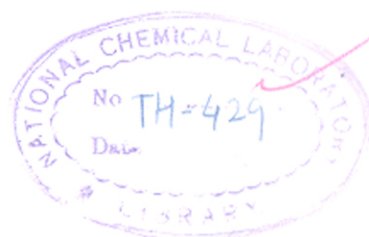


*Genome Characterization in Plants
with Special Reference to
Four Millet Species*

*A Thesis submitted to the
University of Poona
for the degree of
Doctor of Philosophy
(in Chemistry)*

COMPUTERISED

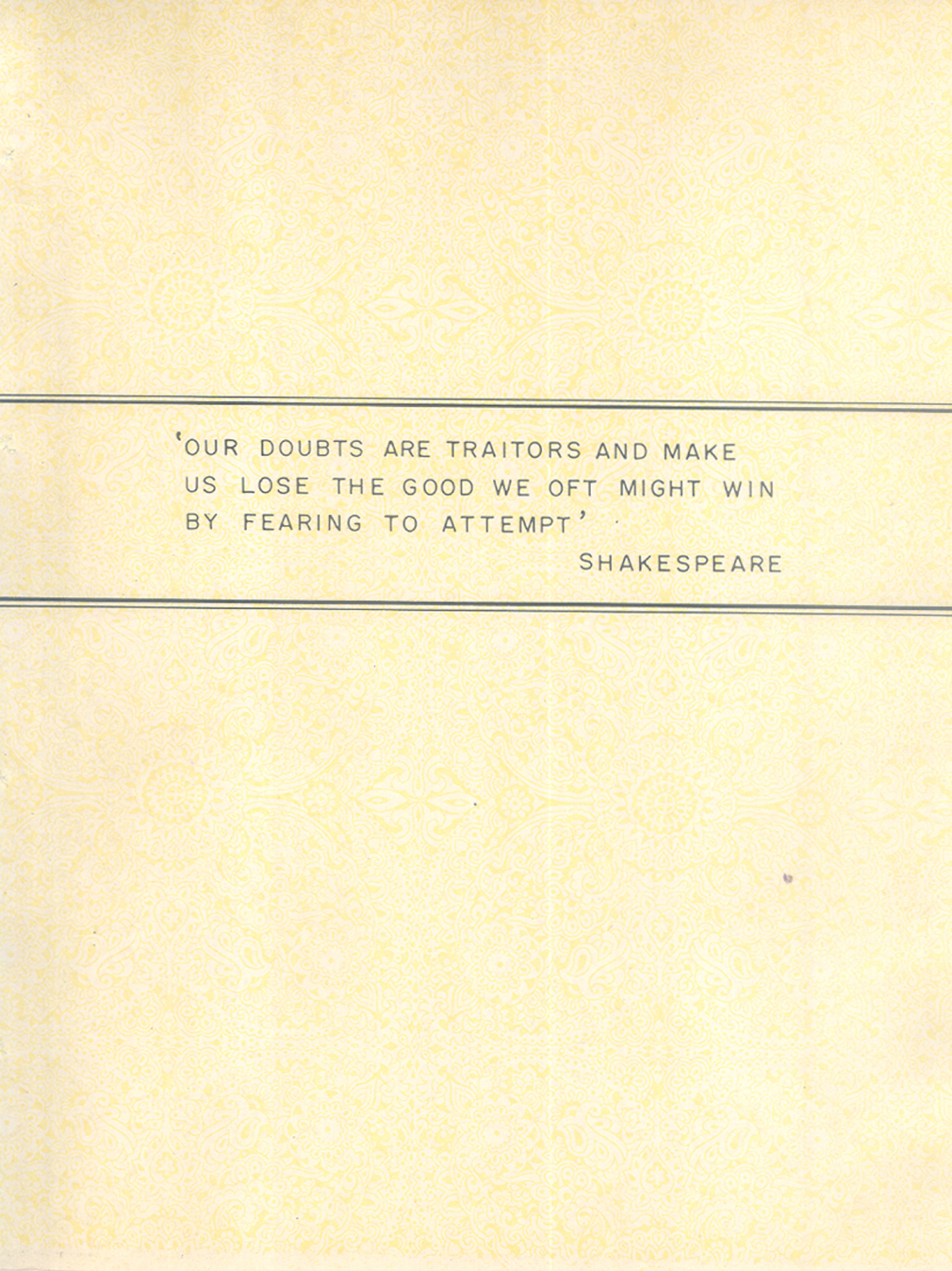


Miss Lakshmi SivaRaman, M. Sc.

581.61(4)
CIN

*Division of Biochemistry
National Chemical Laboratory
Pune - 411 008 (India)*

October, 1984



'OUR DOUBTS ARE TRAITORS AND MAKE
US LOSE THE GOOD WE OFT MIGHT WIN
BY FEARING TO ATTEMPT'

SHAKESPEARE

CONTENTS

	Page
ACKNOWLEDGEMENT	1
DECLARATION	3
ABSTRACT	4
CHAPTER I	14
CURRENT CONCEPTS OF GENOME ORGANIZATION AND EVOLUTION IN EUKARYOTES	
Types of repeated DNA sequences	14
- Inverted repeated sequences	14
- Satellite DNA sequences	20
- Repeated DNA sequences	24
- Dispersed repeat families	27
- Insertion sequences or mobile genetic elements	29
- Multigene families	32
- Single copy DNA	33
Genome organization in eukaryotes	34
Models of repetitive DNA evolution	39
Scope of thesis	43
Some features of plant species in the present study	46
References	54
CHAPTER II	67
HIGH RESOLUTION THERMAL DENATURATION ANALYSIS OF MILLET DNAs	
Summary	67
Introduction	68
Materials and Methods	71
- Seed material	71

