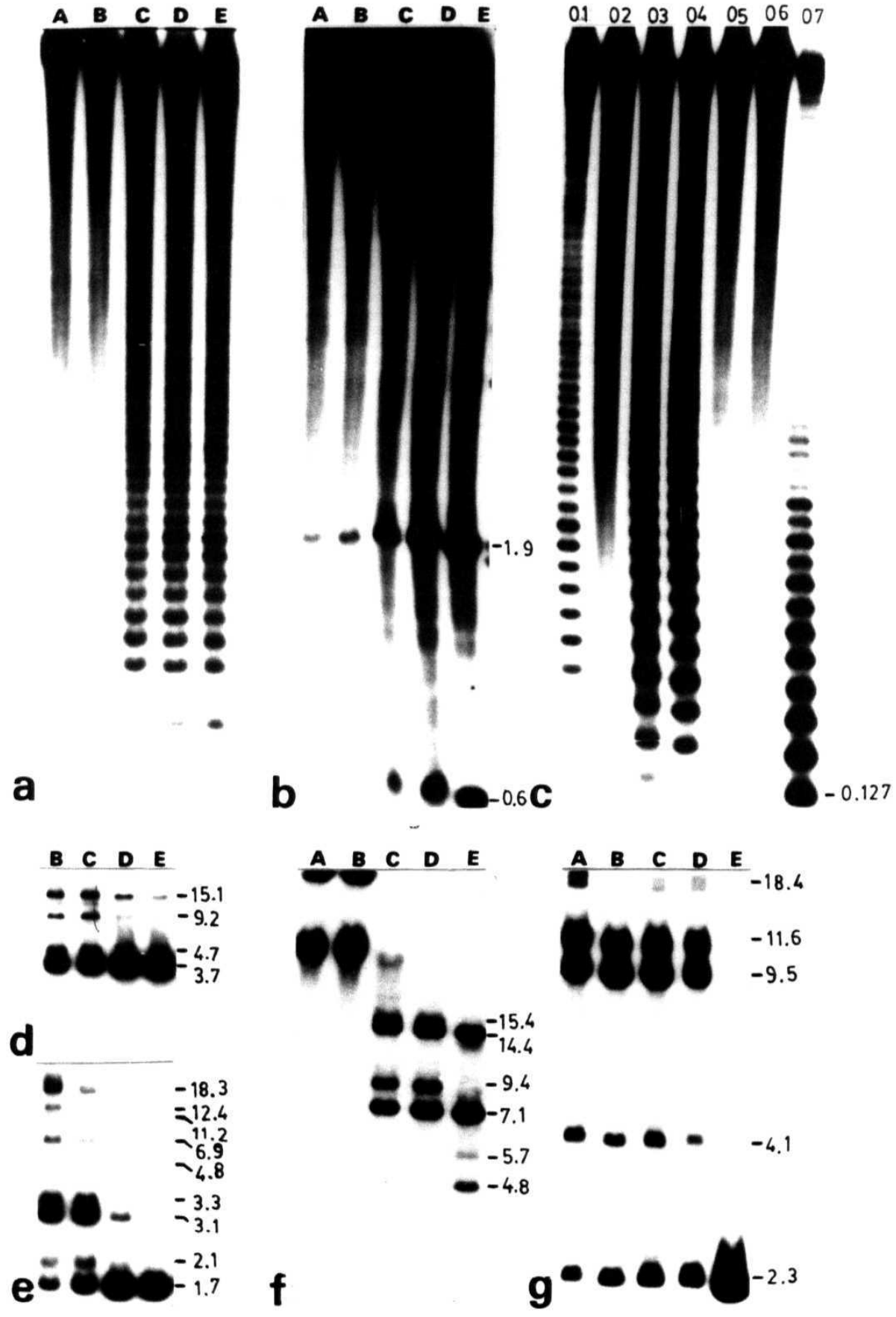
Table III. Copy number, proportion and organization of the cloned pearl millet repetitive DNA sequences in the genome

Clone	Insert size (Kb)	Copy number	Proportion (%)	Organization
PGP 005	3.1	1.9×10^3	0.23	Simple interspersed
PGE 015	3.7	2.0×10^3	0.29	Simple interspersed
PGE 096	9.1	2.0×10^3	0.72	Complex interspersed
PGE 123	6.8	0.6×10^3	0.17	Simple interspersed
PGB 058	2.4	1.5×10^3	0.15	Simple interspersed
PGB 074	1.7	1.6×10^3	0.11	Simple interspersed
PGB 103	1.0	2.3×10^3	0.09	Simple interspersed
PGB 107	0.8	3.8×10^3 *	0.12	Complex interspersed
PGB 582	2.3	1.0×10^3	0.10	Simple interspersed
PGB 625	1.4	1.0×10^4*	0.58	Tandem arrays
PGB 662	1.9	4.3×10^3 *	0.33	Complex interspersed
PGB 727	7.1	0.6×10^3	0.16	Simple interspersed
PGB 788	3.0	1.1×10^3	0.13	Simple interspersed

*Repeat unit length within these inserts are smaller and thus the copy number of repeats is much higher (see text)

Fig 2. Southern blots of time course digested pearl millet genomic **DNA** hybridized to ^{32}P labelled repetitive **DNA** sequences of ***P. g/aucum*** **: a -** PGB 625; **b -** PGB 662; *** c -** PGB 625; **d -** PGE 015; **e -** PGB 074; **f -** PGB 727 and **g -** PGB 582 **(a,b,e,f** & **g** Bam HI digested; **d** - Eco RI digested). Lanes represent DNA samples digested for 5 min (A), 10 min (B), 15 min (C), 20 min (D) and digested to completion (E).

*c) Southern blot hybrids of pearl millet genomic **DNA** following restriction with 01- Bam HI; 02 - Dpn I; 03 - Sau 3AI; 04 - Mbo I; 05 - Msp I; 06 - Hpa II and 07 - Hae **III** respectively, and probed with ^{32}P labelled PGB 625.


A B C D E
a
A B C D E
-1.9
-0.6
b
01 02 03 04 05 06 07
-0.127
c
B C D E
-15.1
-9.2
-4.7
-3.7
d
-18.3
-12.4
-11.2
-6.9
-4.8
-3.3
-3.1
-2.1
-1.7
e
A B C D E
-15.4
-14.4
-9.4
-7.1
-5.7
-4.8
f
A B C D E
-18.4
-11.6
-9.5
-4.1
-2.3
g

Time course digests of genomic DNA on probing with PGB 625 produced a smear of genomic DNA digested for 5 min. The smear progressively turned into a ladder of fragments as digestion time increased. No further change in the extent of restriction was found by increasing the incubation time beyond 20 min. A large proportion of genomic DNA homologous to PGB 625 was spared by Bam HI and appeared as smear from 50 to 5Kb and such DNA is called relic DNA (Bedbrook et al. 1980a).

Two sets of isoschizomers (Msp I & Hpa II and Mbo I, Sau 3AI & Dpn I) and Hae III were used in comparison with Bam HI to examine the methylation pattern of the sequences in the pearl millet genome hybridizing to PGB 625. Digestion with Msp I and Hpa II did not generate distinct bands with the exception of a few very weak bands in Msp I digest against a thick smear (Fig.2 c). Dpn I digest gave a smear without any distinct band whereas Sau3AI and Mbo I digests gave a strong ladder of fragments. Typical ladder was observed in Hae III digests and the striking feature of the ladder was that the signal intensity was not proportional to the size of the multimer. Hae III relic DNA homologous to PGB 625 was least among all the restriction enzymes used From sizes of the fragments in the ladder, an internal progression of a 127 bp repeat length was evident (Fig. 2 c).

Southern blots of pearl millet genomic DNA subjected to increasing periods of digestion with Bam HI was probed with PGB 662. Two intense bands of sizes 0.6 and 1.9 Kb were found against a background smear (Fig. 2 c). Intensities of these two bands increased with the increase in digestion time. Furthermore, smear intensity also increased. Southern blots of time course digested genomic DNA probed with PGB 107 also resulted in a pattern similar to PGB 662. Complete digestion resulted in 10 distinct bands of varying intensities. The sizes of the fragments ranged from 0.5 to 4Kb with low background smear (data not shown) in comparison with that of PGB 662, which is consistent with the low copy number of PGB 107. Other than a thick smear from top to

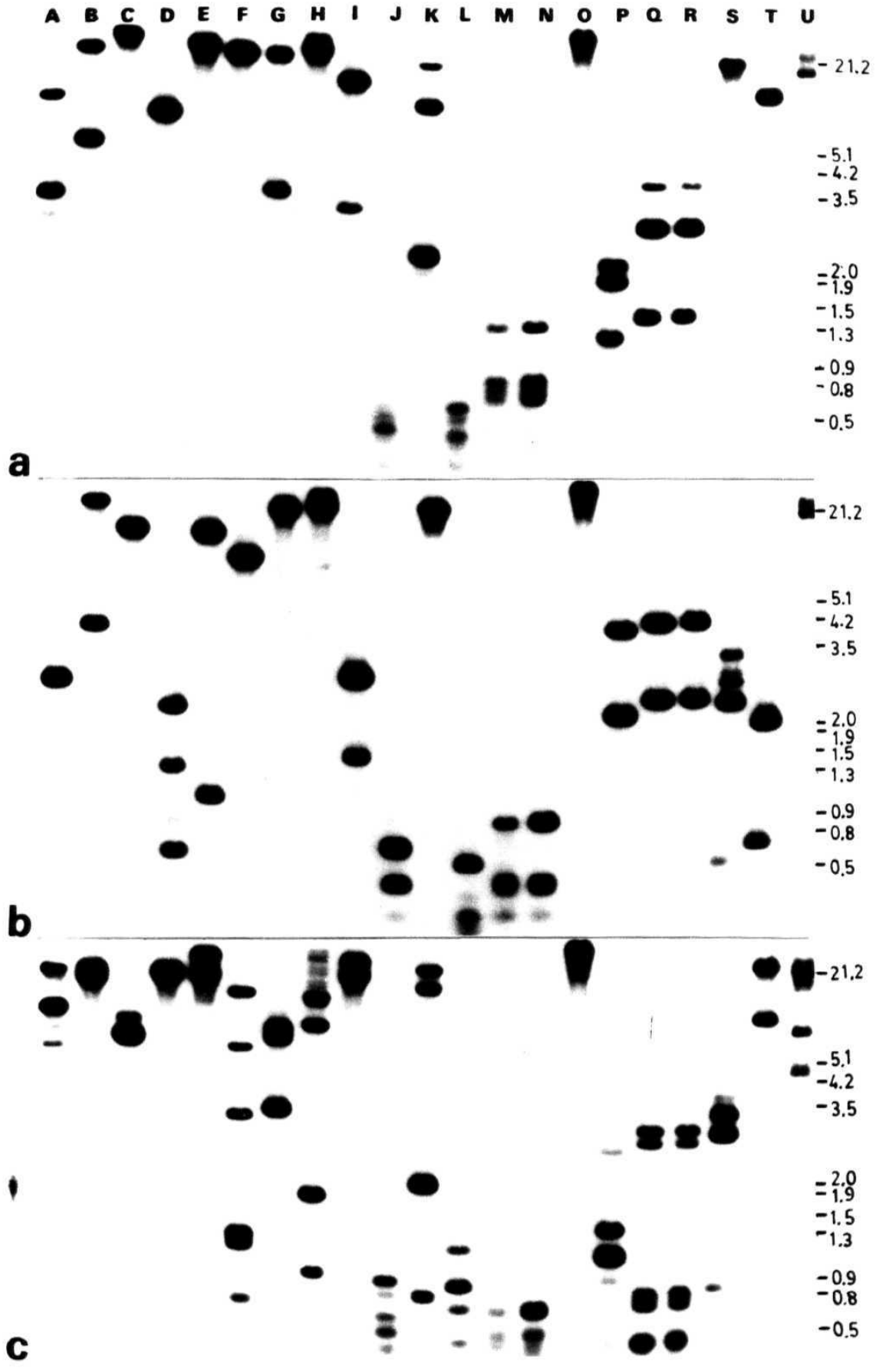
bottom in the lane, no distinct band was detected when Eco RI digested DNA was probed with PGE 096 (data not shown).

Southern blots prepared from pearl millet genomic DNA following time course digestion displayed bands of varying sizes hybridizing to PGE 015, PGB 074, PGB 727 and PGB 582 (Fig.2d-g). Four fragments (15.1, 9.2, 4.7 and 3.7 Kb) were found when genomic DNA was partially digested with Eco RI and probed with PGE 015 (Fig.2 d). Complete digestion with Eco RI (Fig. 2 d) resulted in single 3.7Kb fragment. Partial digest with Bam HI gave nine fragments ranging from 18.3Kb to 1.7Kb hybridizing to PGB 074 (Fig. 2 e) while complete digestion resulted in 1.7Kb fragment only. Bam HI digests hybridizing to PGB 582 showed five fragments ranging from 18.4 to 2.3 Kb (Fig. 2 g) in partial digests while complete digests gave single 2.3Kb fragment only. PGB 727 hybridization pattern with partial digests gave high molecular size fragments while complete digestion resulted in four bands (14.4, 7.1,5.7 and 4.8Kb) homologous to the PGB 727 (Fig. 2 f).

Restriction analysis of genomic DNA

Southern blots of pearl millet genomic DNA following digestions with different enzymes were probed with PGE 123 (Fig 3a), PGB 582 (Fig. 3 b) and PGB *727* (Fig. 3 c) to examine the organization of the basic repeat unit. Genomic DNA restricted with Bgl II, Eco RV, Hind III, Sac I, Xho I and Sal I resulted in single bands hybridizing to PGE 123 (Fig. 3 a lanes C, E, F, H, S & T respectively) differing in sizes while Eco RI fragment (lane D) was of the same size as the insert of the probe. More than one fragment was found in Kpn I, Sau3AI, Mbo I, Hae III, Msp I and Hpa II digests (lanes K, M,N, P, Q & R respectively). Hinf I and Taq I generated many smaller fragments (lanes J & L) resulting in diffused appearance making it difficult to discern their precise number. Restriction profiles revealed by PGB 582 probe (Fig.3 b) consisted of one

Fig 3. Southern blot of pearl millet genomic DNA digested with different restriction enzymes (lanes A-U) and hybridized to pearl millet repetitive DNA sequence probes : a - PGB 123; b - PGB 582; & c - PGB 727. Enzymes used are A - Bam HI; B - Bgl I ; C - Bgl II ; D - Eco RI; E - Eco RV; F -Hind III; G - Pst I; H - Sac I; I - Xba I; J - Hinf I; K - Kpn I; L - Taq I; M -Sau 3AI; N - Mbo I; 0 - Dpn I; P - Hae III; Q - Msp I; R - Hpa II; S - Xho I; T - Sal I & U - Sma I.


A B C D E F G H I J K L M N O P Q R S T U
21.2
5.1
4.2
3.5
2.0
1.9
1.5
1.3
0.9
0.8
0.5
a
21.2
5.1
4.2
3.5
2.0
1.9
1.5
1.3
0.9
0.8
0.5
b
21.2
5.1
4.2
3.5
2.0
1.9
1.5
1.3
0.9
0.8
0.5
c

band each in Bam HI, Pst I and Sac I (lanes A,G & H respectively). Other enzyme digests showed varying numbers of fragments. Hybridization with PGB 727 resulted in variable number of bands in all the enzyme digests. One striking feature of the restriction profiles revealed by the probes hybridizing to more than one band was that the signal intensities of some of the bands in a given profile were not proportional to the fragment size(s).

A 20-fold variation in the copy number among different repetitive DNA sequences was evident (Table III). Sequences homologous to the 12 families of the repetitive sequences together constitute ~ 3.2% of the pearl millet genome.

Distribution of the sequences in the genome

Ladder of fragments generated by time course digestion of pearl millet genomic DNA hybridizing to PGB 625 represent a typical tandem repeat organization. The basic repeat length estimate of PGB 625 family of tandem repeats is 127bp. The repeat unit length of Kpn I family of tandem repeats reported by Ingham et al. (1993) from pearl millet is 140 bp while the Hae III family of repeat reported by Kamm et al. (1994) is 137bp but it has homology to the Kpn I family. Southern blot hybridization pattern of the Bam HI, Hae III, Sau 3AI, Msp I and Hpa II digests of pearl millet genomic DNA to the Hae III monomer reported by Kamm et al. (1994) and to the Bam HI bordered 1.4Kb insert (PGB 625) in the present study are similar and these two repeats may belong to the same family. Although the PGB 625 insert (1.4Kb) copy number estimate was 10^4 (Table III), in view of the ladder of 127bp tandem repeats, the copy number of this repeat is > 10^5. The Bam HI relic DNA homologous to PGB 625 indicates that members of PGB 625 family in pearl millet genome are arranged as long tandem arrays wherein some of the Bam HI recognition sites are modified giving rise to such a pattern. Disproportionality of signal intensities and fragment sizes may indicate random periodicity of unaltered enzyme sites in these tandem arrays. A comparison of the

fragments generated by Bam HI to those of Hae III revealed the disproportionate signal intensities among the ladder of fragments hybridizing to the same probe (PGB 625). Data on DNA sequence of monomers of tandem repeats in several plant species (Martinez-Zaparter et al. 1986; Grellet et al. 1986; Appels et al. 1986; Wu and Wu 1992; Kamm et al. 1995; Li et al. 1996) showed that different copies of the repeating units are not completely identical, but exhibit very high sequence similarity. Such differences among the repeating units arise from base substitution, deletions and insertion during evolution (Grellet et al. 1986; Kamm et al. 1995).

The heterogeneity of the fragments homologous to PGB 107 and PGB 662 is indicative of a complex distribution at diverse sequence environments in the genome. Relic DNA may represent long stretches resulting from altered recognition sites due to methylation or mutation. Smear observed in the present study may reflect interspersion of repetitive sequences among low copy sequences as suggested in other reports (Flavell 1980; Bouchard 1982; Evans et al. 1983). At some genome positions, sequences are arranged in clusters, and discrete bands stem from such clusters. The size of the basic repeat unit of PGB 662 families is ~0.6Kb and the cloned fragment in PBG 662 is likely to be a trimer of the repeat. The basic repeat of PGB 107 family is ~0.5Kb. Thus the copy numbers of the basic repeats of PGB 662 and PGB 107 can now be considered as 1.3×10^4 and 6×10^3 respectively. Data on the genome organization of PGE 096 is inadequate to draw any meaningful conclusion.

Restriction profiles of partially digested genomic DNA homologous to the probes PGE 015, PGB 074 and PGB 582 consisted of many bands and one of them had a size equal to the insert size of the probe used. The signal intensity of the fragment (equal to the insert size) increased with a concomitant decrease and final disappearance of large fragments on complete digestion implying that during the course of digestion, basic repeats belonging to each repetitive sequence family were being successively excised from large stretches thereby increasing the intensity of the band matching with the

size of the cloned fragment. The absence of ladder of fragments in partial digests of genomic DNA homologous to these probes indicate that the basic repeats are not arranged tandemly. Lack of any relationship between the sizes of different fragments and their hybridization signal intensities indicate that they are excised from different genomic locations. Results suggest that inserts of PGE 015, PGB 074 and PGB 582 are mostly arranged in clusters of 2- to 10- mers and such clusters are interspersed at different genomic locations. Differences in the signal intensities of larger fragments reflect the relative proportions of different clusters in the genome.

A 2.0Kb Pst I - Bam HI bordered fragment is common to the Pst I insert of PGP 005 and Bam HI insert in PGB 727 (data not shown). Hybridization of the 7.1Kb and 14.4 Kb Bam HI fragment of genomic DNA to PGP 005 and PGB 727 substantiate the foregoing. The additional 1.1 Kb fragment homologous to PGP 005 and the presence of 4.8Kb and 5.7Kb fragment homologous to PGB 727 are suggestive of their distribution in the genome in different configurations.

In the light of the above, it can be inferred that these 13 sequences fall into two categories, tandem repeat sequences and interspersed sequences. PGB 625 is the only tandem repeat family among the cloned repetitive DNA sequences, PGE 096, PGB 107 and PGB 662 have complex distributions in the genome. Henceforth, they are referred to as complex interspersed sequence and the remaining nine interspersed sequences as simple interspersed sequences (Table III).

Configuration of repeat units

The presence of a single fragment (in an enzyme digest) longer than the insert size of the probe indicates that the enzyme does not have sites within the basic repeat defined by the enzyme used for cloning and it excised monomers or cluster of basic repeat units with or without flanking regions. Such enzymes do not reveal the structure of the basic repeat unit of a family of repetitive sequences. Some of the enzymes, especially tetranucleotide sequence recognizing enzymes, generated more than one

fragment with sizes smaller than the insert sizes indicating that these enzymes have sites within the basic repeat unit defined by the enzyme used for cloning. In many instances, the signal intensities of the fragments observed in such profiles were not proportional to their fragment sizes. This shows that upon digestion with the corresponding enzyme, the basic repeat unit gave rise to a set of fragments and the relative proportion of such fragments were not same. For example, in Msp I or Hpa II digest one would expect the top 3.6Kb fragment hybridizing to PGE 123 to be more intense than the second 2.6Kb fragment if the fragment sizes are taken into consideration. Signal intensity in 2.6Kb band was 5- to 6- fold higher compared to that of 3.6 Kb band suggesting that the 6.3Kb repeating motif of PGE 123 has complex organization and each basic unit is made up of different repeats and the relative proportion of subrepeats vary. Similar conclusion can be drawn from the restriction pattern discernible from hybridization with PGB 582 and PGB 727. Subcloning and sequencing of individual clones are required to resolve their nature.

Methylation status of tandem repeats

Recognition sequence of Hae III is 5'GGCC 3' and the enzyme does not cleave if the internal 'C' is methylated or both the Cs are hemi-methylated. Similarly, restriction site of Msp I or Hpa II is '5 CCGG 3' and Msp I is sensitive to methylation of internal 'C' or hemi-methylation of 'CC' while Hpa II is sensitive to the methylation of either of the 'Cs'. Absence of discrete low molecular size fragments hybridizing to PGB 625 probe following restriction with Msp I and Hpa II may be a reflection of the high levels of methylation at '5 CCGG 3' sites or absence of such sites in the tandem arrays revealed by Hae III digests. Absence of any 5' CCGG 3' sequence in the Hae III 137bp repeat of pearl millet as evident from the sequence data published by Kamm et al. (1994) further substantiates earlier conclusion (page - 32) that the tandem array of (127bp) repeats (PGB 625) in the present study is similar to this 137bp sequence.