	Page
- Chemicals	71
- Treatment and germination of seeds	72
- DNA isolation	72
- Criteria of purity of DNA preparations	74
- DNA sizing	78
- Thermal denaturation	78
- High resolution thermal denaturation analysis	81
Results	
- Yield and purity of DNA preparations	82
- Thermal denaturation of native DNAs	82
- High resolution thermal denaturation analysis	85
Discussion	
- Precautions taken during experimentation	88
- Comparison of DNA melting temperatures of Gramineae plant species	91
- Comparison of high resolution thermal denaturation profiles of millets	91
References	95
CHAPTER III	
NON-LINEAR LEAST SQUARES REGRESSION ANALYSIS OF DNA REASSOCIATION KINETICS OF MILLETS	97
Summary	97
Introduction	98
Materials and Methods	
- DNA isolation	100

	Page
- DNA content determination	101
- DNA shearing and sizing	101
- Hydroxyapatite preparation	101
- DNA reassociation kinetics	102
i) Optical reassociation	
ii) Hydroxyapatite column chromatography	104
- Non-linear least squares regression analysis of reassociation kinetics data	105
Results	
- DNA content	108
- Precautions taken during re-association kinetics studies	108
i) Purity of DNA	108
ii) Fragment length of DNA	108
iii) Control experiments	111
- Present analysis	111
Barn yard millet	111
Fox tail millet	112
Little millet	112
Great millet	116
Discussion	
- Comparison of DNA reassociation kinetics	117
- Salient features of fox tail millet	118
- Relationship between kinetic and cytophotometric estimates of genome size	119
- Relationship between the proportion of repetitive DNA and single copy DNA with nuclear DNA content	121
References	125

	Page
CHAPTER IV	
CHARACTERIZATION OF REPEATED DNA SEQUENCES IN MILLETS	128
Summary	128
Introduction	129
Materials and Methods	132
- Chemicals	132
- Preparation of DNA	132
- Isolation of repetitive DNA fractions	132
- Thermal denaturation of native 0.4 kbp long DNAs and repetitive DNA fractions	133
- Mini Cot curves	134
- Reassociation kinetics studies at different conditions of temperature stringency	134
- Digestion of total millet DNAs with restriction endonucleases	135
- High performance liquid chromatography	135
Results	
- Melting analysis of 0.4 kbp long millet DNAs and isolated repeat fractions	136
- Analysis of mini Cot curves	142
- Analysis of reassociation kinetics at different incubation temperatures	152
- Methylation status of millet DNAs as determined using restriction enzymes	162
- Occurrence of 5-methyl cytosine as determined by HPLC	164
Discussion	
- Possible function of GC rich sequences melting at high temperatures.	168

	Page
- Sequence divergence of repeated DNA sequences in Gramineae	169
- Absence of distinct classes of repetitive DNA in millets	171
- Reassociation kinetics studies at different incubation temperatures	172
References	176
CHAPTER V DNA SEQUENCE ORGANIZATION STUDIES IN MILLET GENOMES	178
Summary	178
Introduction	179
Materials and Methods	
- Chemicals	180
- DNA shearing and sizing	181
- Reassociation kinetics of long DNA fragments	183
- Determination of single strand collapse	183
- Isolation of S1 nuclease resistant repetitive duplexes	184
- Sizing of S1 nuclease resistant reassociation duplexes	185
Results	
- Reassociation kinetics of long DNA fragments	186
- Hyperchromicity of repetitive duplexes of varying lengths	196
- Characterization of S1 nuclease resistant repetitive duplexes	202
i) Duplex content	203
ii) Melting temperature	205
iii) Hyperchromicity	205
iv) Size distribution	207
v) Exact sizing of duplexes in the excluded volume of Agarose A-50 column chromatography	210

		Page
	vi) Repetitive sequence classes reassociating at lower Cot values	210
	- Spacing of repetitive sequence elements	212
	Discussion	
	Modes of interspersion patterns	214
	Thermal stability of repetitive duplexes	216
	Dynamics of sequence interspersion patterns in millets	216
	References	220
CHAPTER VI	SEQUENCE HOMOLOGIES AMONG THE REPEATED DNA SEQUENCES OF MILLET GENOMES	222
	Summary	222
	Introduction	223
	Materials and Methods	
	- Seed material	225
	- Extraction, shearing and sizing of total unlabelled DNAs	226
	- Isolation of repetitive DNA of barn yard millet	226
	- <u>In vitro</u> radio labelling of DNA	226
	- Sizing of labelled DNA	228
	- Reassociation studies	228
	- Hydroxyapatite thermal denaturation analysis	230
	Results	
	- Self reassociation of tracer DNA	231
	- Control experiments	232
	- Use of total nick translated pearl millet DNA as tracer	235
	- Use of nick translated repetitive DNA of barn yard millet as tracer	241

	Page
Discussion	247
Low homology among repeated DNA sequences of millets	247
References	250
CHAPTER VII	
GENERAL DISCUSSION	252
LIST OF PUBLICATIONS	259
PAPERS PRESENTED	260

ACKNOWLEDGEMENTS

I am deeply indebted to my teacher and guide Dr. P.K. Ranjekar whose untiring zeal and constant encouragement at all stages of my work cannot be adequately expressed in mere words. His kind words and patience at all times is what kept me going till the very end.

I profusely thank my friends Vidya and Shirish for their unstinted cooperation and helpful suggestions at all times. I am also grateful for the friendly assistance offered by my colleagues Mrinal, Meena, Shubhada and Madhu Mawal. My special thanks go to Dr. V. Shankar, Scientist, National Chemical Laboratory for examining the manuscript critically and offering his constructive criticism and suggestions. My thanks are also due to Dr. S.R. Gadre, Department of Chemistry, University of Poona for developing various computer programmes which have been useful in a number of analyses.

I will always remember the jolly companionship of my friends Shekhar, Madhu Malick, Medha, Yogesh, Lalitha, Dr.Hendre, Mr. B.G. Patil and other members of the Plant Molecular Biology group.

My special thanks also go to Mr. I.I. Sutar for his invaluable help in the compilation of the thesis. I gratefully acknowledge the technical assistance rendered by Mr. S.T. Dhume and Mr. M.V. Kamthe in handling various instruments and Mr. Pendse in photography. I am thankful to Smt. Indira for the rough drafting of the thesis and to Mr. S. Venkataraman for his neat and flawless typing of the final manuscript.

I thank Dr. V. Jaganathan, Dr. C. Sivaraman and Dr. M.C. Srinivasan for their help and encouragement throughout the period of my research. I also thank Dr. L.K. Doraiswamy, Director, National Chemical Laboratory, for extending all facilities and kind permission to submit this work as a thesis.

I appreciate the help extended by all the members of the Division of Biochemistry, National Chemical Laboratory during the work.

I am grateful to the various institutions for their prompt supply of seed materials used in the present investigation on millet species.

I wish to place in record my indebtedness to the CSIR for the fellowship I received throughout the period of my work.

Last but not the least, I am grateful to my parents for their constant support and encouragement. I also give them the credit of having raised me in an intellectual environment which shaped my interest in science. But for their blessings, this thesis would never have been a reality.

NCL, Poona 411 008

Dated 18th October 1984

Lakshmi Sivaraman

(Lakshmi)

DECLARATION

Certified that the work incorporated in the thesis 'GENOME CHARACTERIZATION IN PLANTS WITH SPECIAL REFERENCE TO FOUR MILLET SPECIES' submitted by Miss Lakshmi Sivaraman was carried out by the candidate under my supervision. Such material as has been obtained from other sources has been duly acknowledged in the thesis.



(P. K. RANJEKAR)

Supervisor

ABSTRACT

ABSTRACT

In family Gramineae, the nuclear genomes of crop plants such as wheat, oat, rye, barley, maize, rice, finger millet and pearl millet have been characterized with respect to their repeated DNA sequences, sequence organization and DNA sequence homologies. The present work describes studies on a few millet plant species known to be cultivated in Asia even under extreme climatic conditions. Molecular studies in these species may reveal mechanisms by which these plants have adapted themselves to the specific climatic conditions. Millet species selected for the investigation are barn yard millet (Echinochloa frumentacea), fox tail millet (Setaria italica), great millet (Sorghum vulgare) and little millet (Panicum miliare). All these millets have a DNA content of less than 5 picograms.

The work carried out in the aforementioned species has been grouped into six main Chapters and the results obtained are summarized as follows:

CHAPTER II

HIGH RESOLUTION THERMAL DENATURATION ANALYSIS OF MILLET DNAs

Melting studies have shown that all the four millet DNAs have smooth, monophasic thermal denaturation profiles

with the melting temperatures and G + C contents in the range of 87.3°C - 87.9°C and 43.9% - 45.1% respectively. A differential analysis of the DNA melting curves has revealed that the millets are rather characteristic in having several distinct components on the G + C rich side of the derivative melting profiles. For convenience, the melting components (about 10 - 14) have been divided into three groups. The first group includes the peak melting components lying in the narrow temperature range of 84-87°C. In all the four millets, the maxima of the differential profiles lie in the narrow range of 85.5 - 85.8°C. Great millet shows an additional prominent peak at 84.3°C and little millet at 86.7°C within this temperature range. The second group includes the A + T rich components melting in the range of 75 - 84°C. These components are not very distinct in the millet species and include small peaks prior to 79°C and mainly shoulders between 79-84°C. The third, major group consists of distinct G + C rich components melting after 87°C in the broad temperature range of 11°C. Barn yard millet shows two prominent peaks between 89-91°C of apparently equal proportion. Fox tail millet and little millet also show two distinct peaks between 88.4 - 90.6°C which differ in their positions and proportion. Great millet is unique in having no distinct melting component between 87.5 - 90.5°C.

All the millets show 3-5 distinct melting blocks after 92.5°C. The species specific differences that are observed on the G + C rich side of the differential profiles may be involved in the speciation events of millets. An attempt is being made to analyze these derivative melting profiles by Gaussian analysis in order to determine the relative proportion of each component. This analysis will be used to determine the genetic relatedness among these plants.

CHAPTER III

NON-LINEAR LEAST SQUARES REGRESSION ANALYSIS OF MILLET DNAs

Computer analysis of the reassociation kinetics data of 0.4 kbp long millet DNAs has revealed the presence of two repetitive and one single copy DNA component in barn yard millet and great millet and one repetitive and one single copy component in fox tail millet and little millet. All these components reassociate with second order kinetics. In addition to these DNA components, a very rapidly reassociating component is present in all the four millet species. Since this component reassociates too rapidly, it could not be resolved with the techniques used in the present investigation. It constitutes 6% of the genome in barn yard millet and great millet and 15-18% in fox tail millet and little millet.

The fast repetitive DNA component present only in barn yard millet and great millet has a reiteration frequency of 7.62×10^3 to 2.00×10^5 and a kinetic complexity of 8.75×10^2 to 6.94×10^3 nucleotide pairs. It represents 13-17% of the total DNA in the two millet species. The slow reannealing repetitive component, with a low reiteration frequency of 9.6×10^1 to 2.9×10^2 , constitutes 9% of the genome in fox tail millet and 24-27% of the genome in barn yard millet, little millet and great millet. The kinetic complexity of this component is estimated to be in the range of 2.84×10^4 to 1.23×10^6 nucleotide pairs. The total repetitive DNA is thus 42%, 26%, 38% and 52% in barn yard millet, fox tail millet, little millet and great millet respectively.

The single copy DNA sequences represent 45 to 50% of the genome with rate constants of 1.25×10^{-3} to 1.06×10^{-2} and kinetic complexities of 3.82×10^7 to 6.81×10^8 nucleotide pairs. About 14% of the genome in barn yard millet and little millet, 18% in great millet and 28% in fox tail millet fails to reassociate at a Cot of 1×10^4 M.s.

From the available data on plant species belonging to family Gramineae and the present data on four millet genomes, a relationship between nuclear DNA content and proportion of repeated and single copy sequences has been

postulated. Initially, both repetitive and single copy DNA sequences increase with increasing nuclear DNA content. However, after the nuclear DNA content value reaches 6 pg, the content of single copy DNA remains more or less constant while that of repetitive DNA continues to increase. Therefore, it appears that repeated DNA sequences contribute significantly to the genome size of these plants species.