Recognition sequence of isoschizomers Mbo I, Sau 3AI and Dpn I is GATC. Adenine methylation is a prerequisite for Dpn I cleavage and is insensitive to 'C methylation. Mbo I is sensitive to adenine methylation and insensitive to cytosine methylation whereas Sau 3AI is sensitive to cytosine methylation and insensitive to adenine methylation. The smear observed in Dpn I digest thus implies no or low adenine methylation in GATC sites in the sequences homologous to PGB 625. Considering the methylation sensitivity, proportion of relic DNA observed in Sau 3AI or Mbo I digests may reflect changes in their recognition sites as suggested earlier(Grellet et al. 1986; Kamm et al. 1995).

3.2 MOLECULAR ANALYSIS OF *PENNISETUM* SPECIES AND PEARL MILLET ACCESSIONS WITH CLONED PEARL MILLET REPETITIVE DNA SEQUENCES

Molecular analysis of *Pennisetum* species differing in their basic chromosome number (P. *ramosum* (x=5), *P. mezianum* (x=8), P. *orientate* (x=9) and an allotetraploid P. *purpureum* (x=7) with one genome common to P. *glaucum*) was carried out with cloned pearl millet repetitive DNA sequences.

Distribution and modulation:

Aliquots of equimolar concentrations of recombinant plasmids carrying different pearl millet repetitive DNA sequence inserts were restricted (for details see section - 2.3.1) and fractionated on agarose gels. Southern blot was hybridized with ^{32}P-labelled pUC 18 DNA to ensure uniform quantity of the vector (pUC 18) among the aliquots (Fig. 4 C). Autoradiograms of the blot following hybridization one after the other with labelled total genomic DNA of *Pennisetum* species (Fig. 4 r, m, o, p & g) were used for analysis. Signal intensity of each cloned fragment resulting from hybridization with genomic DNA from each species as a probe was estimated using UVP gel documentation system. For comparison of signal intensities between species, the ratio of signal intensity of each insert to that of PGP 005 was calculated (Table - IV).

Relative hybridization signal intensities of pearl millet repetitive DNA sequence inserts determined by probing with P. *glaucum* genomic DNA was compared with those of other *Pennisetum* species. Intensities of PGE 123 (0.22 - 0.33), PGB 058 (0.68-0.82), PGB 074 (0.62-1.05), PGB 103 (0.87-1.53), PGB 107 (1.28 - 2.01), PGB 727 (0.24 - 0.37) and PGB 788 (0.48 - 0.74) were of comparable levels across the species (Table IV). Variation for PGE 015 in P. *orientale* was 2.5- fold less as compared to that

Fig 4. Southern blot of restricted repetitive DNA sequence clones hybridized to: **c** - labelled vector DNA and sheared ^{32}P- labelled genomic **DNA(s)** from ***Pennisetum*** species: **g** - *P.* ***glaucum; x -P. ramosum;*** **m -** ***P. mezianum;*** **o -** ***P. orientale;*** **p** - *P.* ***purpureum.*** The clone numbers given at the top of each lane with a prefix P, E or B represent Pst I-, Eco RI- or Bam HI-based pearl millet genomic libraries respectively and the corresponding enzyme used for the digestion of the recombinant clones

-P 005
-E 015
-E 096
-E 123
-B 058
-B 074
-B 103
-B 107
-B 582
-B 727
-B 788
-B 625
-B 662
c
g
r
m
o
p

Table IV: Relative abundance of sequences in *Pennisetum* species homologous to cloned pearl millet repetitive DNA

Clones *	*P.ramosum* (x=5)	*P.mezianum* (x=8)	*P.orientale* (x=9)	*P.purpureum* (x=7)	*P.glaucum* (x=7)
PGP 005	1	1	1	1	1
PGE 015	0.71	0.50	0.39	0.62	1.05
PGE 096	0.60	-	-	-	1.04
PGE 123	0.22	0.26	0.22	0.25	0.33*
PGB 058	0.79	0.73	0.68	0.74	0.82*
PGB 074	0.62	0.83	1.05	1.03	0.87*
PGB 103	0.87	1.34	1.12	1.53	1.21*
PGB 107	1.44	1.21	1.91	1.28	2.01*
PGB 582	0.51	0.72	0.92	0.69	0.56*
PGB 727	0.25	0.37	0.24	0.33	0.29*
PGB 788	0.48	0.74	0.63	0.61	0.58*
PGB 625	-	0.54	0.96	4.45	5.51
PGB 662	0.94	0.43	0.5	0.77	2.29

Relative abundance of different repetitive sequences in relation to the copy number of PGP 005

of P. *glaucum*. Ratios observed for PGE 096, PGB 625 and PGB 662 varied markedly between species. Compared to P. *glaucum,* relative intensity of PGE 096 was nearly half in *P. ramosum* and no signal with other species **DNA.** Relative abundance of PGB 625 insert was similar in *P. purpureum* and *P. glaucum,* was 5- and 10- fold less in *P. orientate* and P. *mezianum* respectively, and was not detectable by the genomic DNA probe of the *P. ramosum.* Relative intensity of PGB 662 was over 2- fold (in P. *ramosum)* to 5 fold less (in P. *mezianum).*

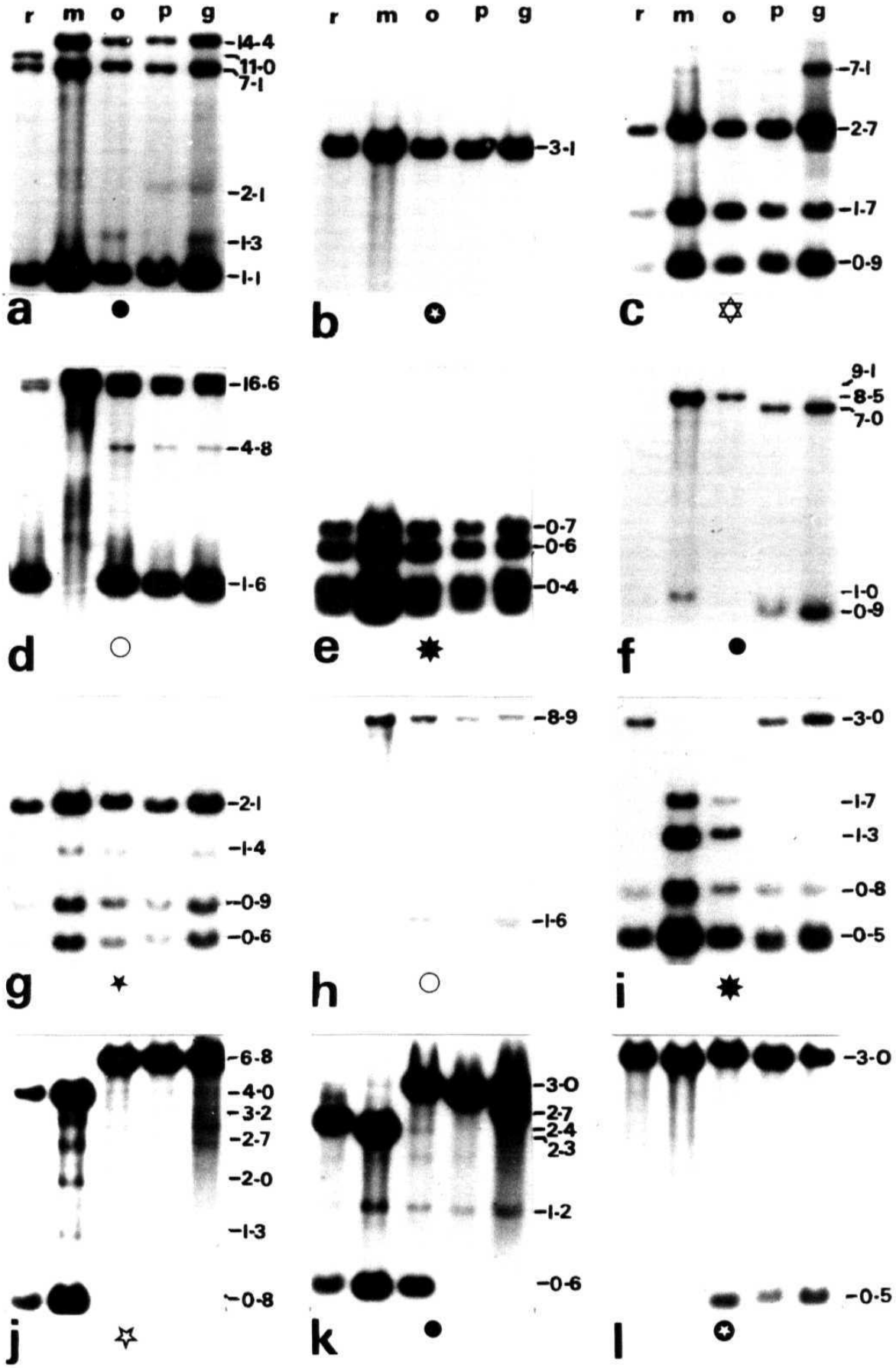
Restriction analysis

Eco RI, Pst I, Bam HI, Eco RV, Sac I, Bgl I, Hind III and Msp I were used for restriction analysis. Southern blots of restriction fragments of genomic DNAs of the five *Pennisetum* species were probed with cloned pearl millet repetitive sequences. PGE 096 did not reveal discrete bands on hybridization with P. *glaucum* genomic DNA and was excluded in further analysis.

Restriction profiles of genomic DNAs and their hybridization patterns with each pearl millet repetitive **DNA** clone are described below

PGE 005:- Digestion with Bam HI (Fig. 5 a) and Bgl I (Fig. 5 d) distinguished P. *ramosum* and P. *mezianum* respectively from the rest of the species. A 9.2 Kb Bam HI fragment exclusive to P. *ramosum* and the absence of a 1.6 Kb Bgl I fragment in P. *mezianum* are the distinguishing features of the sequences homologous to PGP 005. F The few faint bands observed in Bam HI (Fig. 5 a) digests were common to two or three species. A 4.8 Kb Bgl I fragment was found in P. *orientate, P.purpureum* and P. *glaucum.* Two faint bands (2.4 and 1.3 Kb) specific to P. *glaucum* **DNA** were observed in Hind III digest (data not shown). In addition to the four conserved bands (7.1, 2.7, 1.7 and 0.9 Kb), one faint band each of size 3.3 Kb was found in Sac I digests of P *orientate* and P. *glaucum* (Fig. 5 i - not marked). PGP 005 probe did not reveal polymorphism between species in other enzyme digests (Fig. 5 b & e).

Fig 5. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; **o** - *P.* ***orientale***; **p** - *P.* ***purpureum***; **g** - *P.* ***glaucum*** hybridized to pearl millet repetitive DNA sequence probes - PGB 005 (**a - e**); PGE 015 (**f - i**); PGE 123 (**j - l**) and the enzymes used are - Bam HI (●), Bgl I (○), Eco RI (✡), Eco RV (✦), Msp I (✸), Pst I (✪) or Sac I (✡).


r m o p g
14·4
11·0
7·1
2·1
1·3
1·1
a
r m o p g
3·1
b
r m o p g
7·1
2·7
1·7
0·9
c
16·6
4·8
1·6
d
0·7
0·6
0·4
e
9·1
8·5
7·0
1·0
0·9
f
8·9
1·6
h
3·0
1·7
1·3
0·8
0·5
i
2·1
1·4
0·9
0·6
g
6·8
4·0
3·2
2·7
2·0
1·3
0·8
j
3·0
2·7
2·4
2·3
1·2
0·6
k
3·0
0·5
l

PGE 015:- Restriction profile generated by Bam HI (Fig. 5 f) consisted of two bands each in all the species. *P. purpureum* and *P. glaucum* had identical profiles consisting of 7.0 and 0.9 Kb bands. Similarly the profiles comprising the 8.5 Kb and 1.0 Kb bands were common to *P. mezianum* and *P. orientate.* The distinct 1.0 Kb band in *P. mezianum* is very faint in *P. ramosum* and *P. orientate.* A 9.1 Kb band specific to P. *ramosum* is also hazy. Restriction profile revealed by Bgl I (Fig. 5 h) contained a 8.9 Kb band in all the species and one 1.6 Kb band in other species except *P. mezianum.* The 0.8 and 0.5Kb bands revealed by Msp I (Fig 5. i) were found in all the species. In addition, a 3.0 Kb band was present in *P. ramosum, P. purpureum* and *P. glaucum* but two fragments (1.7 Kb and 1.3 Kb) were present in *P. mezianum* and *P. orientale*. Digestion with Eco RI (data not shown) resulted in a 3.6 Kb band common to P. *ramosum* and P. *mezianum* and a 3.7 Kb band common to other species. In addition, one faint band each of sizes 0.4 and 2.0 Kb were observed in P. *mezianum* and P. *orientate* respectively. Digestion with Pst I (data not shown) resulted in two bands each in P. *ramosum, P. mezianum* and P. *orientate* and three bands each in P. *purpureum* and P. *glaucum.* One 3.1 Kb band was common to all the species. Two other bands *{7.7* and 2.4 Kb) were common to P. *purpureum* and P. *glaucum* whereas the size of the second band was characteristic to P. *ramosum* (8.9 Kb), P. *mezianum* (10.7 Kb) and P. *orientate* (8.6 Kb). Eco RV, Sac I and Hind III did not detect variation in the sequences related to PGE 015 among the species and a typical autoradiogram is given in Fig. 5.g. However, very faint 0.9Kb Sac I fragment and 3.0 Kb Hind III fragments were specific to P. *mezianum* and P. *orientate* respectively (data not shown).

PGE 123:- Eco Rl generated (Fig. 5 J) a 6.8 Kb band each in P. *orientale*, *P. purpureum* and P. *glaucum* and two bands each of sizes 4.0 Kb and 0.8 Kb in P. *ramosum* and P. *mezianum.* Varying number of faint bands were observed in all the species. Bam HI restriction profile (Fig. 5 k) consisted of a faint 1.2 Kb band common to all the species, a 0.6 Kb band common to P. *ramosum, P. mezianum* and P. *orientate,*

and a 3.0 Kb fragment common to *P. purpureum* and P. *glaucum.* One band of size 2.4, 2.3, 3.3 and 2.7 Kb were specific to *P. ramosum*, *P. mezianum, P. orientate* and *P. glaucum* respectively.

Restriction profiles of Pst I (Fig. 5 I) comprised of a 3.0 Kb band each in all the species and a 5.0 Kb band in *P. orientale, P. purpureum* and *P. glaucum.* Two Sac I fragments (Fig. 6 a) (16.7 and 5.7 Kb) in P. *ramosum* and *P. mezianum,* and a 25.9 Kb band common to *P. orientate, P. purpureum* and *P. glaucum* were found. The profiles of Bgl I digests (Fig. 6 b) consisted of a 23.2 Kb band in. *P. ramosum* and *P. mezianum* and two bands each of sizes 20.8 and 4.6 Kb in the remaining three species. Msp I generated (Fig. 6 c) a 2.1 Kb high intensity band common to *P. ramosum* and *P. mezianum,* a 0.9 Kb band common in *P. ramosum, P. mezianum* and *P. orientale* and four bands (3.8, 2.9, 1.4 and 0.6 Kb) common to *P. orientale, P. purpureum* and P. *glaucum.* One band each of sizes 9.9 Kb and 20.8 Kb was found in all species in Eco RV and Hind Hi digests respectively (data not shown).

PGB 058:- Restriction profiles following Eco Rl digestion (data not shown) comprised a 6.4 Kb band in P. *ramosum* and P. *mezianum* and a 5.9 Kb band common to the remaining species. A very faint 6.1 Kb Sac I band specific to P. *orientale* was detected (data not shown). Other enzymes used did not reveal variation among the sequences related to PGB 058 (Table V) and a representative autoradiogram is given in Fig. 6 d.

PGB 074:- A (1.6 Kb) Bam Hl fragment (Fig. 6 e) common to P. *ramosum* and P. *mezianum* and another (1.7 Kb) in the remaining species were found. In addition, a distinct 3.1 Kb band was present in P. *glaucum* which was also detectable in P. *mezianum, P. orientale* and P. *purpureum.* Digestion with Msp I (Fig. 6 f) resulted in a conspicuous 0.9 Kb band and a 0.3 Kb band in all the species but a faint 1.9 Kb band each in P. *ramosum, P. mezianum* and a faint 2.8 Kb band each in rest of the species.

Fig 6. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; **o** - *P.* ***orientale***; **p** - *P.* ***purpureum***; **g** - *P.* ***glaucum*** hybridized to pearl millet repetitive DNA sequence probes - PGE 123 (**a - c**); PGB 058 (**d**); PGB 074 (**e & f**); PGB 103 (**g & h**); PGB 107 (**I & j)** PGB 582 (**k & l)** and the enzymes used are : Bam HI (●), Bgl I (○), Eco RI (✡), Hind III (✡), Msp I (✸) or Sac I (✡).

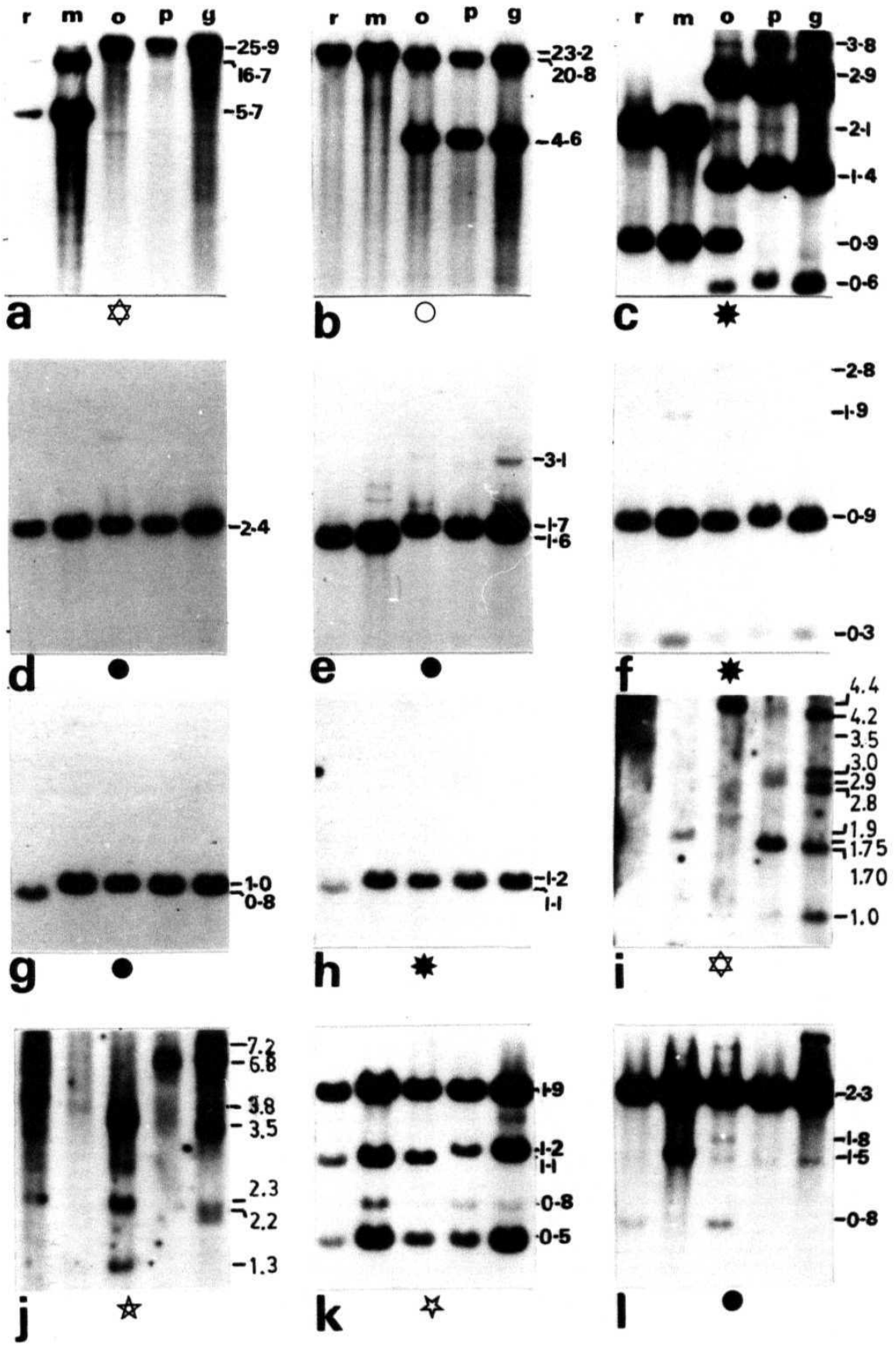

r m o p g
25·9
16·7
5·7
a
r m o p g
23·2
20·8
4·6
b
r m o p g
3·8
2·9
2·1
1·4
0·9
0·6
c
2·4
d
3·1
1·7
1·6
e
2·8
1·9
0·9
0·3
f
1·0
0·8
g
1·2
1·1
h
4.4
4.2
3.5
3.0
2.9
2.8
1.9
1.75
1.70
1.0
i
7.2
6.8
3.8
3.5
2.3
2.2
1·3
j
1·9
1·2
1·1
0·8
0·5
k
2·3
1·8
1·5
0·8
l

Table V: Probe - enzyme combinations revealing sequence diversity of different repetitive DNA sequences in *Pennisetum* species

Probe	Enzymes								
	Eco RI GAATTC	Pst I CTGCAG	Bam HI GGATCC	Eco RV GATATC	Sac I GAGCTC	Bgl I $GCCN_4NGGC$	Hind III AAGCTT	Msp I CCGG	Total
PGP 005	-	-	+	-	-	+	F	-	2
PGE 015	+	+	+	-	F	+	F	+	5
PGE 123	+	+	+	-	+	+	-	+	6
PGB 058	+	-	-	-	F	-	-	-	1
PGB 074	+	-	+	F	-	-	-	+	3
PGB 103	+	-	+	-	-	-	+	+	4
PGB 582	+	-	F	-	F	+	F	+	3
PGB 727	-	-	+	-	-	+	F	+	3
PGB 788	+	-	+	-	-	-	F	+	3
Total	7	2	7	-	1	5	1	7	30

Very faint bands specific to a species or a species sub- set is represented by 'F'

Eco RI digests (data not shown) revealed a 7.3 Kb band common to *P. ramosum* and P. *mezianum* and a 6.7Kb band common to *P. orientale*, *P. purpureum* and P. *glaucum.* Digestion with Eco RV (data not shown) showed a very faint 3.1 Kb band common to P. *mezianum, P. orientale* and P. *glaucum* in addition to a strong 10.1 Kb band common to all the species. Pst I, Sac I, Bgl I and Hind III did not detect polymorphism among the different species for the sequences homologous to PGB 074 (Table V).

PGB 103:- Bam HI generated a 0.8 Kb fragment in P. *ramosum* and a 1.0 Kb fragment in the remaining species (Fig. 6 g). Similarly Msp I digestion (Fig. 6 h) resulted in a 1.1 Kb band characteristic of P. *ramosum* and 1.2 Kb band common to the remaining species. Eco RI digest (data not shown) consisted of three fragments each in all the species, of which a 0.3 Kb fragment was common to all the species. The other two bands (1.7 and 0.6 Kb) were common to P. *mezianum,* P. *orientale* and P. *purpureum* while P. *ramosum* had 1.6 and 0.5 Kb fragments. In addition, a very faint 0.8 Kb band was observed in P. *mezianum.* Hind III restriction pattern distinguished P. *ramosum* from the rest of the species by a characteristic 3.2 Kb band in P. *ramosum* and 2.0 and 1.0 Kb bands in the remaining species (data not shown). Digestion with Pst I, Eco RV, Sac I and Bgl I did not detect variations among the sequences related to PGB 103 in the *Pennisetum* species studied (Table V).

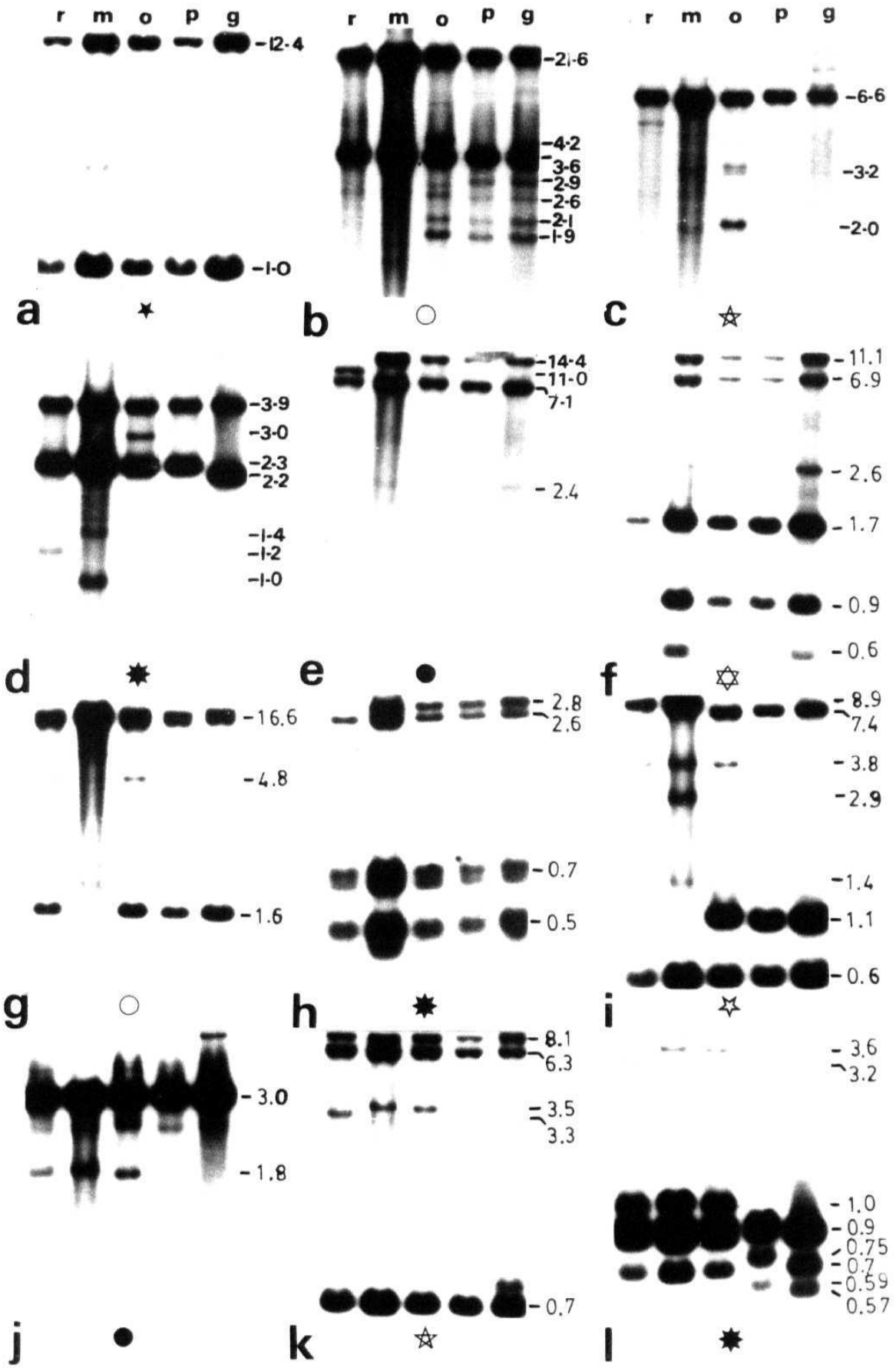
PGB 107:- Bam HI, Sac I and Bgl I restrictions resulted in discrete bands against a faint smear and the restriction profile of DNA hybridizing to PGB 107 were specific to each species. Sac I digests (Fig. 6 i) comprised of two bands each in P. *ramosum* and P. *mezianum,* one in P. *orientale,* three in P. *purpureum* and five in P. *glaucum.* Similarly varying numbers of bands were found among these species following Hind III digestions (Fig. 6 J). Except for a smear, digestion with Eco RI, Pst I, Eco RV and Msp I did not yield any discrete bands (Table V).

PGB 582:- Eco RI restriction profiles (Fig. 6 k) consisted of two prominent bands (1.9 Kb and 0.5 Kb) in all the species. In addition, P. *ramosum*, *P. mezianum* and *P. orientate* had a 1.1 Kb band while *P. purpureum* and *P. glaucum* showed a 1.2 Kb band. A relatively faint 0.8 Kb band was present in *P. mezianum, P. purpureum* and P. *glaucum.* Bam HI generated (Fig. 6 l) a strong 2.3 Kb band and a weak 1.5 Kb band in all the species, in addition to a faint 0.8 Kb fragment in P. *ramosum* and P. *orientale* and another 1.8 Kb band specific to P. *orientale*. Digestion with Eco RV and Pst I did not reveal variation in the sequences homologous to PGB 582 (Table V) and a typical conserved pattern revealed by Eco RV is given in Fig. 7.a.

Restriction profiles generated by Bgl I (Fig. 7 b) consisted of two intense bands each of sizes 21.6 and 3.6 Kb in all the species. Four relatively less intense bands (2.9, 2.6, 2.1 and 1.9 Kb) were common to P. *orientale, P. purpureum* and P. *glaucum* in addition to a 4.2 Kb band specific to P. *orientale.* Hind III generated (Fig. 7 c) a 6.6 Kb fragment in all the species and two bands (3.2 and 2.0 Kb) specific to P. *orientale.* Msp I (Fig. 7 d) gave species-specific fragmentation pattern which consisted of a common 3.9 Kb fragment in all the species and a 2.3 Kb in four species except P. *glaucum* which had a characteristic 2.2 Kb band. Species-specific fragments were present in P. *ramosum* (1.2 Kb), P. *mezianum* (1.4 and 1.0 Kb) and P. *orientale* (3.0 Kb). Sac I generated one intense 23.4 Kb band each in all the species and a faint 0.8 Kb fragment specific to P. *mezianum* (data not shown).

PGB 727:- Restriction with Bam HI resulted in three bands in each species(Fig. 7 e), of which the intense 7.1 Kb and a faint 2.4 Kb bands were present in all the species, and a intense 10.7 Kb band was specific to P. *ramosum,* while the remaining four species displayed a 14.4 Kb fragment. Bgl I generated 16.6 Kb fragment common to all the species and two bands of sizes 4.8 and 1.6 Kb in four species except P. *mezianum* (Fig. 7 g). Msp I generated two doublets of ~0.7 and 0.5 Kb and a 2.6 Kb band in all the species and a 2.8 Kb band in four species except in P. *ramosum* (Fig. 7 h). In

Fig 7. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r** - *P.* ***ramosum***; **m** - *P.* ***mezianum***; **o** - *P.* ***orientale***; **p** - *P.* ***purpureum***; **g** - ***P. glaucum*** hybridized to pearl millet repetitive DNA sequence probes - PGB 582 (**a - d**); PGB 727 (**e - h**); PGB 788 (**i - l**) the enzymes used are - Bam HI (●), Bgl I(○), Eco RI (✰), Eco RV (✶), Hid III (✡), Msp I(✸) or Sac I (✡).


r m o p g
12·4
1·0
a
2i·6
4·2
3·6
2·9
2·6
2·1
1·9
b
6·6
3·2
2·0
c
3·9
3·0
2·3
2·2
1·4
1·2
1·0
d
14·4
11·0
7·1
2·4
e
11.1
6.9
2.6
1.7
0.9
0.6
f
16.6
4.8
1.6
2.8
2.6
0.7
0.5
8.9
7.4
3.8
2.9
1.4
1.1
0.6
g
3.0
1.8
h
8.1
6.3
3.5
3.3
0.7
i
3.6
3.2
1.0
0.9
0.75
0.7
0.59
0.57
j
k
l

addition to the 11.9, 4.8, 2.9, 1.1, 0.6 and 0.3 Kb bands common to all the species, Hind III generated three faint bands specific to P. *glaucum* (data not shown). Eco R I, Eco RV, Pst I and Sac I did not reveal any variation among the sequences related to PGB 727 (Table V).

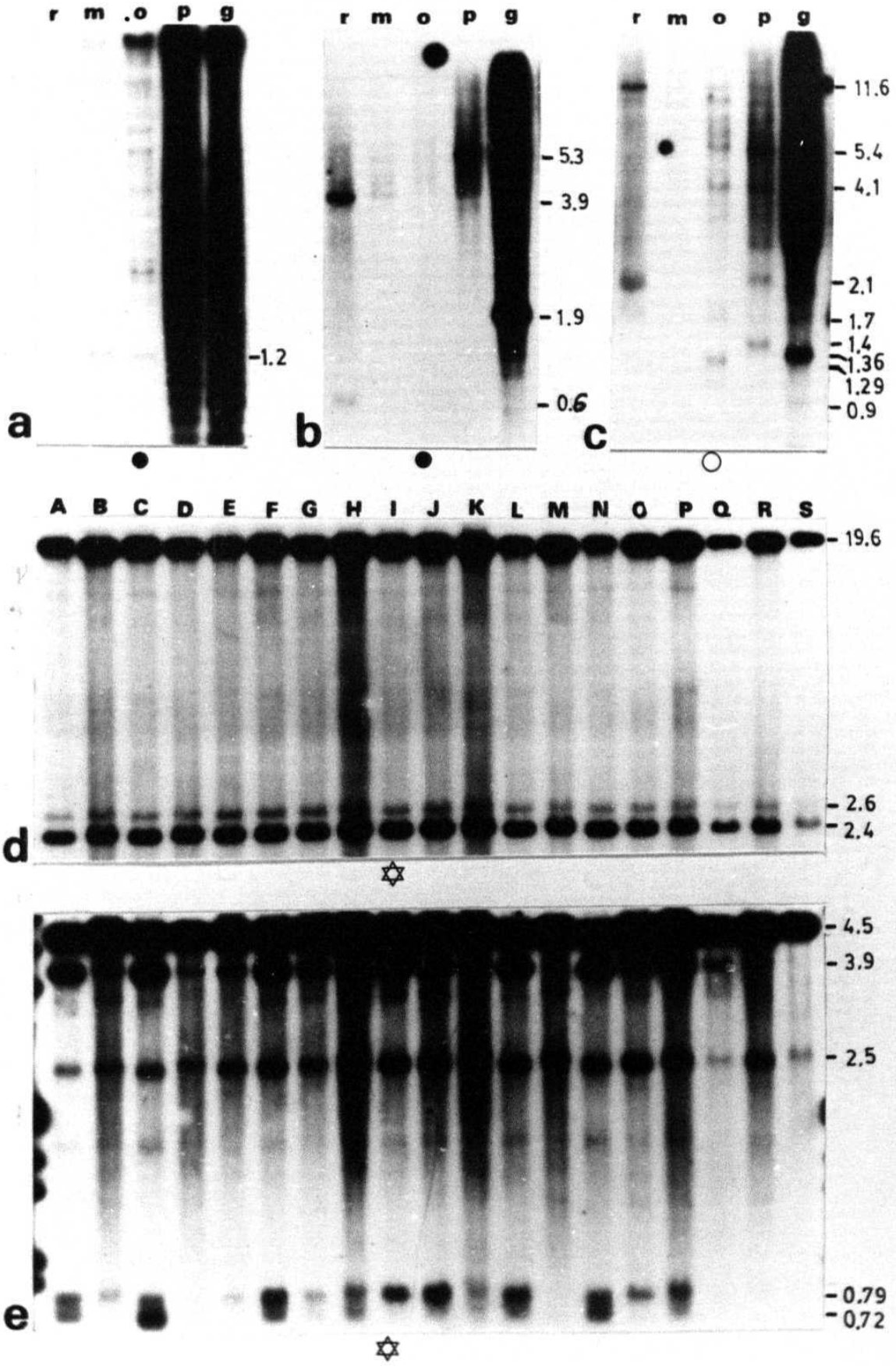
PGB 788:- Digestion with Eco RI (Fig. 7 i) resulted in a conspicuous 0.6 Kb band in all the species, two intense bands each of sizes 7.4 and 1.1 Kb common to P. *orientate, P. purpureum* and *P. glaucum* and a strong 8.9 Kb band shared by *P. ramosum* and *P. mezianum.* Apart from these strong signals, a faint 3.8 Kb band was present in *P. ramosum, P. mezianum* and *P. orientate.* In addition, a 2.9 Kb and a weak 1.4 Kb band were exclusive to *P. mezianum.* Bam HI profiles (Fig. 7 j) consisted of a very intense 3.0 Kb band common to all the species and a 1.8 Kb fragment in P. *ramosum, P. mezianum* and *P. orientate.* Hind III revealed three fragments (8.1, 6.3 and 0.7 Kb) common to all the species, a weak 3.3 Kb band specific to P. *ramosum* and a 3.5 Kb band in P. *mezianum* and P. *orientate* (Fig. 7 k). Msp I digests (Fig. 7 l) gave three very intense bands in each of the species, and the banding pattern differentiated P. *ramosum, P. mezianum* and P. *orientate* from the other two species. Apart from these three bands, a faint 3.2 Kb band specific to P. *ramosum* and a 3.6 Kb band common to P. *mezianum* and P. *orientate* were found. Restriction profiles generated by Pst I, Eco RV, Sac I and Bgl I did not reveal variations in the sequences homologous to PGB 788 (Table V).

PGB 625:- Typical ladder pattern of bands with an internal progression of a 127 bp long monomer was observed in Bam HI digests of P. *glaucum* and P. *purpureum* genomic DNA (Fig. 8 a). P. *orientate* also showed a faint ladder but with a different size monomer. P. *mezianum* showed the faintest ladder and the repeating unit of the ladder was different from P. *orientate, P. purpureum* and P. *glaucum.* After very long exposure, a faint 1.2 Kb band lighted up in P. *ramosum* and it was the only signal.

Fig: 8. Southern blots of restricted genomic DNA from different ***Pennisetum*** species: **r - *P. ramosum*** **m - *P. mezianum***; **o - *P oriantale*** **p - *P. purpureum***; **g - *P. glaucum*** hybridized to pearlmillet repetitive DNA sequence probes - PGB 625 (a); PGB 662 (b & c) and the enzymes used are : Bam HI (●)or Bgl I (○).

Southern blot of Sac I(✡) digested genomic **DNA** of different accessions of *P.* ***glaucum*** hybridized to ^{32}P- labelled pearl millet repetitive DNA sequence clones - PGB 058 (d) and PGB 074 (e).

Accessions are: A - IP 8166; B - IP 8275; C - IP 10401; D - IP 12598; E - IP 12617; F - IP 5009; G - IP 9712; H - GWS 14; I - VCM 36; J - Albino; K - PMC 23A; L - DSA 59-1 A; M - DSA118A; N - DSA 134A; O - DSA 144A; P - ICMA 88004; Q - IP 18292; R - IP 18293; S- IP 18294


r m .o p g
1.2
a
r m o p g
5.3
3.9
1.9
0.6
b
r m o p g
11.6
5.4
4.1
2.1
1.7
1.4
1.36
1.29
0.9
c
A B C D E F G H I J K L M N O P Q R S
19.6
2.6
2.4
d
4.5
3.9
2.5
0.79
0.72
e

observed in Bam HI digested P. *ramosum* **DNA** homologous to PGB 625. Ladder of fragments was found in Eco **RI**, Eco RV and Hind **III** digested genomic DNAs of all except that of P. *ramosum* (data not shown). Nature of the ladder and intensities of the repeating units were same as that revealed by Bam HI. Apart from a smear, no discrete band was observed in Pst I, Sac I, Bgl I and Msp I digests except in P. *ramosum.* In all instances the intensity of the smear was almost identical between P. *glaucum* and P. *purpureum,* less in P. *orientale* and least in P. *mezianum.* In each digest, P. *ramosum* produced a very faint band and the size of such bands varied depending on the enzyme used.

PGB 662:- Bam HI restriction pattern (Fig. 8 b) consisted of a thick smear with two bands of sizes 1.9 and 0.6 Kb in P. *glaucum.* One band each of sizes 5.3 Kb and 3.9 Kb was found against a very weak background smear in P. *purpureum* and P. *ramosum* respectively. A very faint smear with no discrete band was noted in P. *mezianum* and P. *orientale.* Bgl I restriction profiles (Fig. 8 c) comprised a thick smear with a strong band and a faint band in P. *glaucum,* four bands each in P. *purpureum* and P. *orientale* and two bands in P. *ramosum.* Two bands (5.4 and 4.1 Kb) were common to P. *purpureum* and P. *orientale* and one band (2.1 Kb) to P. *ramosum and P. purpureum.* Other bands were specific to each species. A weak smear was observed in P. *mezianum* with no discrete bands. Apart from the smear of different intensities no discrete band was observed in other enzyme digests.

Restriction analysis of pearl millet accessions:-

Southern blots of Bam HI, Eco RI, Sac I Bgl I, Msp I, Hinf I and Hae III digested genomic **DNA** of nineteen pearl millet accessions (Table I) were hybridized with repetitive DNA probes to examine the changes which occurred in each of these sequences among different accessions. All probe-enzyme combinations except PGB 074 in combination with Sac I did not reveal polymorphism among the accessions and a

representative autoradiogram is shown in Fig. 8d. In addition to the conserved bands, PGB 074 in combination with Sac I generated two faint bands (0.79 and 0.72 Kb) in some accessions (Fig. 8 e).

Distribution and modulation of repetitive DNA sequences :

Relative abundance of the sequences related to pearl millet repetitive DNA sequences among *Pennisetum* species (Table IV) indicates that the simple interspersed sequences have almost the same level of abundance across the species. Differences in the relative abundance of PGP 005 and PGB 727 reflect the differences in the copy numbers of different configurations and variation in the abundance of PGB 727 between species indicate their differential distribution. High abundance of PGE 015 in *P. glaucum* and low levels in *P. mezianum* and *P. orientale* indicate duplication of a specific chromosome carrying PGE 015 in *P. glaucum* and dilution effect of extra chromosomes or the absence of this sequence on the extra chromosome in *P. mezianum* and *P. orientate.* The PGE 123 and PGB 058 are likely to be located on the same physical entities for which *P. glaucum* is a secondary polyploid as evident from the higher levels of these two sequences in *P. glaucum.* Similarly the chromosome(s) carrying sequences homologous to PGB 074 and PGB 582 are duplicated in *P. orientate.* PGB 727 and PGB 788 related sequences are likely to be located on the same chromosomes and the variation in their abundance among *P. mezianum, P. orientate* and *P. glaucum* may represent their differential distribution on the extra chromosome(s) in these species. The lower levels of PGE 015, PGE123 and PGB 107 in *P. purpureum* are indicative of the dilution effect of the second genome carrying low copy numbers whereas high abundance of PGB 074, PGB 103 and PGB 582 related sequences in *P. purpureum* may reflect the higher copy numbers in the second genome.