CHAPTER IV

CHARACTERIZATION OF REPEATED DNA SEQUENCES IN MILLETS

Repetitive DNA from the four millet species were isolated on hydroxyapatite and were further characterized for their melting behaviour and optical reassociation properties. The base mismatch as determined from the T_m of the repetitive fractions ranges from 6-8 % in the four millet species.

An interesting observation is the biphasic character of the melting profiles of the repetitive fractions in barn yard millet, fox tail millet and little millet. The proportion of the low melting component (T_m 78 - 84°C) is 92-97% of the total genome while that of the high melting component (T_m 96-100°C) is 3-8%. This high melting fraction is shown to be selectively present in DNA sequences reannealing by Cot 0.1 only and is absent in those forming duplexes in

the Cot range of 0.1 to 10. The presence of a high melting component has also been shown in pearl millet. The three millets in the present study and pearl millet are closely related; are adapted to a dry climate; can survive in scanty rainfall conditions and can grow in poor, thin soil. Its presence has not been shown in other grasses like wheat, oat, rye, barley, maize, rice, finger millet or great millet. The high melting G + C rich component may, therefore, be playing an important role in the ecological adaptations of millets.

Kinetic heterogeneity and thermal stability of DNA fractions enriched in highly repetitive and intermediately repetitive sequences, isolated on hydroxyapatite, were also determined. It is observed that the highly repetitive DNA is thermally more stable than the intermediately repetitive DNA. Reassociation studies of these kinetic fractions have enabled us to detect the presence of very fast reassociating components which otherwise are undetectable while studying the kinetics of total 400 base pairs long DNA.

Digestions of total millet DNAs using the restriction enzymes Msp I, Hpa II and Hae III have helped in determining their methylation status. Extensive digestion with Msp I indicates that several of the external cytosine residues in the sequences 5' CCGG 3' are unmethylated. However, very

little digestion with Hpa II shows that several of the internal cytosine residues in the sequence 5' CCGG 3' are methylated. Extensive digestion with Hae III shows that the dinucleotide 5' Gp.C 3' in the sequence 5' GGCC 3' is unmethylated and such sequences occur with high frequency. The occurrence of 5-methyl cytosine in millet DNAs has also been shown by high performance liquid chromatography.

Reassociation kinetics of 0.4 kbp long DNAs were also studied at different stringency criteria of temperature other than the standard criterion of 62°C to determine the nature of repeated families in millet genomes. Such analysis indicates that barn yard millet, little millet and great millet genomes consist of heterogeneous families of repeated DNA sequences. Comparison of the proportion of 'fossil repeats' and 'true single copy' sequences suggests that the genome of great millet is 'turning over' at a much faster rate than that of barn yard millet and little millet.

CHAPTER V

DNA SEQUENCE ORGANIZATION IN MILLET GENOMES

Sequence interspersions studies were carried out using independent techniques such as reassociation kinetics studies of 1.5 kbp and 20 kbp long DNAs, hyperchromicity measurement of repetitive duplexes isolated on hydroxyapatite column at different fragment lengths, size distribution of repetitive duplexes using single strand specific

S1 nuclease and gel filtration on an agarose A-50 Biogel column. These studies have indicated that almost 85-95% of the genomes of these plants are made up of interspersed sequences. At the highest fragment size (20 kilobase), the total repetitive DNA is not resolved into more than a single component indicating that the slow repetitive DNA sequences and the rapidly renaturing DNA sequences are linked to one another. The single copy DNA size in barn yard millet and great millet lies in the range of 2500 - 3000 nucleotide pairs while that of fox tail millet and little millet is in the range of 1500 - 2000 nucleotide pairs. Repetitive DNA shows three sizes, namely, > 1500 nucleotide pairs, 500-1000 nucleotide pairs and 150-300 nucleotide pairs. Fox tail millet, however, shows predominantly a single size of repetitive DNA i.e. > 1500 nucleotide pairs. Thus barn yard millet, great millet and little millet have a mixed interspersion pattern while fox tail millet has essentially a long period interspersion pattern.

CHAPTER VI

DNA SEQUENCE HOMOLOGIES AMONG REPEATED DNA SEQUENCES IN MILLETS

In vitro DNA-DNA hybridization studies have been carried out to determine the interspecies homology of repeated DNA sequences in millets. When radiolabelled total pearl millet DNA is used as tracer, the homology between the repeated sequences of pearl millet to those of the other millets is

as low as 8-14%. When repetitive DNA of barn yard millet is used as tracer, the homology increases to 34-37% except in fox tail millet where it is still only 12.07%. Such low homologies among repeated sequences of millets and some of the earlier reported values in cereals tend to support the idea of polyphyletic origin of grasses.

A comparison of the present data on millet genomes with the available data in other temperate Gramineae plant species reveals several contrasting features, namely,

- i) Low nuclear DNA content (< 5 picograms)
- ii) Low proportion of repetitive DNA ($< 60\%$)
- iii) Presence of G + C rich DNA sequences in the highly repetitive DNA fractions, with great millet being an exception.
- iv) Moderate sequence divergence among repeated DNA sequences.
- v) Occurrence of heterogeneous repeat families.
- vi) Presence of either mixed or long period interspersed patterns.
- vii) Extremely low homology among repeated DNA sequences.

Based on the above data, it appears that the genomes of the four millets under the present investigation are 'turning over' at a much lower rate than those of wheat, rye and maize. It is further postulated that some environmental