Reiteration of sequences related to PGE 096, PGB 625 and PGB 662 vary over 10-fold across species which is a reflection of the modulation of these sequences

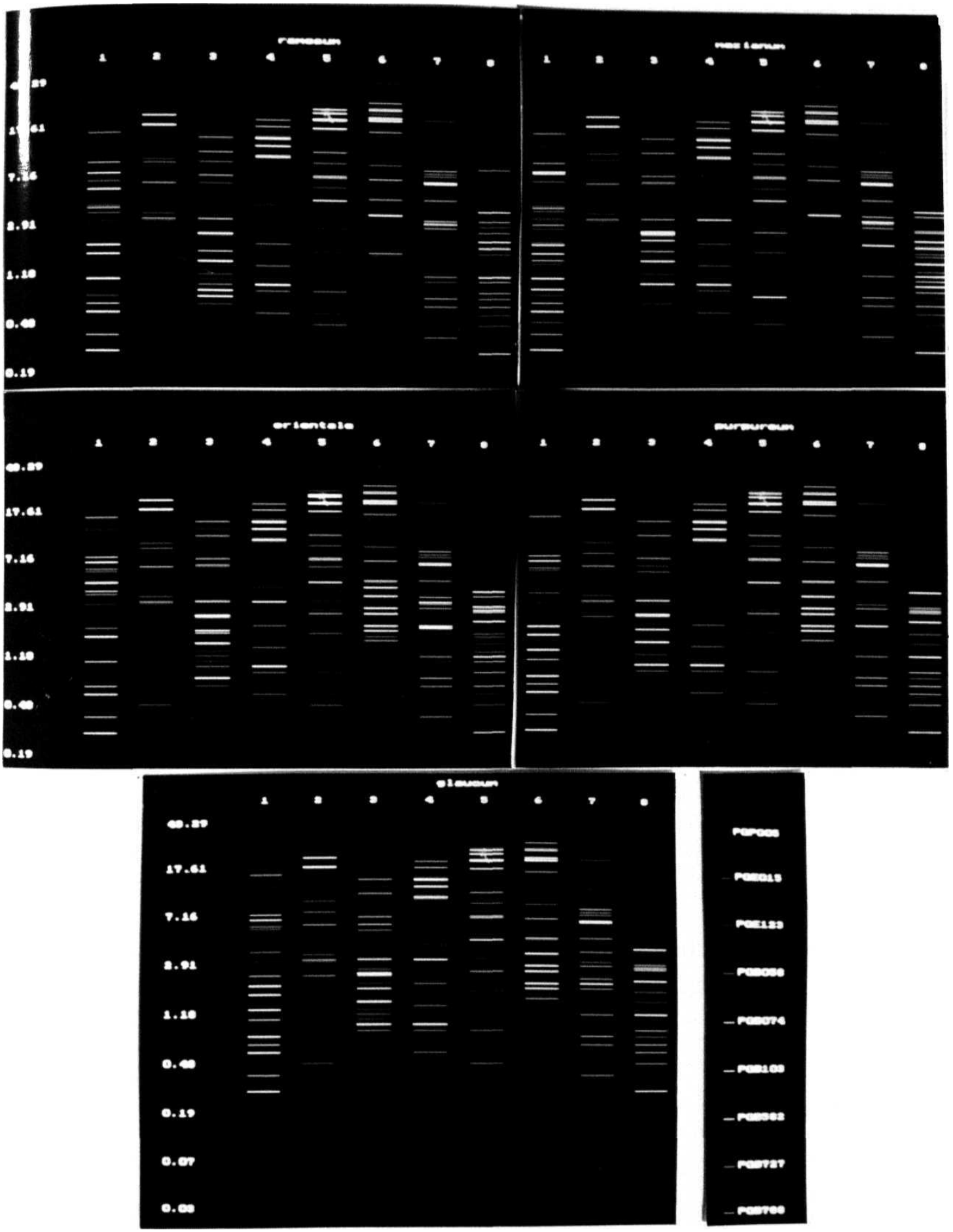
during evolution. The results show that the abundance of sequences homologous to PGB 625 is maintained while the PGE 096 and PGB 662 related sequences are deamplified following polyploidization of *P. purpureum*. Modulation in the repetitive DNA sequences is known among related species. Sequences related to several repetitive sequences of *H. vulgare* are in low abundance in wild species (Sonina et al. 1989; Vershinin et al. 1990; Belostotsky and Ananiev 1990). Sequences related to pHch1 and pHch3 are distributed at varying levels in 19 species of *Hordeum* representing the basic genomes H, I, X and Y (Ferrer et al.1995). Sequences related to tandemly repeated pAs1 family of *Aegilops squarrosa* are 10- to 100- fold less abundant in several related species (Nagaki et al. 1995). Sequences homologous to RS-1 family of oats were found distributed in 18 species at varying levels (Fabijanski et al. 1990). Sequences homologous to the satellite DNA of perennial oats is in high copy number in sub-genus *Helictotrichon,* in low abundance in species of sub-genera *Pratavenastrum* and *Pubavenastrum*, and absent in species of the related genera, *Arrhenatherum* and *Avena* (Grebenstein et al.1995). Among *Oryza* species, sequences related to repetitive DNA isolated from one genotype are found modulated in species with different genomes (Wu and Wu 1992: Mawal et al. 1995). Sequences related to RRS7 family of tandemly repeated DNA isolated from *O. alta* were found tandemly arranged in AA, BB, BBCC and EE genomes, and at a detectable level in F genome (Li et al. 1996).

Southern blot analysis of cloned fragments with genomic DNA as probe is found to be an efficient method of investigating the distribution and abundance of homologous sequences. Data on the restriction analysis of genomic DNA corroborate these findings.

Nature and extent of variation in repetitive DNA sequences:

Restriction fragment length variations revealed by different enzymes differed markedly (Fig. 9). Eco RI, Bam HI, and Msp I revealed maximum polymorphism in the

Fig 9. Species-wise panels of DNA fragments generated by each enzyme hybridizing to the set of probes were merged using a C-based computer programme. Fragments hybridizing to each probe are colour coded (given at right). Numbers 1-8 represent digests of Eco RI, Pst I, Bam HI, Eco RV, Sac I, Bgl I, Hind III and Msp I respectively.


1 2 3 4 5 6 7 8
7.16
2.91
1.18
0.48
0.19
orientale
17.61
purpureum
40.29
0.07
0.08
PGP008
PGE015
PGE123
PGE056
PGE074
PGE109
PGE082
PGE727
PGE768

sequences homologous to seven out of the nine repetitive DNA probes(Table V).

Sequences related to PGE 015 and PGE 123 diverged maximally among the sequences examined while sequences related to PGB 058 are highly conserved across the species. Other sequences have shown almost same extent of divergence among the species. The four enzymes which detected divergence among the sequences related to PGB 103 stem from the variations limited to *P. ramosum*. Repetitive DNA sequences diverge rapidly and are highly variable. Divergence, homogenization and fixation of variants occur during the evolution of repetitive DNA sequences (Dover 1982). Results obtained in the present study suggest that different repetitive DNA sequences differ in the extent of their variation / divergence. This may reflect differential selection pressure on the variants within a family.

Restriction polymorphism of simple interspersed sequences:

Sizes and distribution of conserved fragments among *Pennisetum* species are summarized in Table VI. Fragments conserved across all the species are given (Tables VI & VII) separately from those which have group-wise affinity. Based on the fragments detected in the respective digests and the hybridisation signal intensities, the shared fragments can be categorized into (i) single fragment showing intense hybridization signal and (ii) fragments with faint signal in addition to the intense bands resulting from variation in a proportion of the family of repetitive sequences (Table VII). Similarly the faint species-specific fragments (Table VII) imply that changes occurred in a low proportion of sequences in a family of sequences during or after the species divergence.

Out of the total fragments detected by the nine simple interspersed sequences in combination with eight enzymes(Fig. 10), 45% were conserved across the species, over 33% were common to two or more species and over 21% were species-specific (Table VII). Variation in the distribution of shared fragments between species (Table VII) may represent the phylogenetic affinity among species. Repetitive sequence variation

Table VI. Size(s) and distribution of conserved fragments among *Pennisetum* species *

Probe-Enzyme	Fragments (Kb)					
	Conserved	Limited to				
	r m o p g	r	m	o	p	g
PGP 005-Bam HI	7.1/1.1.	-	-	1.3	2.1	2.1/1.3
PGP 005-Eco RI	7.1/2.7/1.7/0.9/0.5	-	-	3.3	-	3.3
PGE 015-Pst I	3.1	-	-	-	7.7/2.4	7.7/2.4
PGE 123-Pst I	3.0	-	-	0.5	0.5	0.5
PGB 074-Eco RV	10.1	-	3.1	3.1	-	3.1
PGB 074- Msp I	0.9/0.3	1.9	1.9	2.8	2.8	2.8
PGB 582-Eco RI	2.0/0.5	1.1	1.1/0.8	1.1	1.3/0.8	1.3/0.8
PGB 582-Bgl I	21.6/3.6	-	-	2.9/2.6	2.9/2.6	2.9/2.6
				2.1/1.9	2.1/1.9	2.1/1.9
PGB 727-Msp I	2.6/0.7/0.5	-	2.8	2.8	2.8	2.8
PGB 788-Eco RI	0.6/0.3	0.8/3.8	0.8/3.8	7.4/3.8	7.4/1.0	7.4/1.0
PGB 788-Bam HI	3.2	1.8	1.8	1.8	-	-
PGB 788-Hind III	8.1/6.3/0.7	-	3.5	3.5	-	-

r - *P. ramosum*, m - *P. meziamum*, o - *P. orientate*, p - *P. purpureum*, g - *P. glaucum*.

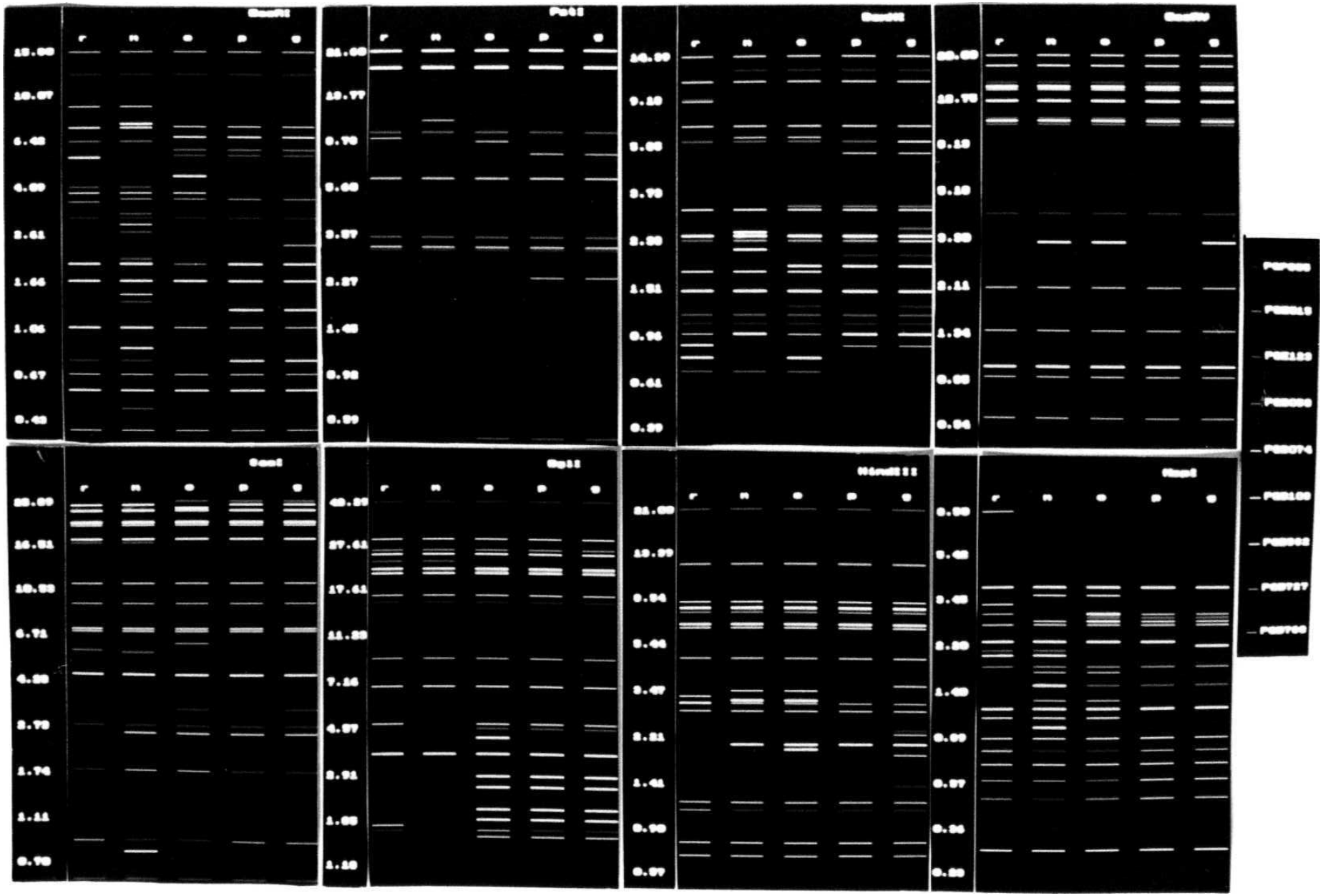
Table VII. Nature and distribution of restriction polymorphism among *Pennisetum* species*

Probe	Fragments																			
	Conserved	**Shared															Species specific			
	rmopg	rm	ro	rmo	mog	ropg	rpg	mo	rmop	mopg	mpg	opg	og	pg			r	m	o	g
PGP 005	22					2				1				2	1		1	1		2
PGE 015	14	1		1		1	1	3				1		4			2	3	3	
PGE 123	5	7		2								10						4		1
PGB 058	8	1										1							1	
PGB 074	6	3			1							4						2		
PGB 103	5									6							5	2		
PGB 582	12		1	1					1		1	4					2	4	7	1
PGB 727	27					2				3							1			3
PGB 788	10	1		6				2				2		3			2	2		2
Total	109	13	1	10	1	5	1	5	1	10	1	22	2	8			13	18	11	9

*T - *P. ramosum*, m - *P. mezianum*, o - *P. orientale*, p - *P. purpureum*, g - *P. glaucum*.

**represented singly or in groups.

Fig 10. Enzyme-wise panels of DNA fragments generated by each enzyme hybridizing to the set of probes were merged in each species using a C-based computer programme. Fragments hybridising to each probe are colour coded (given at right). r, m, op and g represent *P. ramosum, P. mezianum, P. orientate, P. purpureum,* and *P. glaucum* respectively



has been used to examine the phylogenetic hypothesis on the origin of *Sorghum* species. Based on cross hybridization pattern it was concluded that S. *bicolor* and S. *halepense* belong to section *Eusorghum* and S. *versicolor* which diverged from them belongs to *Parasorghum* (Hoang - Tang et al. 1991). Grouping of six wild species of *Hordeum* based on the copy number and genome organization similarity of the sequences homologous to six highly repetitive DNA isolated from *H. vulgare* was in good agreement with the data on interspecific crossing and on chromosome pairing in hybrid (Vershinin et al. 1990). From the fragments detected by repetitive sequences among 19 *Hordeum* species , Ferrer et al. (1995) concluded that species that share a basic genome show more similar fragment hybridisation pattern than species with different genomes.

Molecular probes:

The diagnostic probe-enzyme combinations that reveal distinct species-specific bands with relative ease are summarized in Table VIII. These are vital to identify the respective genome in alien genetic background. For example, *P. ramosum* could be differentiated from other species by the Bam HI fragments hybridizing to the probe, PGP 005 and Bgl I fragments of P. *mezianum* homologous to the same probe are distinct from other species. PGE 123 probe differentiates *P. ramosum* and *P. mezianum* from the other three in the sizes of their genomic fragments generated by Eco RI, Bam HI, Pst I, Sac I, Bgl I and Msp |. Deploying these probes in our laboratory species-specific fragments and sequence rearrangements could be identified in the interspecific hybrids, *P. glaucum* X P. *mezianum* and P. *glaucum X P. orientate* and their derivatives. Changes in the structural organization of rye and barley repetitive sequences have been reported in hybrids of *H. vulgare* X *Secale cereale* and in the tri-generic hybrid *(H. vulgare* X *Triticum timopheevi*) X *S.cereale* (Svitashev et al. 1995). Vershinin et al (1992) showed structural rearrangements of some highly repetitive DNA sequences in barley interspecific hybrids only a few years old. In *Arabidopsis suecica,*

Table VIII. Diagnostic probe-enzyme combinations to distinguish *Pennisetum* species *

Probe	Eco RI	Bam HI	Pst I	Sac I	Bgl I	Msp I
PGP 005	–	r/mopg	–	–	m/ropg	–
PGE 015	rm/opg	r/mo/pg	r/m/o/pg	–	m/ropg	mo/rpg
PGE 123	rm/opg	rm/opg	rm/opg	rm/opg	rm/opg	rm/opg
PGB 058	rm/opg	–	–	–	–	–
PGB 074	rm/opg	rm/opg	–	–	–	–
PGB 103	r/mopg	r/mopg	–	–	–	r/mopg
PGB 582	rmo/pg	–	–	–	–	rmop/g
PGB 727	–	r/mopg	–	–	m/ropg	r/mopg
PGB 788	rm/opg	–	–	–	–	–

* r - *P. ramosum*, m - *P. mezianum*, o - *P. orientale*, p - *P. purpureum*, g - *P. glaucum*.

Table IX. Pair-wise similarity (%)among *Pennisetum* species in their repetitive DNA homologous to cloned pearl millet sequences

Species:	*P. mezianum* (x=8)	*P. orientale* (x=9)	*P. purpureum* (x=7)	*P. glaucum* (x=7)
P. ramosum (x=5)	80	67	67	64
P. mezianum (x=8)		73	68	67
P. orientale (x=9)			87	87
P. purpureum(x=7)				97

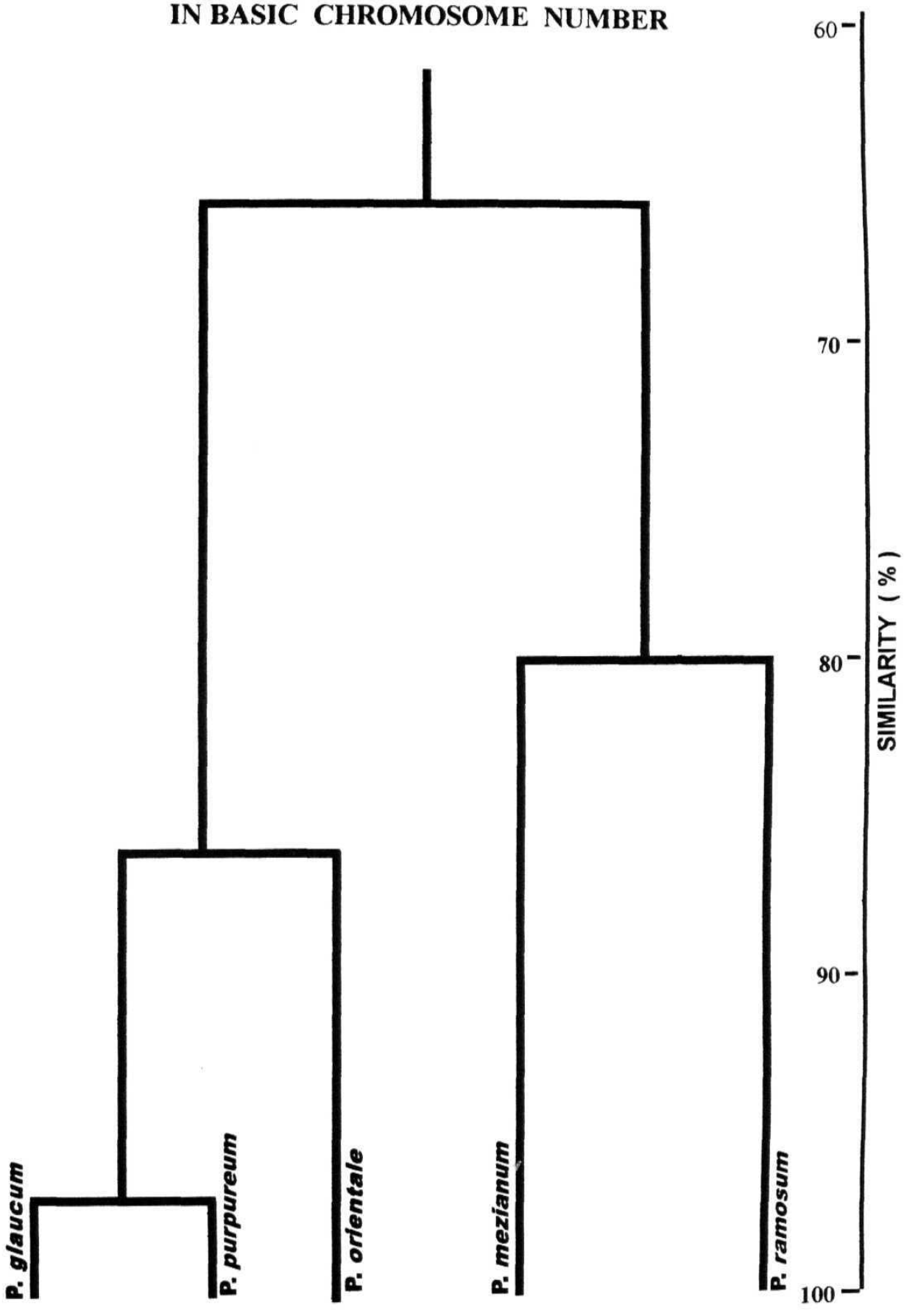
presumed to be a hybrid between *A. thaliana* and *A. arrenosa,* the sequence organization of Bam HI repeat of *A. arrenosa* is conserved whereas the pAal1 repeat sequence of *A. thaliana* had undergone considerable changes (Kamm et al. 1995).

Phylogenetic Analysis

Data on the extent of conservation of sequences homologous to cloned repetitive **DNA** sequences among the five *Pennisetum* species have shown ~2/3 common stem indicating the extent of conservation. The sequence divergence patterns revealed by the restriction polymorphism were used to look at pair-wise similarity index (Table IX). The tetraploid species *P. purpreum* has 97% similarity with P. *glaucum* (one of its diploid progenitor). These two species in turn have 87% similarity with *P. orientate.* The other two species *P. ramosum* and P. *mezianum* have 80% similarity, but both have 64 to 68% similarity with P. *glaucum* and P. *purpureum.* Data on the nature and distribution of different restriction fragments hybridizing to different probes were utilized to construct a phylogenetic tree of *Pennisetum* species (Fig 11) using PHYLIP programme and merged with that of other cereals (Fig 20 discussed in page - 83).

Fig **11.** Phylogeny of *Pennisetum* species differing in basic chromosome number

PHYLOGENY OF *PENNISETUM* SPECIES DIFFERING IN BASIC CHROMOSOME NUMBER



3.3 MOLECULAR ANALYSIS OF CEREAL GENOMES WITH PEARL MILLET REPETITIVE DNA SEQUENCES

Distribution and modulation:

Southern blots of restricted repetitive DNA clones, hybridization with cereal genomic DNA(s) as probes, densitometric analysis of signal intensities generated by each cereal genomic DNA probes and determination of relative intensities were carried out as described in section (2.3.1) and comparisons were made within and between species.

Hybridization signal intensities of the set of cloned pearl millet repetitive DNA sequences with each cereal genomic DNA as probe are given in Figs. 12 and 13. Relative signal intensity corresponding to each clone with that of PGP 005 generated by cereal genome is summarized in Table X. Over 2-fold variations in the relative intensities were observed for PGE 015 (0.50-1.23), PGE 123 (0.11-0.36), PGB 058 (0.39-0.95), PGB 074 (0.51-1.79), PGB 103 (0.77-2.20), PGB 582 (0.31-1.43), PGB 727 (0.17-0.47) and PGB 788 (0.20-0.80). None of the cereal genomes showed detectable signal intensity with PGB 107. The relative intensity of PGE 096 was over 30-fold less in sorghum as compared to *P. glaucum* and no signal was detectable in the other cereal genomes. Ratios of PGB 625 signal was very low among the cereals as compared to *P. glaucum.* While sorghum and maize had very low levels of PGB 662 signal, no signal was detectable with the other cereal genomes.