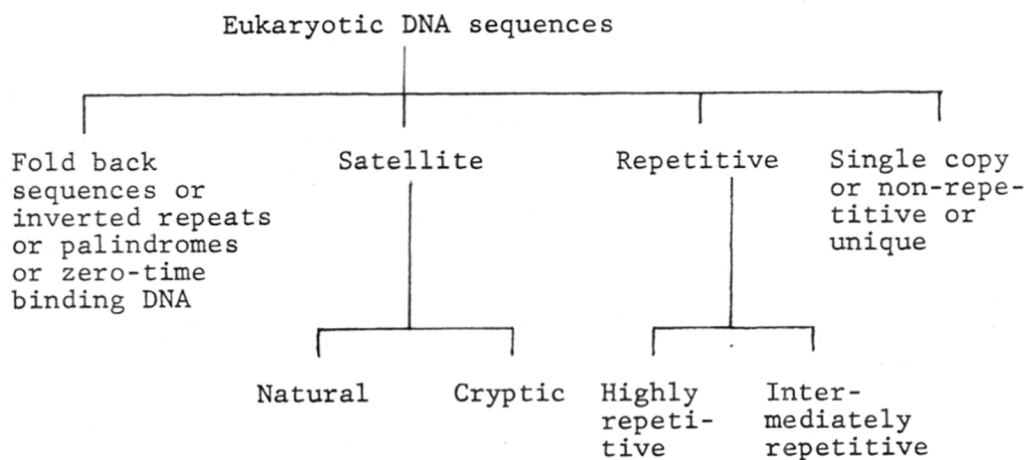
factor is imposing a functional constraint on the millet genomes and thus is able to maintain such low turn over rates. The turn over rates reflect on the genome size and are useful in predicting plant responses to climate, in plant breeding and in analysis of niche differentiation in plant communities. Temperature is one such environmental factor which affects genome size. Plant species growing in warm and arid climatic conditions like the present millets show small genomes which probably help them in ecological adaptations.

CHAPTER I

CURRENT CONCEPTS OF GENOME ORGANIZATION
AND EVOLUTION IN EUKARYOTES.

Most of our current concepts of structure and organization of eukaryotic DNAs are based on DNA reassociation kinetics experiments and equilibrium centrifugation of DNAs in cesium chloride and cesium sulfate gradients. Based on these studies, the nuclear genomes of eukaryotes have been shown to consist of mainly four types of DNA sequences. (Figure 1.1).

Figure 1.1



TYPES OF REPEATED DNA SEQUENCES

Inverted repeat sequences

Generally these sequences represent 1-5% of the total genome (1-19, Table 1.1) and are usually dispersed with other DNA sequences. They can be observed as hairpin

TABLE 1.1 REASSOCIATION KINETICS STUDIES IN HIGHER PLANTS

DNA Components of the plant	Fraction of genome	Copies per 1C genome	Kinetic complexity (base pairs)	Reference
1	2	3	4	5
<u>DICOTYLEDONS</u>				
<u>Arachis hypogaea</u>				
Highly repeated	0.12	38,000	2.9×10^5	1
Intermediately repeated	0.15	6,700	9.6×10^6	
Rare repeats	0.37	200	2.0×10^9	
Single copy	0.36	1	-	
<u>Benincasa hispida</u>				
Very fast	0.12	-	-	
Slow	0.36	142	6.70×10^5	2
Single copy	0.35	1	9.20×10^7	
<u>Brassica nigra</u>				
Very fast	0.03	-	-	3
Fast	0.25	22,640	5.1×10^3	
Slow	0.29	587	2.2×10^5	
Single copy	0.47	1	2.2×10^8	
<u>Coccinia indica</u>				
Very fast	0.14	-	-	
Slow	0.11	176	1.70×10^5	2
Single copy	0.59	1	1.59×10^8	
<u>Cucumis sativus</u>				
Very fast	0.16	-	-	
Slow	0.20	267	1.50×10^5	2
Single copy	0.37	1	7.40×10^7	

... Contd.

TABLE 1.1 (contd.)

1	2	3	4	5
<u>Glycine max</u>				
Very fast	0.05	-	-	
Fast	0.28	3,850	9.4×10^4	4
Slow	0.24	153	2.0×10^6	
Single copy	0.40	1	5.6×10^8	
<u>Gossypium hirsutum</u>				
Highly repeated	0.08	-	3.8×10^6	5
Repeated	0.27	134	9.7×10^8	
Single copy	0.61	1	2.9×10^{11}	
<u>Linum usitatissimum</u>				
Zero time binding	0.03	-	-	6
Satellite	0.11	-	-	
Fast repeats	0.29	-	-	
Intermediate repeats	0.15	-	-	
Single copy	0.35	1	3.5×10^8	
<u>Luffa acutangula</u>				
Very fast	0.08	-	-	2
Slow	0.51	319	1.00×10^6	
Single copy	0.27	1	1.69×10^8	
<u>Luffa cylindrica</u>				
Very fast	0.02	-	-	2
Fast	0.16	52,900	7.00×10^2	
Slow	0.33	90	8.50×10^5	
Single copy	0.36	1	8.37×10^7	
<u>Matthiola incana</u>				
Very fast	0.07	-	-	3
Highly repetitive	0.20	973,000	2.8×10^2	
Slow	0.42	192	3.0×10^6	
Single copy	0.31	1	4.8×10^8	

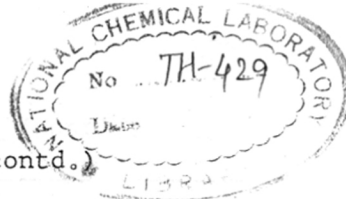


TABLE 1.1 (contd.)

1	2	3	4	5
<u>Nicotiana tabacum</u>				
Very fast	0.02	-	-	7
Fast	0.06	12,400	4.6×10^3	
Intermediate	0.59	252	2.5×10^6	
Single copy	0.22	1	2.8×10^8	
<u>Petroselinum sativum</u>				
Very fast	0.05	-	-	8
Highly repetitive	0.13	136,000	1.8×10^3	
Fast	0.48	3,000	3.0×10^5	
Slow	0.16	42	7.4×10^6	
Single copy	0.12	1	2.2×10^8	
<u>Phaseolus mungo</u>				
Rapidly reassociating	0.06	-	-	9
Intermediately reassociating	0.29	833	2.8×10^9	
Slow	0.66	1	5.3×10^8	
<u>Phaseolus aconitifolius</u>				
Rapidly reassociating	0.05	-	-	
Intermediately reassociating	0.36	857	1.0×10^6	9
Slow	0.59	1	1.4×10^9	
<u>Pisum sativum</u>				
Very fast	0.04	-	-	10
Fast	0.46	10,000	2.1×10^5	
Slow	0.33	3,000	5.0×10^6	
Single copy	0.21	1	4.5×10^9	
<u>Trichosanthes anguina</u>				
Very fast	0.06	-	-	2
Fast	0.19	3,260	1.80×10^4	
Slow	0.40	29	4.12×10^6	
Single copy	0.23	1	6.97×10^7	

... Contd.

581.169(043)
SIV

TABLE 1.1 (contd.)

1	2	3	4	5
<u>Vigna radiata</u>				
Very fast	0.02	-	-	
Fast	0.05	10,000	2.8×10^3	11
Slow	0.28	175	9.0×10^5	
Single copy	0.54	1	4.7×10^8	
<u>MONOCOTYLEDONS</u>				
<u>Allium cepa</u>				
Very fast	0.07	-	-	12
Fast	0.41	21,600	2.9×10^5	
Slow	0.36	225	2.4×10^7	
Single copy	0.06	1	9.3×10^8	
<u>Eleusine coracana</u>				
Very rapidly reassociating	0.17	-	-	13
Rapidly reassociating	0.07	1,500	1.7×10^4	
Intermediately reassociating	0.25	140	7.0×10^5	
Single copy	0.42	1	1.6×10^8	
<u>Hordeum vulgare</u>				
Very rapidly reassociating	0.08	-	-	14
Rapidly reassociating	0.34	-	-	
Slow repeats	0.30	-	-	
Single copy	0.28	-	-	
<u>Oryza sativa</u>				
Very rapidly reassociating	0.09	200,000	1.0×10^3	13
Rapidly reassociating	0.19	12,000	3.9×10^4	
Intermediately reassociating	0.25	190	3.2×10^6	
Single copy	0.48	1	1.2×10^9	

.... Contd.

TABLE 1.1 (contd.)

1	2	3	4	5
<u>Pennisetum amaricanum</u>				
Rapidly renaturing	0.23	-	-	16
Middle repetitive	0.47	298	3.1×10^5	
Single copy	0.17	1	3.3×10^7	
<u>Secale cereale</u>				
Highly repeated (I)	0.10	6×10^5 to 6×10^6	1.7×10^2 to 1.7×10^3	17
Highly repeated (II)	0.33	37,500	2.7×10^4	
Intermediately repeated	0.31	850	1.7×10^6	
Single copy	0.27	1	1.2×10^9	
<u>Triticum aestivum</u>				
Rapidly reannealing	0.04- 0.10	-	-	18
Intermediately reannealing	0.83	4,300	-	
Slow reannealing	0.12	1	-	
<u>Zea mays</u>				
Highly repeated	0.42	79,333	3.6×10^4	19
Middle repeated	0.17	1,133	2.5×10^6	
Non-repeated	0.30	1	2.8×10^9	

(-) indicates that the precise value has not been cited in the reference.

structures or bubbles under electron microscope (4,20,21). Ribosomal RNA as well as heterogeneous nuclear RNA from mammals have also been shown to contain a low percentage of foldback sequences (22-24). Palindromic sequences flank most of the transposable elements and probably are involved in transposition of genetic material (25). They are also supposed to function as binding sites for regulatory proteins (26, 27) and cleavage sites for restriction enzymes (28). These sequences are additionally implicated in the mechanism of gene expression and/or processing of the primary transcripts (29, 30).

Satellite DNA sequences

Many of the very highly repeated DNA sequences are present in long tandem arrays and appear as 'satellites' after equilibrium density gradient centrifugation of DNA in neutral CsCl (natural satellites) (31) or in Cs₂SO₄ containing heavy metals such as Ag⁺ or Hg⁺⁺ (cryptic satellites). Both natural and cryptic satellites occur in many plant species (31-33, Table 1.2).

In plants, satellites are particularly abundant in Cucurbitaceae and Rutaceae. Satellite DNAs are often heterogeneous in base composition and can be separated into subfractions differing in their melting properties, reassociation rates and degrees of methylation (51, 56-60).

TABLE 1.2 SATELLITE DNA SEQUENCES IN HIGHER PLANTS

Species	Satellite DNA % of total genome	Reference
1	2	3
NATURAL SATELLITES		
DICOTYLEDONS		
<u>Ammi visnaga</u>	-	33
<u>Aquilegia alpina</u>	-	34
<u>Brassica rapa</u>	21	35
<u>Brassica pekinensis</u>	17	35
<u>Brassica</u> (9 sp.)	37	36
<u>Bryonia dioica</u>	5	35
<u>Calendula officinalis</u>	-	35
<u>Citrullus vulgaris</u>	3	35
<u>Citrus</u> (4 sp.)	19-23	35
<u>Cucumis melo</u>	30	35,37
<u>Cucumis sativus</u>	16-28	35,38,39
<u>Cucurbita pepo</u>	16-18	35
<u>Daucus carota</u>	-	33,40
<u>Drimys piperata</u>	7	35
<u>Hammamelis mollis</u>	6	35
<u>Fortunella</u> sp.	24	35
<u>Gossypium</u> sp.	-	41
<u>Lagenaria vulgaris</u>	9	35
<u>Linum usitatissimum</u>	15	35,5
<u>Lobularia maritima</u>	32	35
<u>Lupinus angustifolius</u>	40-75	42
<u>Luffa cylindrica</u>	6	35
<u>Lycopersicon esculentum</u>	6	43
<u>Oenothera fruticosa</u>	-	34
<u>Phaseolus aconitifolius</u>	-	44
<u>Phaseolus aureus</u>	5	35

.... Contd.

TABLE 1.2 (contd.)

1	2	3
<u>Phaseolus coccineus</u>	24	35
<u>Phaseolus mungo</u>	-	44
<u>Phaseolus vulgaris</u>	19	35,45,46
<u>Raphanus sativus</u>	10	38
<u>Sinapis alba</u>	-	47
<u>Solanum crispum</u>	6	35
<u>Solanum tuberosum</u>	4	35
<u>Tropaeolum majus</u>	-	48
MONOCOTYLEDONS		
<u>Cymbidium hybrid</u>	15	49
CRYPTIC SATELLITES		
DICOTYLEDONS		
<u>Citrus paradisi</u>	-	32
<u>Cucumis melo</u>	-	32
<u>Cucumis sativus</u>	-	32
<u>Cucumis pepo</u>	-	32
<u>Linum usitatissimum</u>	-	32
<u>Vicia faba</u>	-	32,50
MONOCOTYLEDONS		
<u>Avena sativa</u>	-	32
<u>Hordeum vulgare</u>	8	32,51
<u>Scilla campanulata</u>	-	32
<u>Scilla sibirica</u>	19	32,50
<u>Secale cereale</u>	-	52
<u>Sinapis alba</u>	12	53
<u>Triticum aestivum</u>	12	51,54,55

(-) indicates that the precise value has not been cited in the reference.