Restriction analysis:

Southern blots of Eco RI, Pst I, Bam HI and Eco RV restricted genomic DNA of cereals were hybridized with cloned pearl millet repetitive sequence probes and the restriction profiles generated by each enzyme-probe combination are described

Fig 12. Southern blot of restricted repetitive DNA sequence clones hybridized to: c - labelled vector DNA and sheared ^{32}P labelled genomic DNA(s) from: **P** - **pearl millet** (*P. glaucum*); **S** - **sorghum** (S. *bicolor* var CSH - 11); **M** - **maize(** *Z.mays* var cm 119x cm120); **R** - **rice** (*O. sativa* var basmati). The clone numbers given at the top of each lane with a prefix P ,E or B represent Pst I-, Eco RI-, or Bam HI-based pearl millet genomic libraries respectively, and the corresponding enzyme used for the digestion of the recombinant clones

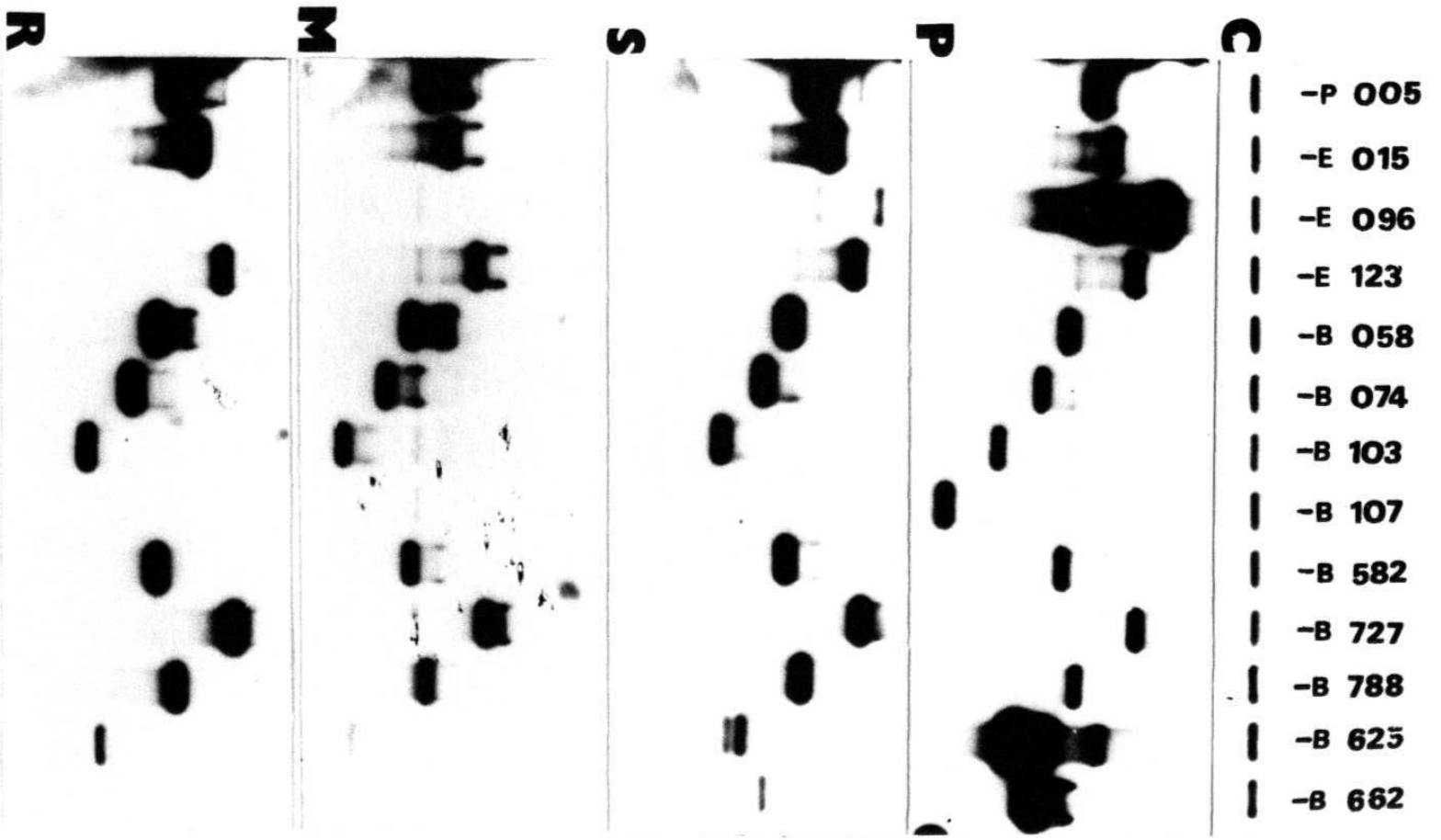

R
M
S
P
C
-P 005
-E 015
-E 096
-E 123
-B 058
-B 074
-B 103
-B 107
-B 582
-B 727
-B 788
-B 625
-B 662

Fig 13. Southern blot of restricted repetitive DNA sequence clones hybridized to: c - labelled vector DNA and sheared ^{32}P labelled genomic DNA(s) from **P** - **pearl millet** *(P. glaucum*); W - **wheat**(*T. aestivum* var Chinese spring); **B** - **barley** *(H. vulgare* var betzes); **O** - **oats** (*Avena sativa),* RY - **rye** (S. *cereale).* The clone numbers given at the top of each lane with a prefix P ,E or B represent Pst I-, Eco RI-, or Bam HI-based pearl millet genomic libraries respectively, and the corresponding enzyme used for the digestion of the recombinant clones

-P 005
-E 015
-E 096
-E 123
-B 058
-B 074
-B 103
-B 107
-B 582
-B 727
-B 788
-B 625
-B 662
C
P
W
B
O
RY

Table X. Relative abundance of sequences in cereal genomes homologous to cloned pearl millet repetitive DNA

Clone*	Pearl millet	Sorghum	Maize	Rice	Wheat	Barley	Oats	Rye
PGP 005	1	1	1	1	1	1	1	1
PGE 015	1.05	0.92	0.77	1.23	0.80	0.52	0.50	0.60
PGE 096	1.04	0.03	-	-	-	-	-	-
PGE 123	0.33	0.24	0.28	0.36	0.20	0.14	0.14	0.11
PGB 058	0.82	0.74	0.54	0.95	0.68	0.59	0.74	0.39
PGB 074	0.87	0.88	0.73	1.79	0.87	0.54	0.96	0.51
PGB 103	1.21	1.07	1.53	2.20	1.44	0.77	1.25	0.80
PGB 107	2.01	-	-	-	-	-	-	-
PGB 582	0.56	0.52	0.71	1.43	0.68	0.33	0.55	0.31
PGB 727	0.29	0.26	0.29	0.47	0.35	0.17	0.25	0.21
PGB 788	0.58	0.47	0.55	0.80	0.67	0.20	0.36	0.26
PGB 625	5.51	0.63	0.58	0.12	0.84	0.06	0.34	0.07
PGB 662	2.29	0.11	0.23	-	-	-	-	-

*Relative abundance of different repetitive sequences in relation to the copy number of PGP 005.

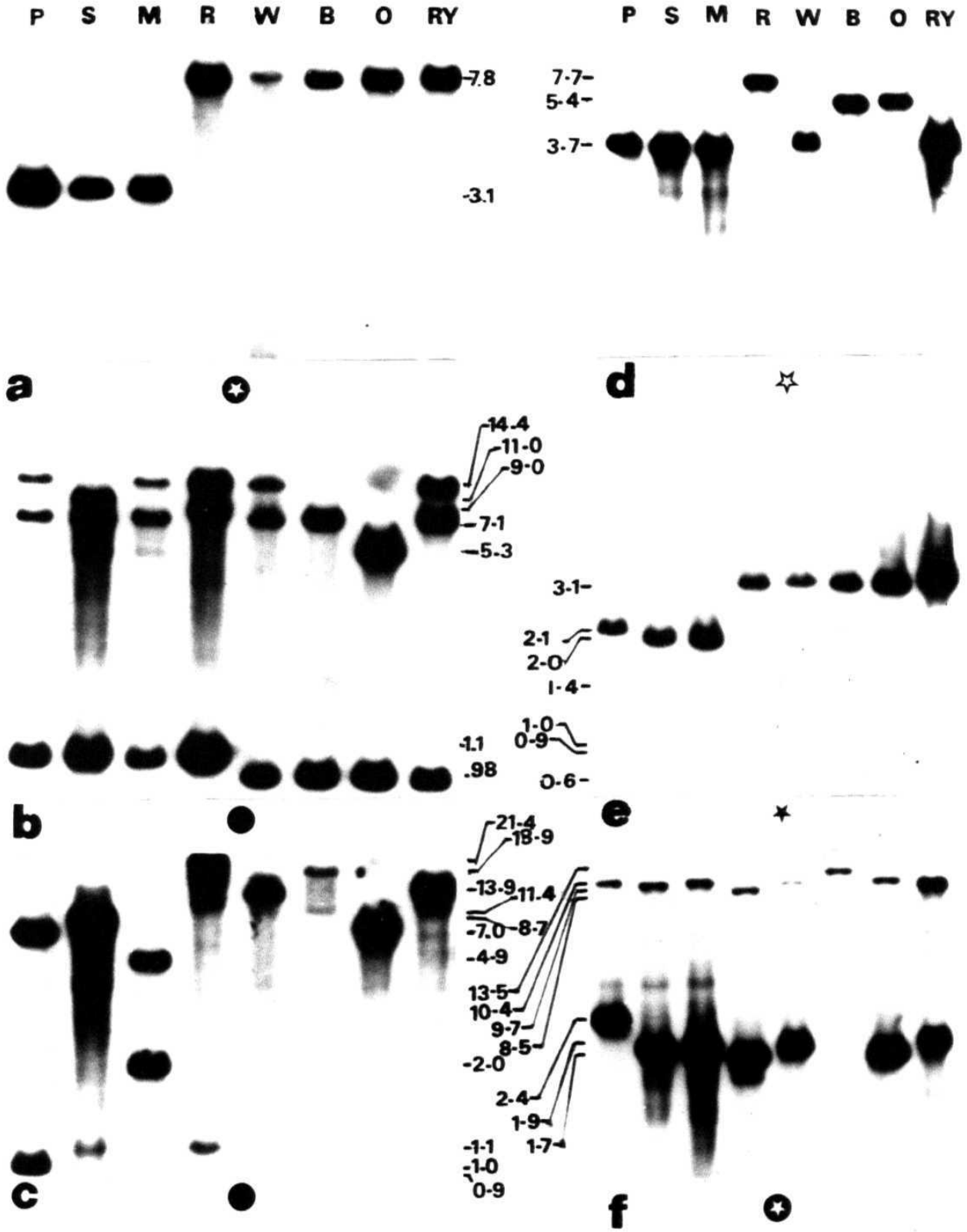
below. PGE 096 and PGB 107 are not distributed at adequate levels among cereals (Table X).

PGP 005:- Pst I digestion resulted in a 3.1 Kb fragment each in pearl millet, sorghum and maize and 7.8 Kb fragment each in the remaining cereals (Fig. 14 a). Restriction profiles generated by Bam HI (Fig. 14 b) consisted of a 14.4 Kb fragment each in pearl millet, maize, rice, wheat and rye and a 7.1 Kb fragment each in pearl millet, sorghum, maize, wheat, barley and rye. Fragments 11 Kb, 9 Kb and 5.3 Kb were specific to sorghum, rice and oats respectively. A common 1.1 Kb fragment in pearl millet, sorghum, maize and rice (tropical cereals) and another fragment (0.9 Kb) common to wheat, barley, oats and rye (temperate cereals) were present. Eco RI generated (data not shown) a 12.8 Kb fragment in all the cereals and faint bands specific to sorghum (2.7 Kb), wheat (2.9 Kb) and oats (2.1 Kb). Eco RV restriction profiles did not reveal any polymorphism.

PGE 015:- Bam HI restriction pattern (Fig. 14 c) revealed strong bands specific to pearl millet (7.0 Kb), sorghum (8.7 Kb), maize (4.9 and 2.0 Kb), rice (21.4 Kb), barley (18.9), oats (7.6 Kb) and a faint 18.9 Kb band in rice. An intense 13.9 Kb band was common to wheat and rye and a faint 11.4 Kb band was common to rice and barley. In addition, a faint 1.1 Kb band common to sorghum, maize and rice and a dense 1.0 Kb band in pearl millet common to winter cereals were also present.

Eco RI generated a 3.7 Kb fragment common to pearl millet, sorghum, maize, wheat and rye, a 5.4 Kb fragment in barley and oats and a 7.7 Kb fragment specific to rice (Fig. 14 d). Restriction profiles generated by Eco RV (Fig. 14 e) consisted of a strong 3.1 Kb fragment in rice, wheat, barley, oats and rye, a 2.0 Kb fragment in sorghum and maize and a 2.1 Kb fragment specific to pearl millet. In addition, a 0.6 Kb fragment in tropical cereals, a 1.0 Kb fragment in sorghum and maize and 1.4 Kb and 0.9 Kb fragments specific to pearl millet were present. Pst I generated intense bands of sizes 1.9 Kb in sorghum, maize, wheat and rye, 1.7 Kb in rice and oats, and a 2.4 Kb

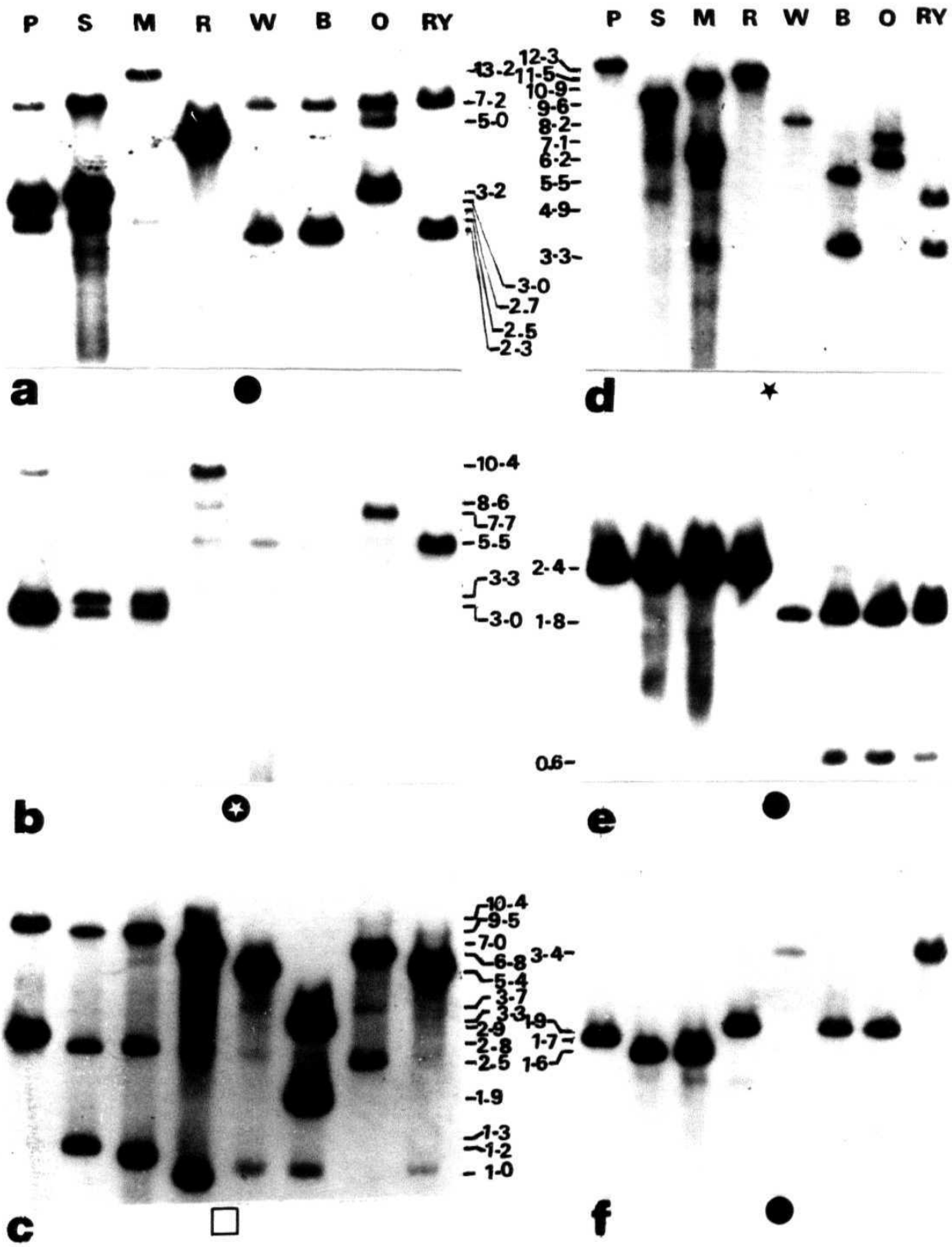
Fig 14. Southern blots of restricted genomic DNA from different cereals: **P** - **pearl millet (***P. glaucum)*; **S** - **sorghum** (*S. bicolor* var CSH - 11); **M** - **maize** (*Z. mays* var cm 119 x cm 120); **R** - **rice** (*O . sativa* var basmati); **W** - **wheat** (*T . aestivum* var chinese spring); **B** - **barley** (*H. vulgare* var betzes); **O** - **oats** (*A. sativa*); **RY** - **rye** (*S. cereale*) hybridized to pearl millet repetitive DNA sequence probes: PGP 005 (**a** & **b**); PGE 015 (**c** - **f**); and the enzymes used are Bam HI (●), Eco RI (✡), Eco RV (✸), or Pst I (✪).


P S M R W B O RY
7·8
-3.1
P S M R W B O RY
7·7-
5·4-
3·7-
a
14·4
11·0
9·0
7·1
5·3
1·1
.98
d
3·1-
2·1
2·0
1·4-
1·0
0·9
0·6-
b
21·4
18·9
13·9
11·4
7·0
8·7
4·9
13·5
10·4
9·7
8·5
2·0
2·4
1·9
1·7
1·1
1·0
0·9
e
c
f

fragment specific to pearl millet (Fig. 14 f). Relatively faint high molecular size fragments common to sorghum, maize, wheat, oats and rye (9.7 Kb), 10.4Kb specific to pearl millet, 8.5 Kb specific to rice and 13.5 Kb specific to barley were also present.

PGE 123:- Bam HI generated a 7.2 Kb fragment each in pearl millet, sorghum, wheat, barley, oats and rye, a 2.7 Kb fragment common to sorghum and maize, a 3.2 Kb fragment common to sorghum and oats, one 5.0 Kb fragment common to rice and oats and one 2.3 Kb fragment common to wheat, barley and rye (Fig. 15 a). In addition, 3.0 and 2.5 Kb fragments in pearl millet, and a 13.2 Kb fragment specific to maize and two faint bands specific to rice were also found(not marked in Fig. 15.a) . The Pst I generated profiles (Fig. 15 b) consisted of a 5.5 Kb fragment in all (differing in intensities) except in oats, a 10.4 Kb fragment in tropical cereals, a 3.0 Kb fragment common to pearl millet, sorghum and maize, a 3.3 Kb fragment in sorghum and maize and a faint 8.6 Kb fragment common to rice, barley and rye. A 7.7 Kb fragment was specific to oats. In addition, Xba I restriction analysis with PGE 123 revealed two fragments each (9.5 and 2.8 Kb) in sorghum and maize and a 5.4 Kb fragment common to wheat and rye and a faint 1.0 Kb fragment common to rice, wheat, barley and rye (Fig. 15 c). Additional fragments observed in sorghum (1.3 Kb), maize (1.2 Kb), barley (3.3 and 1.9Kb) and the profile of bands observed in pearl millet (10.4 and 2.9 Kb) and oats (6.8, 3.7 and 2.5 Kb) were specific. The Eco RV digestion profiles (Fig. 15 d) consisted of fragments specific to pearl millet (12.3 Kb), sorghum (9.6 Kb), maize (10.9 and 3.3 Kb), rice (11.5 Kb), wheat (8.2 Kb), barley (5.5 Kb) oats (7.1 Kb) and rye (4.9 Kb). In addition, a 6.2 Kb fragment common to maize and oats and a 3.3 Kb fragment common to maize, barley and rye were present. The Eco RI generated restriction profile contained (data not shown) a 6.4 Kb fragment common to tropical cereals, two fragments each of sizes 3.2 Kb and 2.5 Kb common to sorghum and maize and a 4.7 Kb fragment common to wheat, barley and rye. An additional 3.7 Kb fragment observed in rice and the single 11.2 Kb band generated in oats were specific to them. One faint 0.4 Kb fragment was common to winter cereals.