The cryptic satellite of Sinapis alba is rather unique in lacking very short repeats and methylation (53). In general, the basic repeating unit of plant satellite DNAs is longer than that in animals (47, 58, 61, 62).

In situ localization of satellite DNAs in plant chromosomes has shown that these sequences are generally distributed in the heterochromatic regions of chromosomes, chromocentres or nucleolar organizer regions (48,50,52,63). Studies on the relationship between satellite DNA and heterochromatin can throw light on the origin and function of satellite DNA sequence (64-67). Satellite DNA sequences have been implicated in chromosome condensation, homologous chromosome recognition and pairing, providing specific attachment sites for spindle fibres and nuclear membrane and maintenance of proper association between particular regions of nonhomologous chromosomes during interphase, in addition to their role as genetic regulators, spacers and mutation accumulating loci (65).

In many instances, differential replication of satellite DNA sequences has been observed under certain physiological conditions, and thus satellite sequences could be involved in response to physiological variations (68-70). All of these studies, however, are correlative and even closely related species often differ with respect to the

TABLE 1.3 REPEATED DNA SEQUENCES

content of satellite DNA sequences (Table 1.2). Satellite DNA sequences are considered to be very sensitive indicators of evolutionary divergence. It is suggested that they may be involved in establishing fertility barriers during speciation (64,66).

Repeated DNA sequences

Most of the higher eukaryotic genomes are composed of different classes of repeated DNA sequences differing in proportion, kinetic complexity and frequency of reiteration (Tables 1.1, 1.3) (1-19, 71-93).

The highly repeated DNA sequences which generally occur as long tandem arrays, are repeated $10^3 - 10^6$ times, and have a kinetic complexity of $10^3 - 10^5$ base pairs. The moderately repeated DNA sequences occur about $10^1 - 10^3$ times in the genomes and have kinetic complexities in the range of $10^5 - 10^6$ base pairs. These sequences are generally dispersed throughout the genome and appear to be more conserved than the highly repeated DNA sequences. Thermal denaturation studies have indicated that wide variations exist in the degree of sequence divergence in repetitive DNAs of different plant species. For example, in rye, wheat, barley, tobacco, finger millet and maize, upto 15% base mismatch is observed in repetitive DNA duplexes (6,13,15,17-19) but in rice, the base mismatching is only

TABLE 1.3 REPEATED DNA SEQUENCES IN HIGHER PLANTS

Species	Repetitive DNA % of total genome	References
1	2	3
DICOTYLEDONS		
<u>Anacyclus</u> sp.	70.0-89.0	71
<u>Anemone</u> sp.	53.0-67.0	72
<u>Anthemis</u> sp.	63.0-67.0	71
<u>Arachis hypogaea</u>	64.0	1
Beans	40.0-70.0	73-75
<u>Brassica pekinensis</u>	57.0	3
<u>Cinnamomum camphora</u>	37.3	76
<u>Cucumis sativus</u>	55.0	77
<u>Decaisnea fargesii</u>	48.5	76
<u>Glycine max</u>	60.0	78,4
<u>Gossypium hirsutum</u>	40.0-50.0	5,79
<u>Helianthus annuus</u>	69.0	80
<u>Helianthus tuberosus</u>	45.0	81
<u>Lathyrus</u> sp.	56.0-72.0	82
<u>Linum usitatissimum</u>	56.0	6
<u>Liriodendron tulipifera</u>	47.5	76
<u>Magnolia soulargiana</u>	39.5	76
<u>Matthiola incana</u>	69.0	3
<u>Nicotiana tabaccum</u>	70.0	7
<u>Petroselinum sativum</u>	50.0-70.0	8
<u>Phaseolus</u> (4 sp.)	35.0-47.0	9,46
<u>Pisum sativum</u>	65.0-75.0	9,83-85
<u>Raphanus sativus</u>	45.0	77
<u>Tropaeolum majus</u>	70.0	48
<u>Vicia</u> (3 sp.)	70.0-90.0	86

... Contd.

TABLE 1.3 (contd.)

1	2	3
<u>Vigna radiata</u>	35.0	10
MONOCOTYLEDONS		
<u>Allium</u> (7 sp.)	44.0 - 60.0	87
<u>Allium</u> (3 sp.)	83.0 - 60.0	86
Daffodil lily	50.0	88
<u>Eleusine coracana</u>	49.0	13
<u>Hedera helix</u>	44.0 - 56.0	89
<u>Lilium henryii</u>	60.0	90
Orchids (4 sp.)	45.0	63
<u>Oryza sativa</u>	52.0	13
<u>Pennisetum americanum</u>	69.0	16
	54.0	13
<u>Poa annua</u>	87.0	91
<u>Poa trivialis</u>	82.0	91
<u>Scilla sibirica</u>	40.0	92
Wheat, oat rye, barley	70.0 - 83.0	15,17,18,93
<u>Zea mays</u>	78.0	91
	60.0	19

1% which is the lowest reported so far (13). Phaseolus species show an intermediate degree of base mismatch (8, 46).

Many functional roles have been assigned to repeated DNA sequences. Most of the early suggestions were based on cytological observations. A clear evidence for the functional involvement of repeated DNA sequences in meiosis comes from the studies on crossing over in Lilium microsporocytes, wherein it is observed that a fraction of moderately repeated DNA sequences is synthesized during repair synthesis in pachytene (94,95). In addition to the roles in chromosome mechanics, certain fractions of repeats have been assigned functions which will be described in the following sections. However, significance of the much larger fraction of repetitive DNA remains, more or less, unknown or speculative.