Fig 15. Southern blots of restricted genomic DNA from different cereals: **P** - **pearl millet (***P. glaucum)*; **S** - **sorghum** (*S. bicolor* var CSH - 11); **M** - **maize** (*Z. mays* var cm 119 x cm 120); **R** - **rice** (*O . sativa* var basmati); **W** - **wheat** (*T . aestivum* var chinese spring); **B** - **barley** (*H. vulgare* var betzes); **O** - **oats** (*A. sativa*); **RY** - **rye** (*S. cereale*) hybridized to pearl millet repetitive DNA sequence probes: PGE 123 (**a** -**d**); PGB 058 (**e**) ; PGB 074 (**f**) and the enzymes used are Bam HI (●), Eco RV (✶), Pst I (✪) or Xba I (□)


P S M R W B O RY
12-3
11-5
10-9
9-6
8-2
7-1
6-2
5-5
4-9
3-3
2-4
1-8
0.6
13-2
7-2
5-0
3-2
3-0
2-7
2-5
2-3
10-4
8-6
7-7
5-5
3-3
3-0
9-5
7-0
6-8
5-4
3-7
3-3
2-9
2-8
2-5
1-9
1-7
1-6
1-3
1-2
1-0
a
b
c
d
e
f

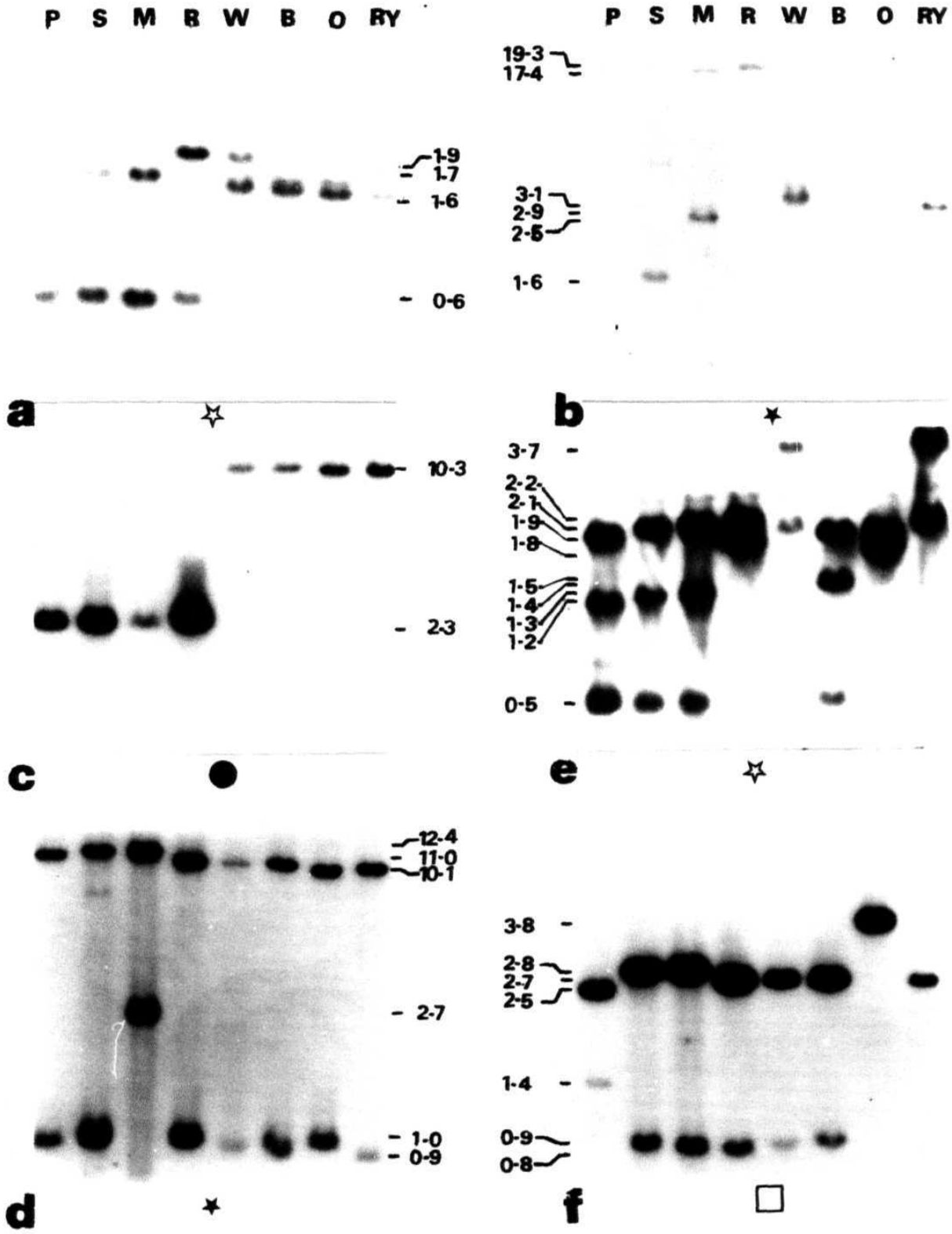
PGB 058:- Bam HI digests gave a 2.4 Kb fragment each in tropical cereals and 1.8 Kb and 0.6 Kb fragments in the winter cereals (Fig. 15 e). Patterns revealed by other enzymes were not informative.

PGB 074:- Bam HI restriction profiles (Fig. 15 f) consisted of a 1.7 Kb fragment specific to pearl millet, a 1.6 Kb fragment common to sorghum and maize, a 1.9 Kb fragment common to rice, barley and oats and a 3.4 Kb fragment in wheat and rye. Restriction profiles generated by other enzymes were not informative.

PGB 103:- Eco RI digestion resulted in a 1.7 Kb fragment common to pearl millet, sorghum and maize, a 0.6 Kb fragment in the tropical cereals and a 16Kb fragment in winter cereals. A 1.9 Kb fragment was common to rice, wheat and rye (Fig. 16 a). A 19.3 Kb fragment common to pearl millet, rice and barley and a 17.4 Kb fragment common to sorghum, maize and oats in addition to a fragment each specific to sorghum (1.6 Kb), maize (2.5 Kb), wheat (3.1 Kb) and rye (2.9 Kb) were observed in Eco RV digests (Fig. 16 b). Digestion with Bam HI resulted in a 4.6 Kb fragment common to sorghum and maize, a 5.1 Kb fragment common to winter cereals, a 0.9 Kb fragment specific to pearl millet and a 5.3 Kb fragment in rice (data not shown).

PGB 582:- Bam HI generated a 2.3 Kb fragment each in tropical cereals and one 10.3 Kb fragment common to winter cereals (Fig. 16 c). Restriction pattern generated by Eco RV comprised of a 12.4 Kb fragment common to pearl millet, sorghum and maize, a 11.0 Kb fragment common to rice, wheat, barley and rye and a 10.1 Kb fragment specific to oats (Fig. 16 d). A 1.0 Kb fragment common to pearl millet, sorghum, rice, wheat, barley oats and one fragment each specific to maize (2.7 Kb) and rye (0.9 Kb) were also noticed. Digestion with Eco Rl resulted in a 1.9 Kb fragment each in pearl millet, barley and oats, a 2.1 Kb fragment each in sorghum, maize, wheat and rye, a 0.5 Kb fragment each in pearl millet, sorghum, maize and barley ,a 1.8 Kb fragment common to rice and oats and a 3.7 Kb fragment common to wheat and rye (Fig. 16 e).

Fig 16. Southern blots of restricted genomic DNA from different cereals: **P** - **pearl millet (***P. glaucum)*; **S** - **sorghum** (*S. bicolor* var CSH - 11); **M** - **maize** (*Z. mays* var cm 119 x cm 120); **R** - **rice** (*O . sativa* var basmati); **W** - **wheat** (*T . aestivum* var chinese spring); **B** - **barley** (*H. vulgare* var betzes); **O** - **oats** (*A. sativa*); **RY** - **rye** (*S. cereale*) hybridized to pearl millet repetitive DNA sequence probes: PGB 103 (**a** & **b**) ; PGB 582 (**c** - **f**) and the enzymes used are Bam HI (●), Eco RI (✡), Eco RV (▲) or **Xba I** (□)


P S M R W B O RY
1·9
1·7
1·6
0·6
a
10·3
2·3
c
12·4
11·0
10·1
2·7
1·0
0·9
d
P S M R W B O RY
19·3
17·4
3·1
2·9
2·5
1·6
b
3·7
2·2
2·1
1·9
1·8
1·5
1·4
1·3
1·2
0·5
e
3·8
2·8
2·7
2·5
1·4
0·9
0·8
f

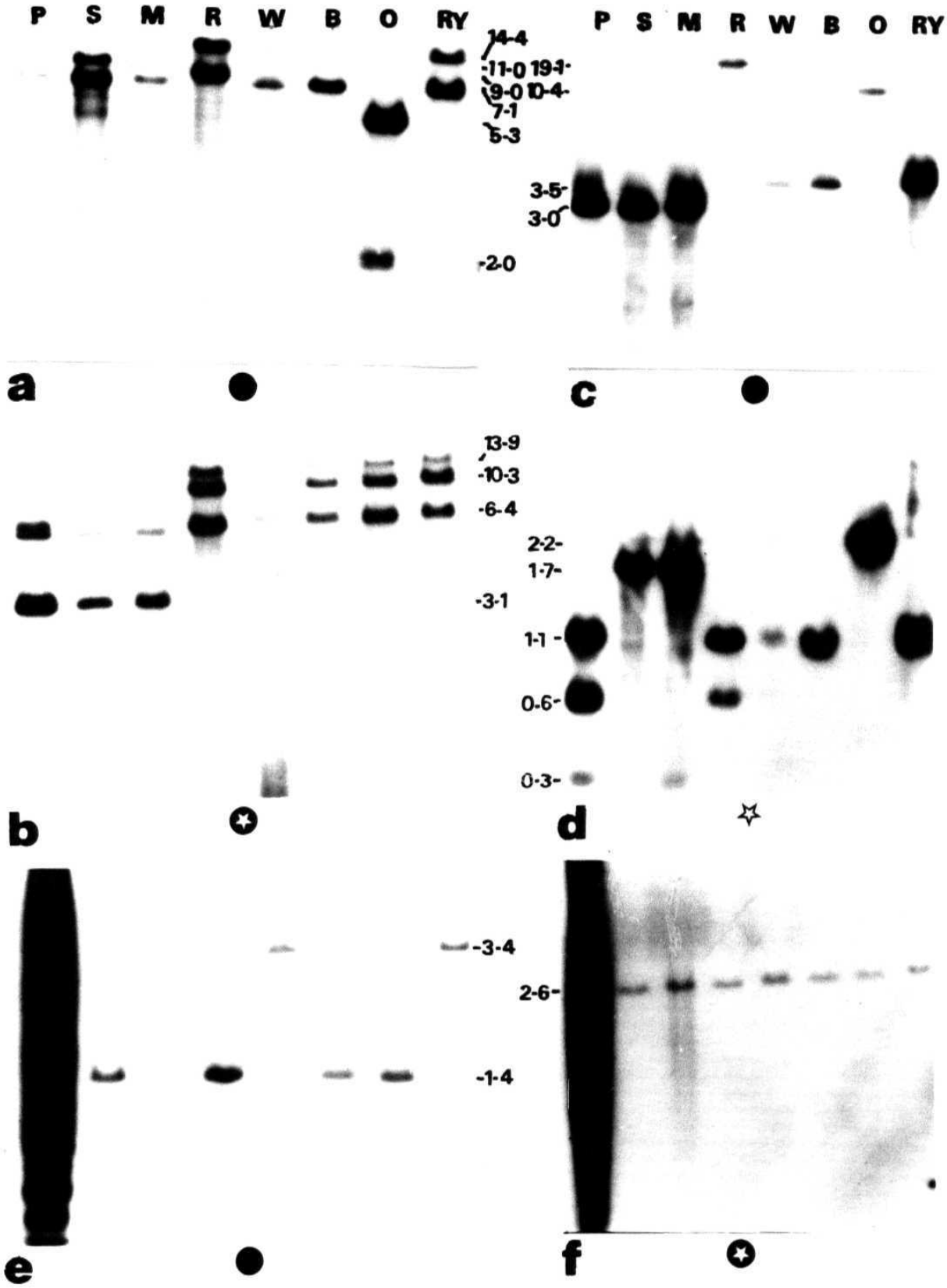
Additional fragments each specific to pearl millet (1.2 Kb), sorghum (1.3 Kb), maize (1.4 Kb), rice (2.2 Kb) and barley (1.5 Kb) were present. Xba I digests (Fig. 16 f) revealed a 2.8 Kb fragment common to sorghum and maize, a 2.7 Kb fragment in rice, wheat, barley and rye, a 0.9 Kb fragment common to sorghum, wheat and barley and a 0.8 Kb fragment common to maize, rice and rye. Two fragments (2.5 Kb and 1.4 Kb) in pearl millet and a single fragment in oats (3.8 Kb) were specific. Pst I did not reveal informative banding profiles.

PGB 727:- High molecular weight fragments generated by Bam HI (Fig. 17 a) were identical to that of PGP 005. Bam HI did not generate conserved low molecular weight fragments homologous to PGB 727. However, a 2.0 Kb specific fragment was observed in oats. The Pst I generated a 6.4 Kb fragment conserved in pearl millet, rice, wheat, barley, oats and rye, a 13.9 Kb fragment in all the cereals (Fig. 17 b), except pearl millet, a 6.1 Kb and a 3.1 Kb fragment common to pearl millet, sorghum and maize and a 10.3 Kb fragment common to rice, wheat, barley, oats and rye. Eco RI and Eco RV did not reveal informative restriction profiles.

PGB 788:- Restriction profiles generated by Bam HI (Fig. 17 c) comprised of a 3.0 Kb fragment common to pearl millet, sorghum and maize, a 3.5 Kb fragment common to wheat, barley and rye and one specific fragment each in rice (19.1 Kb) and oats (10.4 Kb). Eco RI digests (Fig. 17 d) gave a 1.1 Kb fragment common to pearl millet, rice, wheat, barley and rye, a 0.6 Kb fragment each in pearl millet and rice, a faint 0.3 Kb fragment each in pearl millet, sorghum and maize a 1.7 Kb fragment each in sorghum and maize and a single 2.2 Kb specific to oats. Pst I and Eco RV did not reveal informative restriction profiles.

PGB 625:- Digestion with Bam HI resulted in a ladder of fragments in pearl millet DNA (Fig. 17 e). Nature and the organization of the Bam HI ladder of bands detected by PGB 625 is given in page - 29. On prolonged exposure one faint band each of size

Fig 17. Southern blots of restricted genomic DNA from different cereals: **P - pearl millet (***P. glaucum)*; **S - sorghum** (*S. bicolor* var CSH - 11); **M - maize** (*Z. mays* var cm 119 x cm 120); **R - rice** (*O . sativa* var basmati); **W - wheat** (*T . aestivum* var chinese spring); **B - barley** (*H. vulgare* var betzes); **O - oats** (*A. sativa*); **RY - rye** (*S. cereale*) hybridized to pearl millet repetitive DNA sequence probes: PGB 727 (**a & b**); PGB 788 (**c & d**); PGB 625 (**e**); PGB 662 (**f**) and the enzymes used are Bam HI (●), Eco RI (✡), or Pst I (✪).


P S M R W B O RY
14-4
11-0
19-1
9-0
10-4
7-1
5-3
2-0
3-5
3-0
a
c
13-9
10-3
6-4
3-1
2-2
1-7
1-1
0-6
0-3
b
d
3-4
2-6
1-4
e
f

1.4 Kb in sorghum, maize, rice, barley and oats and a faint 3.4 Kb fragment each in wheat and rye were lighted up. Other enzymes did not reveal informative profiles.

PGB 662:- Digestion with Pst I resulted in a thick smear in *P. glaucum* (Fig. 17 f). A very faint fragment of size 2.6 Kb was observed in all other species. Similar patterns was observed in Eco RI and Eco RV digests (data not shown). Bam HI digests of related cereal genomic DNA did not produce signal, while a thick smear was observed in *P. glaucum.*

Distribution and modulation:-

Two to three- fold less abundance of these sequences in barley, oats and rye as compared to *P. glaucum* (Table X) is indicative of the dilution effect of these sequences with the increase in their genomic size. The higher level of PGB 074, PGB 103, PGB 582, PGB 727 and PGB 788 related sequences in wheat (Table X) suggests either amplification of these sequences following polyploidization or their abundance in one of their progenitor genomes. Consistently higher relative abundance in rice is not surprising considering the small genome size (Moore et al. 1995) and this signifies their conservation. PGE 096 related sequences are modulated during evolution resulting in a very low level in sorghum to undetectable levels in the other cereal genomes. The sequences related to PGB 625 are maintained at a very low level in the other cereals while those related to PGB 662 are at a very low level in sorghum and maize and are not detectable in the other cereal genomes.

Repeated sequences usually have very limited distribution among related taxa. The interspersed repetitive sequences isolated from *Thinopyrum elongatum* are of abundance in *Agropyron cristatum, Pseudorogneria spicta* and *Thinopyrum junceiforme* whereas they are of low abundance in *Dasypyrum villosum, Psathyrostachys juncea, Secale cereale* and *Triticum aestivum(* McIntyre et al.1987). The interspersed sequences reported from wheat have the same dispersion patterns in rye and also occur in barley in a different configuration, but are absent in maize and

oats (Szurmak and Dobrzanska 1993). The rye RXX630 family of interspersed sequences are present in wheat but are absent in barley (Mao et al. 1994). The comparable levels of abundance of simple interspersed sequences in the present study indicate that these sequences might have evolved in an ancestral genome common to the cereals and were conserved over evolution.

Other cereal genomic DNA probes did not hybridize to PGB 107 and thus can serve as *Pennisetum* specific probes. Genome specific interspersed sequences have been reported in barley (Sonina et al. 1989), rye (Rogowsky et al. 1991) and *Oryza* species (Zhao et al. 1989; Kiefer-Meyer et al. 1995).

Nature and extent of variation in repetitive DNA sequences.

Bam HI revealed polymorphism for sequences homologous to all the simple interspersed sequences (Figs. 18 & 19). Eco RI revealed variations in five sequences while Pst I and Eco RV were useful for four sequences.

As in *Pennisetum* species, the sequences related to PGE 015 and PGE 123 diverged maximum and PGB 058 diverged very little among the cereals. The relative levels of conservation / divergence of other sequences are comparable among cereals. These further substantiate earlier findings that simple interspersed sequences have different levels of conservation and tolerance to variation. Moreover, there exists a selection pressure on repeated sequences and all the sequences in a genome are not rapidly diverging and variable.

Restriction polymorphism of simple interspersed sequences:

Unlike in *Pennisetum* species, a large proportion of the fragments were intense in cereals. The ancestral configurations are replaced with variants from different lineages during different evolutionary periods. Dover (1982) hypothesized that the molecular drive in genomes ensures progressive homogenization of sequences and fixation of variants in a family of sequences. In closely related *Pennisetum* species, larger

Fig 18. Cereal-wise panels of DNA fragment generated by each enzyme in hybridizing to the set of probes were merged using a C - based computer programme. Fragments hybridizing to each probe are colour coded (given at right). Numbers 1-4 represent Eco **RI**, Pst **I**, Bam HI and Eco RV respectively.

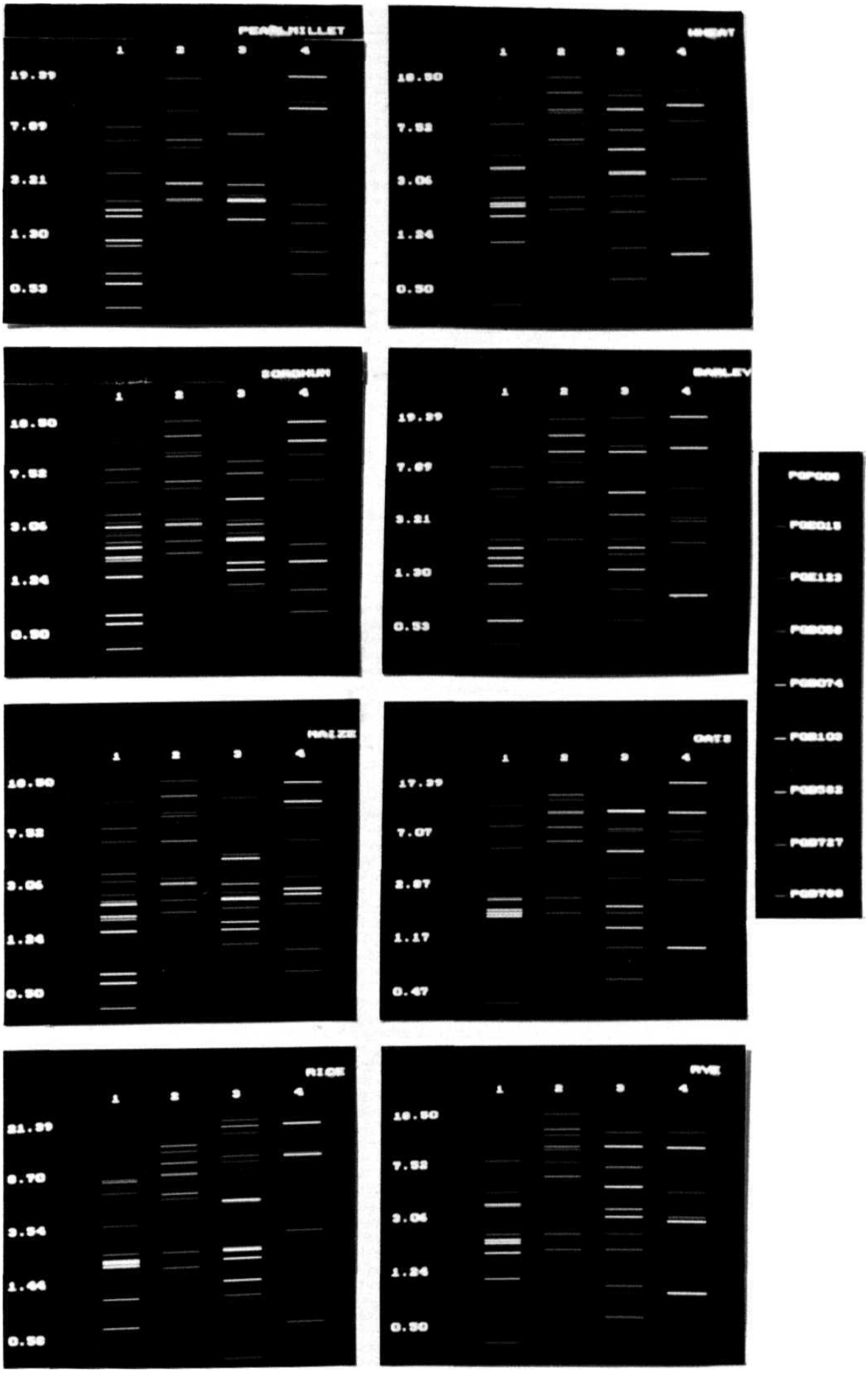
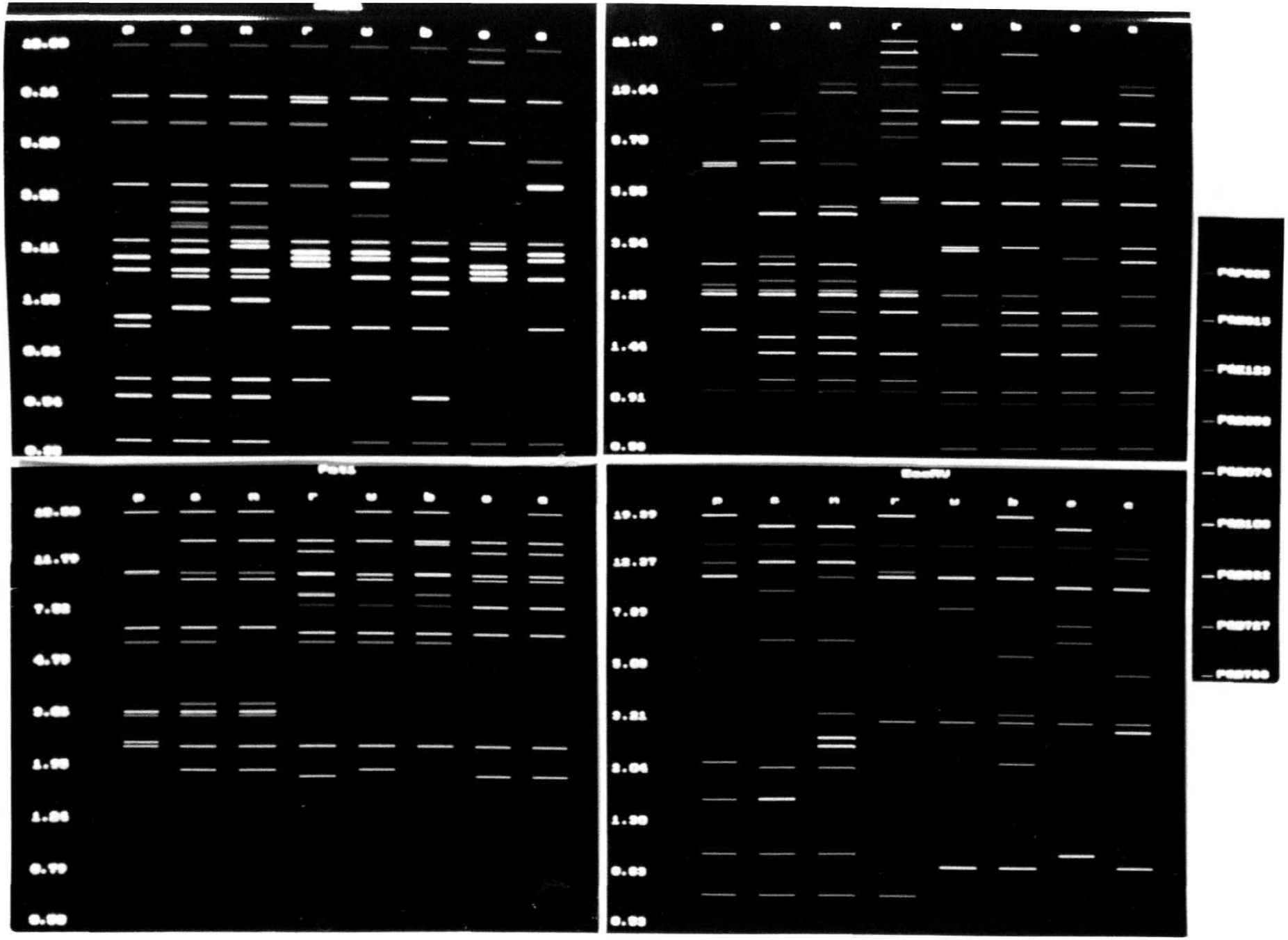

PEARL MILLET
1 2 3 4
19.39
7.69
3.21
1.90
0.53
WHEAT
1 2 3 4
18.90
7.52
3.06
1.24
0.90
SORGHUM
1 2 3 4
18.90
7.52
3.06
1.24
0.90
BARLEY
1 2 3 4
19.39
7.69
3.21
1.90
0.53
PSP008
PSR013
PSR123
PSR098
PSR074
PSR103
PSR582
PSR727
PSR708
MAIZE
1 2 3 4
18.90
7.52
3.06
1.24
0.90
OATS
1 2 3 4
17.39
7.07
2.87
1.17
0.47
RICE
1 2 3 4
21.39
8.70
3.54
1.46
0.98
RYE
1 2 3 4
18.90
7.52
3.06
1.24
0.90

Fig 19. Enzyme-wise panels of DNA fragment generated by each enzyme hybridizing to the set of probes were merged in each species using a C-based computer programme. Fragments hybridizing to each probe are colour coded (given at right). p, s, m, r, w, b, o and c represent pearl millet, sorghum, maize, rice, wheat, barley, oats and rye respectively.



proportion of different repetitive sequences are maintained in the configuration as that of their common ancestral form.

Out of the total number of fragments detected by nine interspersed sequences, 54% were shared by two or more cereals (Table XI) and 45% were species-specific (Table XII). Contrary to the 45% conserved fragments among *Pennisetum* species, only 0.7% were conserved among cereals. Polymorphism observed among *Pennisetum* species shows that variations occurred in the sequence configurations during or after the divergence of the evolutionary intermediate from which *Pennisetum* species have evolved. Comparison of the data on cereals shows the variation occurred in the ancestral configurations in different lineages. Higher level of sequence diversity observed among cereals compared to the diversity among species of *Pennisetum* is expected considering the shared lineages of different species within a genus. Present results are consistent with this interpretation.

Pair-wise distribution of polymorphic fragments (Table XIII) revealed different levels of conservation of fragments among their cereals examined. A closer look at the number of such fragments suggests that pearl millet, sorghum, maize and rice fall into one group while wheat, barley, oats and rye into another. Certain groups of cereals have more number of conserved fragments(non-random). Such groups of cereals are as follows: pearl millet, sorghum and maize (11.1%); sorghum and maize (11.!%); pearl millet, sorghum, maize and rice (10%); wheat, barley, oats and rye (10%); wheat, barley and rye (5.5%) and rice, wheat, barley, oats and rye (4%). Pair-wise distribution of such conserved (non-random) fragments are given in parenthesis preceded by the number of shared fragments (Table XIII). These fragments represent variants of a family of sequences conserved in lineages. Closely related cereals such as wheat and rye, and sorghum and maize have more number of such conserved lineages . It can be concluded that the number of conserved lineages are indicative of their closer phylogenetic affinities . The remaining shared fragments are present in all the cereal pairs and do not show affinity to a particular group of cereals suggesting that these fragments represent random divergence of variants in a family of sequences.

Table XI. Size(s)(in Kb) and distribution of polymorphic fragments among cereals

Probe	Enzyme			
	Eco RI	**Pst I**	**Bam HI**	**Eco RV**
PGP 005	-	3.1-PSM 7.8-RWBORY	7.1-PSMWBRy 14.4-PMRWRy 1.1-PSMR 0.9-WBORy	-
PGE 015	3.7-PSMWR 5.4-BO	9.7-SMWORy 1.9-SMWRy 1.7-RO	1.0-PWBORy 1.1-SMR 11.4-RB 13.9-Wry	0.6-PSMR 2.0/ 1.0-SM 3.1-RWBORy
PGE 123	3.2-/2.5SM 0.4-WBORy 4.7-WBRy 6.4-PSMR	5.5-PSMRWBRy 8.6-RBRy 10.4-PSMR 3.0-PSM 3 .3-SM	7.2-PSWBORy 5.0-RO 3.2-SO 2.7-SM 2.3-WBRy 2.4-PSMR 1.8/ 0.6-WBORy	3.3-MBRy 6.2-MO
PGB 074	-	-	1.9-RBO 3.4-WRy 1.6-SM	-
PGB 103	1.9-RWRy 0.6-PSMR 1.7-PSM 1.6-WBORy	-	4.6-SM 5.1-WBORy	17.4-SMO 19.3-PRB
PGB 582	0.5-PSMB 1.9-PBO 2.1-SMWRy 3.7-WRy 1.8-RO	-	2.3-PSMR 10.3-WBORy	12.4-PSM 11.0-RWBRy 1.0-PSRWBO
PGB 727	-	13.9-SMRWBORy 6.4-PRWBORy 3.1 / 6.1-PSM 10.3-RWBORy	7.1-PSMWBRy 14.4-PMRWRy	-
PGB 788	1.1-PRWBRy 0.6-PR 0.3-PSM 1.7-SM	-	3.0-PSM 3.5-WBRy	-

* P - Pearl millet, S - Sorghum, M - Maize, R - Rice, W - Wheat, B - Barley, O - Oats, Ry - Rye

Table XII. Species - specific fragments among cereals* detectable with the pearl millet repetitive DNA probes

Probe	Fragments(Kb)			
	Eco RI	Pst I	Bam HI	Eco RV
PGP 005	S- 2.7 W-2.9 O-2.1	S-11.O	- R-9.0 O-5.3	
PGE 015	R-7.7	P-10.4 / 2.4 R-8.5 B-13.5	P-7.0 S-8.7 M-4.9 / 2.0 M-21.4 O-7.6 B-18.9	P-2.1 / 1.4 / 0.9
PGE 123	R-3.7 O-11.1	O-7.7	M-13.2 P-3.0/2.5 R-1.4 / 1.1	P-12.3 S- 9.6 M-0.9 / 3.3 R-11.5 W-8.2 B-5.5 O-7.1 Ry-4.9
PGB 074	-	-	P-1.7	-
PGB 103	-	-	P-0.9 R-5.3	S-1.6 S-2.5 W-3.1 Ry-2.9
PGB582	P-1.2 S-1.3 M-1.4 R-2.2 B-1.5	-	-	O-10.1 M-2.7 Ry-0.9
PGB 727	-	-	S-11.0 R-9.0 O-5.3 / 2.0	-
PGB 788	O-2.2	-	R-19.1 O-10.4	-

*P- Pearl millet, S - Sorghum, M - Maize, R - Rice, W - Wheat, B- Barley, O - Oats, Ry - Rye

Table XIII. Pair-wise distribution of shared fragments among cereals

	Sorghum (n=10)	Maize (n=10)	Rice (n=12)	Wheat (n=21)	Barley (n=7)	Oats (n=21)	Rye (n=7)
Pearl millet(n=7)	22(15)	22(15)	15(7)	9(0)	9(0)	3(0)	8(0)
Sorghum (n=10)		36 (24)	12 (7)	10 (0)	7(0)	7(0)	9(0)
Maize(n=10)			13(7)	10(0)	6(0)	4(0)	11(0)
Rice(n=12)				11(3)	12(3)	9(3)	11(3)
Wheat(n=21)					21(14)	14(10)	29(14)
Barley(n=7)						16(10)	22(14)
Oats (n=21)							13(10)

Conserved fragments are in parenthesis(see text for details)

Table XIV. Pair-wise similarity (%) among cereal genomes in their repetitive DNA homologous to cloned pearl millet sequences

	Sorghum	Maize	Rice	Wheat	Barley	Oats	Rye
Pearl millet	51	55	37	21	25	6	22
Sorghum		80	32	22	17	15	20
Maize			33	23	18	14	26
Rice				28	34	25	29
Wheat					59	43	86
Barley						52	64
Oats							42

Results from the present study indicate that some variants in a family of sequences are conserved in certain lineages while other variants tolerate variations and diverge rapidly.

Fragments of sequences related to PGB 625 and PGB 662, which lighted up after prolonged exposure, represent their maintenance as single / low copy in these cereal genomes. Further characterization of PGB 625 and PGB 662 related sequences in the cereals along with their flanking regions may provide more insight in understanding the mechanism of "saltatory" amplification (Britten and Kohne 1968).

Phylogenetic analysis

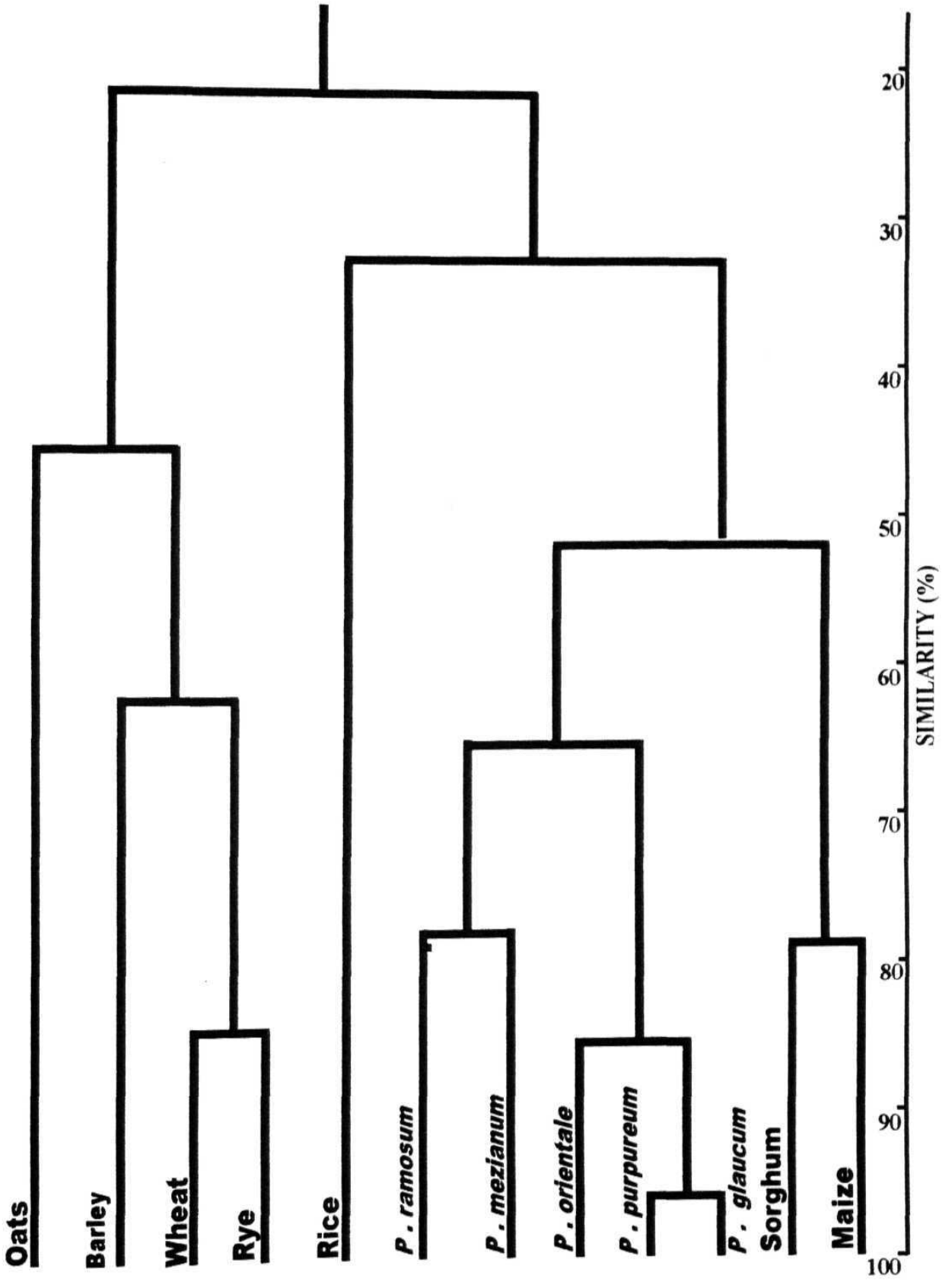
Sequence similarity index was estimated among cereals in pairs from the restriction polymorphisms (Table XIV). Wheat and rye have 86% similarity and both have an average of 61% similarity with barley. These three cereals as a group have nearly 45% similarity with oats. Sorghum and maize have 80% similarity for the sequences examined and these two cereals have - 52% similarity with pearl millet which in turn have ~35% stem common with rice. Present data on the divergence of conserved repetitive sequences thus differentiate tropical from temperate cereals with ~20% common stem.

Phylogenetic tree of the cereals (Fig 20) was constructed based on the mean similarity index (Table XIV) using PHYLIP programme.

Phylogenetic relationship based on the changes occurring in simple interspersed sequences shows that the first divergence of the ancestral form of these cereals led to forms, one of which was common to the tropical cereals (Paniceae and Andropogoneae), while another was common to temperate cereals(Triticeae and Avenae) in the present context. Oats diverged very early in the evolution of the temperate group of cereals, followed by barley and finally wheat from rye. Rice diverged much earlier in the evolution of tropical cereals, followed by pearl millet and finally sorghum from maize.

Fig 20. Phylogeny of cereals

PHYLOGENY OF CEREALS



Phylogenetic relationship among temperate cereals (*A venae* and *Triticeae)* and those of tropical cereals belonging to Andropogoneae, Paniceae and *Oryzae,* as revealed by variations in repetitive sequences is consistent with their cytotaxonomical and cytogenetic relationships. Within Paniceae, *Pennisetum* species have >60% common stem (Fig 20).

Phylogenetic relationships based on the divergence path of conserved repetitive sequences are consistent with the relationships of temperate cereals reported by Flavell et al. 1977. This parallelism between the two trees suggests that the origin of new classes of repetitive sequences and the divergence of conserved sequences are governed by evolutionary time periods.

Repeated sequences serve as vital tools in studying phylogenetic relationships (Dvorak et al. 1988) since the phylogenetic trees based on the repetitive DNA divergence are consistent with the taxonomic grouping of the species (Gupta et al. 1989; Dvorak and Zhang 1992; Ferrer et al. 1995). The estimation of phylogenetic relationships among distantly related taxa based on repetitive DNA divergence has not been reported earlier. Results from the present study demonstrate the applicability of cloned repetitive sequences in elucidating phylogenetic relationships of distantly related taxa irrespective of their genome sizes or repeated DNA contents .

SUMMARY AND CONCLUSIONS

Sequences belonging to 12 families of repetitive sequences were isolated from genomic libraries of pearl millet following colony hybridization and were characterized. Distribution, modulation and divergence of these repetitive sequences were examined among *Pennisetum* species differing in their basic chromosome number (P. *ramosum* (x=5), P. *mezianum* (x=8), P. *orientale*(x=9) and an allotetraploid P. *purpureum* (x=7) with one genome common to P. *glaucum (x=7))* and related cereals (sorghum, maize, rice, wheat, barley, oats and rye). PGP 005, a Pst I derived repetitive sequence used in the present study, had partial homology with PGB 727.

Abundance of these sequences varied from 0.6 x 10^3 (PGB 727) to 1.0 x 10^4 (PGB 625) and together constituted ~3.2% of the pearl millet genome. PGB 625 is organized as tandem repeats and all the other sequences are interspersed in pearl millet genome. Three interspersed sequences, PGE 096, PGB 107 and PGB 662 have very complex organization (complex interspersed sequences) distributed at diverse sequence environments. The other interspersed sequences are relatively less dispersed (simple interspersed sequences) in the genome. Restriction analysis revealed that the sizes of basic repeat unit of the complex interspersed sequences, PGB 107 (0.8 Kb), PGB 625 (1.4 Kb) and PGB 662 (1.9 Kb) are 500 bp, 127 bp and 600 bp respectively. Preliminary investigations suggest the sub-repeat nature for the basic repeat unit (cloned fragment) of the simple interspersed sequences.

Although the abundance of simple interspersed sequences varies, they do not show an order of magnitude differences among *Pennisetum* species and cereals while tandem repeat and complex interspersed sequences are extensively modulated. Repeated sequences reported from cereals such as barley (Sonina et al. 1989; Vershinin et al. 1990; Belostotsky and Ananiev 1990; Hueros et al. 1990), wheat (Dobrzanska and Szurmak 1992; Szurmak and Dobrzanska 1993, rye (Bedbrook et al. 1980a; Xin and Appels 1987; Mao et al. 1994), *Aegilops squarrosa* (Nagaki et al. 1995), *Thinopyrumelongatum* (Mc lyntyre et al. 1987), oats (Fabijanski et al. 1990; Grebenstein et al. 1995), rice (Wu and Wu 1992; Mawal et al. 1995; Li et al. 1996) have different

levels of modulation. The present study demonstrates the occurrence of conserved repetitive sequences among cereal genomes.

Minimum chromosome number of genus *Pennisetum* is 5 and is represented in P. *ramosum*. Species with basic chromosome numbers 7, 8 and 9 are secondary polyploids evolved from an ancestor through duplication of specific chromosomes. Differential abundance of some of the simple interspersed sequences among *Pennisetum* species differing in their basic chromosome number is supportive of their secondary polyploid nature. These sequences may be present on different chromosomes (segments) in different doses in different species (genomes). Further experiments such as fluorescence *in situ* hybridization using suitable aneuploid lines are required to assign these sequences onto specific chromosome(s). The high relative abundance of simple interspersed sequences in rice with small a genome size and their comparatively low abundance in other cereals with large genome size is indicative of the dilution with effect of the increase in their genome size. This implies that the abundance of these sequences is maintained among cereals in spite of the tremendous range of variation in their genomic DNA content (Bennett et al. 1982) and the apparent variation in their abundance could be attributed to differential amplification or the dilution effect with increase in genome size.

PGB 625 is present at comparable level in P. *purpureum*, while it is at a very low level in other *Pennisetum* species and cereal genomes. Sequences homologous to PGB 662 are at low levels in all *Pennisetum* species and cereals. Kamm et al. (1994) reported that sequences related to the Hae III family of pearl millet were not detected in fox tail millet, wheat, rye and barley. Sequences related to PGB 625 are detected in all the *Pennisetum* species and cereals studied and are likely to belong to the Hae III family. It is apparent that these two sequences amplified in pearl millet ancestor from a low copy sequence which remained as a low copy sequence in other lineages. Distribution aspects of the highly repetitive sequences across distantly related taxa, as carried out during the present study, are not addressed in many of the reports. The mechanism involved in the sudden amplification, "saltatory" amplification (Britten and

Kohne 1968) of sequences, are not known. Further characterization of these sequences along with the homologous sequences and their flanking regions in other cereals would be vital in understanding the features of sequences which undergo "saltatory" amplification.

PGE 096 and PGB 662 are deamplified while the abundance and organization of PGB 625 is conserved in P. *purpureum,* which has one genome common with *P. glaucum,* following interspecific hybridization and polyploidization. McClintock (1978) first suggested that remote hybridization, which places the parental genome in an unusual genomic environment, is a form of stress. Further analysis of pearl millet repetitive sequences in *P. purpureum* might reveal the basis of selective deamplification during speciation.

PGB 107, which is genus specific and highly variable among the *Pennisetum* species examined, is an ideal marker for genome analysis and genome typing in *Pennisetum.* However, other *Pennisetum* species also have to be examined for its wider application. Although the PGE 096 sequence is specific to *P. glaucum* and *P. ramosum,* it is not useful as a marker since it does not reveal discrete bands with the enzymes tested.

PGE 123 and PGE 015 related sequences are relatively more diverged while PGB 058 related sequences are highly conserved among *Pennisetum* species and related cereals. Other sequences show intermediate levels in their divergence. The results suggest that the nature and extent of changes are not uniform among the repeated sequences studied, which in turn may suggest that some of these conserved sequences are under selection pressure.

Since each family of simple interspersed sequences is at conserved abundance across the cereals, the term 'family' may be referred to all cross hybridizing sequences from different cereals and therefore the sequences from each of the cereals represent a sub-family. On the assumption that all the members of a particular repeated sequence family were initially derived from a common ancestral sequence and transmitted through different lineages, the characterization of the structural and organizational differences

between sub-families documents the types of molecular changes which can occur and accumulate in the chromosomal DNA. Comparison of the nature and extent of divergence of simple interspersed sequences within *Pennisetum* species and between cereals thus could contribute to our understanding of the evolution of repeated sequences.

Conservation of simple interspersed sequences (100%) in pearl millet accessions, 45% among *Pennisetum* species and 0.7% among cereals indicates that these sequences undergo slow evolutionary turnover despite their conserved abundance. Shared fragments and species-specific fragments among *Pennisetum* species in addition to the conserved fragments indicate that a small proportion of a family of sequences has undergone variation. On the other hand a large proportion of species-specific and shared fragments are intense and no fragments are conserved across the cereals. When these two observations are taken together (along with the fact that species of a genus are more closely related than different tribes of a family), it can be concluded that divergence begins in a small proportion of sequences and the existing configurations are slowly replaced with the variants as evolution proceeds. Further, the species divergence is parallel with the extent of fixation of variants. This supports the idea that repetitive sequences evolve by fixation of new variants, successive divergence and by concerted evolution (Bedbrook et al. 1980a; Dover 1982; Grellet et al. 1986; De kochko et al. 1991). Polymorphism observed among cereals provides additional information about the kinds of variants in a family of sequences. The species-specific fragments and the fragments which are found across some of the tropical and temperate cereals represent random divergence of some variants. Other variants are conserved only in a group of cereals and the number of such conserved variants in the group was found to be parallel with their phylogenetic affinities. It can be concluded from the data that phylogenetic distance increased with increased divergence in the conserved (variants) sequences. Investigators in this field have not looked into this aspect of the conservation of variants and their role in phylogenetic relationship due to the lack of cloned sequences which are conserved across distant taxa.

Sequence divergence patterns revealed by the restriction polymorphism were used to estimate the pair-wise similarity index. The high similarity found between P. *glaucum* and P. *purpureum* is expected since P. *purpureum* has one genome common with P. *glaucum.* The present study revealed that P. *glaucum* is more related to P. *orientate* than to P. *ramosum* and P. *mezianum*.

The temperate cereals are grouped under the sub-family Festucoideae of Gramineae family. Oats comes under the tribe Aveneae and barley, wheat and rye are grouped in the tribe Triticeae. Wheat and rye are further grouped into the sub-tribe Triticineae. Pearl millet, sorghum and maize are grouped under sub-family Panicoideae. Pearl millet belongs to the tribe Paniceae and the other two are under the tribe Andropogoneae. Rice is grouped under the sub-family Oryzoideae and tribe Oryzeae (Langer and Hill 1991). Similarities of repeated DNA sequences revealed by the molecular analysis in the present study are consistent with aforesaid taxonomical classification of cereals. However, rice has more affinity towards cereals belonging to the sub-family Panicoideae with 35% similarity.

Information from the comparative mapping programme involving rice revealed conservation of low copy sequences which are by and large genic (Moore et al. 1995). This suggests that there is a basic line of essential functional genes among species and the changes in genome size in related species are due to variations in the amount of repeatitive sequences (Sagai-Maroof et al. 1996). From these observations, one would be tempted to attribute a role to the repetitive sequence divergence in speciation. This concept may further strengthen the occurrence of species-specific and genome-specific sequences. The conserved abundance of simple interspersed sequences across cereals and their slow evolutionary turn over, parallel with species divergence, suggests the significance of selection pressure in speciation. Dover (1982) proposed that the concerted pattern of fixation of variants and the replacement of existing members of a family of sequences with a single new variant member by molecular drive could be the probable explanation for the origin of species discontinuities and biological novelty.

New variants in the families of conserved repetitive sequences slowly replace the ancestral configuration and the phylogenetic distance between species increases as the fixation of new variants progresses.

Different members of a family of repetitive sequences have different rates of divergence during evolution.

Phylogenetic tree constructed at the intertribal level on based repetitive sequence / similarity is consistent with the taxonomical classification.

The simple interspersed sequences developed in the present study are 'universal cereal probes' ideal for phylogenetic analysis of cereals and for use as molecular markers.

REFERENCES

Ahn, S.N. and Tanksley, S.D. 1993. Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci USA 90: 7980-7984.

Appels, R., Moran, L.B. and Gustafson, J.P. 1986. Rye heterochromatin I. Studies on clusters of the major repeating sequence and the identification of a new dispersed repetitive sequence element. Can J Genet Cytol 28:645-657.

Bedbrook, JR., Jones, J., O'Dell, M., Thompson, R.D. and Flavell, R.B. 1980a. A molecular description of telomeric heterochromatin in *Secale* species. Cell 19: 545-560.

Bedbrook, JR., O' Dell, M., and Flavell, R.B. 1980b. Amplification of rearranged repeated DNA sequences in cereal plants. Nature 288:133-137.

Belostotsky, D. A. and Ananiev, E.V. 1990. Characterization of relic DNA from barley genome. Theor Appl Genet 80: 374-380.

Bennett, M. D., and Smith, J.B. 1976 Nuclear DNA amounts in angiosperms. Philos Trans R Soc Lon Ser B 274: 227-274.

Bennett, M. D. Smith, J.B. and Heslop-Harrison, J.S. 1982. Nuclear DNA amounts in angiosperms. Philos Trans R Soc Lon B 216: 179-182.

Berlani, R. E., Davis, R.W. and Walbot, V. 1988a. Genomic organization of two families of highly repeated nuclear DNA sequences of maize selected for autonomous replicating activity in yeast. Plant Mol Biol 11: 161-172.

Berlani, R. E., Walbot, V. and Davis, R.W. 1988b. Sequence analysis of three fragments of maize nuclear DNA which replicate autonomously in yeast. Plant Mol Biol 11: 173-182.

Bouchard, R.A. 1982. Moderately repetitive DNA in evolution. Int Rev Cytol 76: 113-193.

Bournival, B., Obanni, M., Abad, A., Ohm, H. and Mackenzie, S. 1994. Isolation of a new species - specific repetitive sequence from *Thinopyrum elongatum* and its use in the studies of alien translocations. Genome 37: 97-104.

Britten, R.J. and Kohne, D.E. 1968. Repeated sequences in DNA. Science 161: 529-540.

Chooi, W. Y. 1971. Comparison of the DNA of six *Vicia* species by the method of DNA-DNA hybridization. Genetics, 68: 213-230

De Kochko, A., Kiefer, M.C., Cordesse, F., Reddy, A.S and Delseny, M. 1991. Distribution and organization of a tandemly repeated 352bp sequence in the *Oryzae* family. Theor Appl Genet 82: 57-64.

Deshpande, V.G. and Ranjekar, P.K. 1980. Repetitive DNA in three *Gramineae* species with low **DNA** content. Hoppe-reyleris Z Physiol Chem 361 :S 8 1223-1233.

Devos, K.M., Millan, T. and Gale, M. 1993. Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. Theor Appl Genet 85: 784-792.

Dobrzanska, M. and Szurmak, B. 1992. A moderately repeated DNA sequence of wheat and rye genomes. Plant Mol Biol 18: 603-605.

Dover, G. A. 1982. Molecular drive: a cohesive mode of species evolution. Nature 299: 111-117.

Dvorak, J., McGuire, P. E. and Cassidy. B. 1988. Apparent source of the `A' genomes of wheat inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. Genome 30: 680-689.

Dvorak, J. and Zhang, H.-B. 1990. Variation in repeated nucleotide sequences sheds light on the phytogeny of the wheat 'B' and 'G' genomes. Proc Natl Acad Sci USA 87: 9640-9644.

Dvorak, J. and Zhang, H.-B. 1992. Reconstruction of the phylogeny of the genus *Triticum* from variation in repeated nucleiotide sequences. Theor App Gent 84: 419-429.

Evans, I.J , James, A.M. and Barnes, S R 1983. Organization and evolution of repeated DNA sequences in closely related plant genomes. J Mol Biol 170: 803-826.

Fabijanski, S., Fedak, G., Armstrong, K. and Altosaar, I. 1990. A repeated sequence probe for the 'C' genome in *Avena* (oats) Theor Appl Genet 79: 1-7.

Ferrer, E., Loarce, Y. and Hueros, G. 1995. Molecular characterization and chromosome location of repeated DNA sequences in *Hordeum* species and in the amphiploid tritordeum (X *Tritordeum* Ascherson et Graebner). Genome 38: 850-857.

Flamm, W.G. 1972. Highly repetitive sequences of **DNA** in chromosomes. Int Rev Cytol 32: 2- 49

Flavell, R. 1980. The molecular characterization and organization of plant chromosomal DNA sequences. Ann Rev Plant Physiol 31: 569-96.

Flavell, R.B., Bennett, M.D , Smith, J.B. and Smith, D.B. 1974 Genome size and the proportion of repeated nucleotide sequence DNA in plants Biochem Genet 12: 257-269.

Flavell, R.B., Rimpau, J. and Smith, D.B 1977. Repeated sequence DNA relationships in four cereal genomes. Chromosoma 63: 205-222.

Francis, H.A., Leitch, A. R. and Koebner, R.M.D. 1995. Conversion of a RAPD-generated PCR product, containing a novel dispersed repetitive element, into a fast and robust assay for the presence of rye chromatin in wheat. Theor Appl Genet 90: 636-642.

Grebenstein, B., Grebenstein, 0., Saner, W. and Hemleben, V. 1995. Characterization of a highly repeated DNA component of perennial oats *(Helictotrichon,* Poaceae) with sequence similarity to a A-genome-specific satellite DNA of rice *(Oryza).* Theor Appl Genet 90: 1101-1105.

Grellet, F., Delcasso, D., Panabieres, F. and Delseny, M. 1986. Organization and evolution of a higher plant aphoid like satellite DNA sequence. J Mol Biol 187: 495-507.

Guidet, F., Rogowsky, P., Taylor, C, Song, W. and Langridge, P. 1991. Cloning and characterization of a new rye-specific repeated sequence. Genome 34: 81-87.

Gupta, P.K., Fedak, G., Molnar, S.J. and Wheatcroft, R. 1989. Distribution of *Secale cereale* DNA repeat sequence among 25 *Hordeum* species Genome 32: 383-388.

Harcourt, R.L. and Gale, M.D. 1991. A chromosome-specific DNA sequence which reveals a high level of RFLP in wheat. Theor Appl Genet 81: 397-400.

Helentjaris, T. 1993. Implications for conserved genomic structure among plant species. Proc Natl Acad Sci USA 90: 8308-8309.

Hoang-Tang, Dube, S. K., Liang, G.H. and Kung, S.-D. 1991. Possible repetitive DNA markers for *Eusorghum* and *Parasorghum* and their potential use in examining phylogenetic hypotheses on the origin of *Sorghum* species. Genome 34: 241-250.

Hueros, G., Loarce, Y. and Ferrer, E. 1993. A structural and evolutionary analysis of a dispersed repetitive sequence. Plant Mol Biol 22: 635-643

Hueros, G., Monte, J V and Ferrer, E. 1990. *Hordeum chilense* repetitive sequences: Genome characterization using biotinylated probes. Theor Appl Genet 80: 24-32.

Hulbert, S.H., Richter, T.E., Axtell, J D. and Bennetzen, J.L. 1990. Genetic mapping and characterization of *Sorghum* and related crops by means of maize DNA probes. Proc Natl Acad Sci USA 87: 4251-4255.

Huttley, G. A., MacRae, A F. and Clegg, M. T. 1995. Molecular evolution of the *Ac/Ds* transposable-element family in pearlmillet and other grasses. Genetics 139:1411-1419.

Ingham, L.D., Hannah, W. W., Baier, J.W. and Hannah, L.C. 1993. Origin of the main class of repetitive DNA within selected *Pennisetum* species. Mol Gen Genet 238:350-356.

Jones, E.S., Liu, C. J., Gale, M. D., Hash, C.T. and Witcombe, JR. 1995. Mapping quantitative trait loci for downy mildew resistance in pearl millet. Theor Appl Genet 91:448-456.

Jones, J. D. G and Flavell, R. B. 1982b. The structure, amount and chromosomal localization of defined repeated DNA sequences in the species of genus *Secale.* Chromosoma 86: 613-641

Kamm, A., Galasso, I., Schmidt, T. and Heslop-Harrison, J.S. 1995. Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationship between *Arabidopsis* species. Plant Mol Biol 27: 853-862.

Kamm, A., Schmidt, T and Heslop-Harrison, J.S. 1994. Molecular and physical organization of highly repetitive undermethylated DNA from *Pennisetum glaucum.* Mol Gen Genet 244: 420-425.

Kiefer-Meyer, M.C., Reddy, A.S. and Delseny, M. 1995. Characterization of a dispersed repetitive **DNA** sequence associated with the CCDD genome of wild rice. Genome 38: 681-688.

Kim, N.-S., Whelan, E.D.P., Fedak, G. and Armstong, K. 1992. Identification of a *Triticum - Lophopyrum* noncompensating translocation line and detection of *Lophopyrum* DNA using a wheatgrass specific molecular probe. Genome 35: 541-544.

Kumar, L. S., Gupta, V.S. and Ranjekar, P.K. 1990. Identification and partial characterization of two species-specific repeat families in the great millet *(Sorghum vulgare, Poaceae)* genome. PI Syst Evol 171: 249-257.

Kurata, N., Nagamura, Y., Yamamoto, K., Harushima, Y., Sue, N., Wu, J., Antonio, B. A., Shomura, A., Shimizu, T., Lin, S. Y., Inoue, T., Fukuda, A., Shimano, T., Kuboki, Y., Toyama, T., Miyamoto, Y., Kirihara, T., Hayasaka, K., Miyao, A., Monna, L, Zhong, H. S , Tamura, Y , Wang, Z. X., Momma, T., Umehara, Y., Yano, M., Sasaki, T. and Minobe, Y. 1994a. A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genetics 8: 365-376.

Kurata, N., Moore, G , Nagamura, Y., Foote, T., Yano, M., Minobe, Y. and Gale, M. 1994b. Conservation of genome structure between rice and wheat. Bio/Technol 12: 276-278.

Langer, R. H. M. and Hill, G. D. 1991. Agricultural plants. Cambridge University Press. pp 40- 139.

Lapitan, N.L.V. 1992. Organization and evolution of higher plant nuclear genomes. Genome 35: 171-181.

Li, X-B , Liang, C.-Z., Wu, H.-G., Zhai, W.-X., Huang, N. and Zhu, L.-H. 1996. Isolation and identification of non-specific tandemly repeated DNA sequence in *Oryza* species. Theor Appl Genet 92: 702-708.

Liu, C. J., Witcombe, JR., Pittaway, T.S., Nash, M., Hash, C. T., Busso, C. S. and Gale, M.D. 1994. An RFLP-based genetic map of pearl millet *(Pennisetum glaucum).* Theor Appl Genet 89: 481-487.

Liu, Y.-G., Ikeda, T.M. and Tsunewaki, K. 1992 Moderately repeated, dispersed, and highly variable (MRDHV) genomic sequences of comon wheat usable for cultivar identification. Theor Appl Genet 84: 535-543.

Mao, L., Zhai, W., Hu, H. and Zhu, L. 1994. Cloning and characterization of a repetitive sequence in rye *(Secale cereale).* Plant Science 100: 51-57.

Martinez-Zaparter, J., Estelle, M. and Sommerville, C. 1986. A highly repeated DNA sequences in *Arabidopsis thaliana.* Mol Gen Genet 204: 417-423.

Mawal, Y., Lagu., M., Moon, E., Change, S., Chung, M.-C, Wu, H.-K., Gupta, V., Ranjekar, P. and Wu. R. 1995. Bam HI and Hind III repetitive DNA families in the rice genome. Genome 38:191-200.

McClintock, B. 1978. Mechanism that rapidly reorganize the genome Stadler Genet Sym 10:25-48.

McIntyre, C. L, Clarke, B. C. and Appels, R. 1987. Amplification and dispersion of repeated DNA sequences in the *Triticeae.* PI Syst Evol 160: 39-59.

McNeil, D., Lagudah, E.S., Hohmann, U. and Appels, R. 1994. Amplification of DNA sequences in wheat and its relatives: the Dgas 44 and R 350 families of repetitive sequences. Genome 37: 320-327.

Metzlaff. M , Troebner, W., Baldauf, F , Schlegel, R. and Cullum, J. 1986. Wheat specific repetitive DNA sequences - construction and characterizaton of four different genomic clones. Theor Appl Genet 72: 207-210.

Miksche, J. P. and Hotta, Y. 1973. DNA base composition and repetitive DNA in several conifers. Chromosoma 41: 29-36.

Moore, G., Devos, K. M., Wang, Z. and Gale, M. D. 1995. Grasses, line up and form a circle. Current Biology 5: 737-739.

Nagaki, K, Tsujimoto, H., Isono, K. and Sasakuma, T. 1995. Molecular characterization of a tandem repeat, Afa family, and its distribution among Triticeae. Genome 38: 479-486.

Peacock, W.J., Dennis, E.S., Rhoades, M.M. and Pryor, A.J. 1981. Highly repeated DNA sequence limited to knob heterochromatin in maize. Proc Natl Acad Sci USA 78: 4490-4494.

Pecchioni, N., Stanca, A.M., Terzi, V. and Cattivelli, L. 1993. RFLP analysis of highly polymorphic loci in barley. Theor Appl Genet 85: 926-930.

Perez-Vicente, R., Petris, L., Osusky, M., Potrykus, I. And Spangenberg, G. 1992. Molecular and cytogenetic characterization of repetitive DNA sequences from *Lolium* and *Festuca*: applications in the analysis of *Festulolium* hybrids. Theor Appl Genet 84: 145-154.

Rayburn, A. L. and Gill, B. S. 1986. Repeated DNA sequences in *Triticum* (*Poaceae*): chromosomal mapping and its bearing on the evolution of B and G genomes. PI Syst Evol 159:229-235.

Rogowsky, P.M., Manning, S., Liu, J.-Y. and Langridge, P. 1991. The R173 farmily of rye specific repetitive **DNA** sequences: a structural analysis. Genome 34: 88-95.

Saghai-Maroof, M. A., Yang, G.P., Biyashev, R.M , Maughan, P.J. and Zhang, Q. 1996. Analysis of the barley and rice genomes by comparative RFLP linkage mapping. Theor Appl Genet 92: 541-551.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning, A laboratory Manual. Cold Spring Harbour Laboratory Press. Cold Spring Harbour, NY.

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

Smith, D.B. and Flavell, R.B. 1974. The relatedness and evolution of repeated nucleotide sequences in the genomes of Gramineae species. Biochem Genet 12: 243-256.

Sonina, N. V., Lushnikova, A. A., Tihonov, A P. and Ananiev, E.V 1989. Dialect-I, species-specific repeated DNA sequence from barley, *Hordeum vulgare.* Theor Appl Genet 78: 589-593.

Southern, EM. 1970. Base sequence and evolution of guinea-pig a -satellite DNA. Nature 227: 794- 798

Svitashev, S. K., Vershinin, A. V., Trunova, S. A., Pershina, L. A. and Shumny, V. K. 1995. Molecular analysis of the genomes of wide hybrids in cereals. Herditas 122: 25-31

Szurmak. B. and Dobrzanska. M 1993. A large repeat of the dispersion pattern common to wheat and rye genomes. Plant Mol Biol 21: 919-921.

Talbert, L E., Kimber, G , Magyar, G.M. and Buchanan, C. B. 1993. Repetitive DNA variation and pivotal-differential evolution of wild wheats. Genome 36: 14-20.

Van Deynze, A. E., Nelson, J.C., O'Donoughue, L. S., Ahn, S.N., Siripoonwiwat, W., Harrington, S. E., Yglesias, E.S., Braga, D. P., McCouch, S. R. and Sorrells, ME. 1995b. Comparative mapping in grasses. Oat relationships. Mol Gen Genet 249: 349-356.

Van Deynze, A. E., Nelson, J.C., Yglesias, E.S., Harrington, S. E., Braga, D. P., Mc Couch, S. R. and Sorrells, ME. 1995a. Comparative mapping in grasses. Wheat relationships. Mol Gen Genet 248: 744-754.

Vershinin, A.V., Salina, E. A. and Svitashev, S.K. 1992. Is there a connection between genomic changes and wide hybridization? Hereditas 116: 213-217.

Vershinin, A.V., Salina, E.A., Solovyov, V.V. and Timofeyeva, L.L. 1990. Genomic organization, evolution and structural peculiarities of highly repetitive DNA of *Hordeum vulgare.* Genome 33: 441-449.

Wang, Z.X., Kurata, N., Saji, S., Katayose, Y. and Minobe, Y. 1995. A chromosome 5-specific repetitive DNA sequence in rice *(Oryza sativa L.)*. Theor Appl Genet 90: 907-913.

Whitkus, R., Doebley, J. and Lee. M. 1992. Comparative genome mapping of sorghum and maize. Genetics 132: 1119-1130.

Wu, H.-K., Chung, M-C, Wu, T., Ning, C-.N. and Wu. R. 1991. Localization of specific repetitive DNA sequences in individual rice chromosomes. Chromosoma 100: 330-338.

Wu, JR., Hum, J. and Bonner, J. 1972. Size and distribution of the repetitive segments of the Drosophila genome. J Mol Biol 64: 211- 219

Wu, T. and Wu, R. 1992. A novel repetitive DNA sequence in the genus *Oryza.* Theor Appl Genet 84:136-144.

Xin, Z.-Y. and Appels, R. 1988. Occurrence of rye *(Secale cereale)* 350-family DNA sequences in *Agropyron* and other *Triticeae*. PI Syst Evol 160:65-76.

Zhang, H.-B. and Dvorak, J. 1991. The genome origin of tetraploid species of *Leymus (Poaceae - Triticeae)* inferred from variation in repeated nucleotide sequences. Amer J Bot 78:871-884

Zhang, H.-B. and Dvorak, J. 1992. The genome origin and evolution of hexaploid *Triticum crassum* and *Triticum syriacum* determined from variation in repeated nucleotide sequences. Genome 35: 806-814.

Zhang-, H.-B., Dvorak, J. and Waines, J. G.1992. Diploid ancestry and the evolution of *Triticum kotschyi* and *T. peregrinum* examined using variation in repeated nucleotide sequences. Genome 35: 182-191

Zhao, X., Wu, T., Xie, Y. and Wu, R. 1989. Genome-specific repetitive sequences in the genus *Oryza.* Theor Appl Genet 78: 201-209.