Dispersed repeat families

Discovery of restriction endonucleases and development of molecular cloning and DNA sequencing techniques have revolutionized the study of structure of complex eukaryotic genomes. Digestion of DNA with specific restriction endonucleases and the determination of the digestion pattern by electrophoretic methods has led to the identification of two different classes of highly repeated DNA sequences in mammals. Dispersed families with unit lengths under 500 base pairs are found

TABLE 1.4 DISPERSED REPEAT FAMILIES IN MAMMALS

Type of family	Organism	Copy number	Length of sequences (base pairs)	Reference
1	2	3	4	5
Alu I	Human	100,000-500,000	300	96-98
Alu-equivalent	Galago	-	-	98
	Lemur	-	-	98
	New world monkey	-	300	99
	Chinese hamster	-	140,96	96,98
	Mouse	40,000-80,000	130,190	96
	Rat	-	-	96
C family	Rabbit	-	-	100
	Chicken	-	-	96
R family	Mouse	100,000	475	101,102
Hinf	Human	50-100	319	103
Kpn I	Primates	$6 \times 10^4 - 10^5$	1200,1500	104-106
			1800,1900	107
EcoRI	Human	-	340	108
MIF	Mouse	50,000	1300	109,110,111
	Apodemus	-	-	109,110,111
	Rat	-	-	109,110,111
	Chinese hamster	-	-	109,110,111
EcoRII	Mouse	-	1500,1700	112
Xba I	Human	-	850,1300,	105
	Gibbon	-	1900	105
	Great ape	-	-	105
	Old world monkey	-	1300,850	105
Hind III	Human	-	1900	113,114
Bam HI	Mouse	-	5600	115

(-) indicates that the precise value has not been cited in the reference.

in hundreds of thousands of copies, and are termed as SINES (short interspersed repeated sequences). The other class, composed of repeat units several kilo base pairs in length, occurring about 10^4 (or fewer) times in the genome, is termed as LINES (long interspersed repeated sequences) (96-115, Table 1.4). The LINES and/or SINES in the mammalian genomes have been suggested to have various functions due to their widespread occurrence. For example, the 4.5 S and 7S small nuclear RNAs of mouse and human cultured cells are considered as the Alu member transcripts (96). Kpn I sequences are also transcribed in HeLa cells by RNA polymerase III into abundant heterogeneous species of RNA (106). Alu I and Kpn I members are often interspersed with structural genes and could be involved in recombination and gene conversion events (104, 116). The involvement of these sequences has also been suggested in DNA replication (117), genome reorganization and RNA processing (118). It is also assumed that these sequences have a role in definition of discrete chromosomal and nuclear domains that are distinct from those occupied by the predominant centromeric repeats (113).

Insertion sequences or mobile genetic elements

To date many genetic systems under the control of mobile sequence elements have been established both in

TABLE 1.5 MOBILE DNA SEQUENCES IN HIGHER PLANTS

System	Plant	Gene affected	Reference
Ds-Ac	Maize	Shrunken locus coding for sucrose synthetase, waxy locus determining amylose content of the endosperm	119-125
En-1	Maize	Loci involved in anthocyanin synthesis	126
Ds-Mp	Maize	Loci involved in anthocyanin synthesis	127
Uq	Wheat	Strains infected with Wheat Streak Mosaic Virus	128
Cin 1	Maize		129
3.4 Kb unit	Soybean	L1 lectin gene	130,131

(-) indicates that the precise value has not been cited in the reference.

pro/ and eukaryotes. The plant mobile genetic elements are listed in Table 1.5, (119-131). The most well-studied transposons in higher plants are those in maize, namely (i) the Ac-Ds insertion elements affecting the 'shrunk' locus that encodes the enzyme sucrose synthetase and thus leading to mutants with collapsed endosperm (119-122), (ii) the Ac-Ds system active at 'waxy' loci determining the amylose content of pollen and endosperm tissue (123-125), and (iii) the En-I (126) and Ds-Mp (127) systems causing instability at the loci responsible for anthocyanin synthesis. The mutant waxy loci have been cloned to isolate Ac and Ds elements (125). Ds element is found to be 4.1 and 2.0 kilo base pairs in length, while Ac element is 4.3 kilo base pairs long. The 4.1 kilo base pairs long Ds element is highly homologous to Ac elements, while the 2.0 kilo base pairs long Ds is homologous to the ends of Ac element. Sequences homologous to the ends of Ac are present in many copies in the maize genome, while those homologous to centre of Ac are very few. The Ac element at waxy locus is structurally different from other genomic Ac-like sequences as shown by restriction enzyme digestions.

In majority of instances, the mobile sequences lead to inactivation of genes. These sequences are also known to mobilize genes and activate silent genes lacking

normal promoters (132-134) in addition to being involved in certain developmental pathways (25). The intervening sequences in eukaryotic split genes are generally flanked by short-imperfect direct repeats (136). It is suggested that the fluidity of DNA rearrangement resulting from the frequent transposition events can offer considerable scope for adaptations (135, 137). Mobile elements can cross barriers which normally **separate cells, organisms or species**. Phage Mu, retroviruses and T-DNA of Agrobacterium tumefaciens are interspecific and even intergeneric vectors or agents for genome reorganization (138). In addition, a single mobile element in Drosophila (P element) can behave differently in different cell types (138). Flavell et al. (139) suggest that the movement of short pieces of DNA to new sites is a relatively common phenomenon in plant chromosome evolution.

Multigene families:

Structural genes have generally been considered as unique sequences. However, variety of multigene families are now known in which the gene sequences are present in sufficient copies to be included in moderately repetitive DNA. They may even show up as discrete components on buoyant density centrifugation in cesium chloride or cesium sulfate gradients. The examples of plant multigene families include ribosomal RNA genes (34, 140-150), storage protein genes (151-155), actin genes of soybean (156, 157) and genes coding

for leghaemoglobin (158, 159).

Many multigene families such as those of ribosomal RNA genes are organized in arrays of tandemly repeating units which include a large amount of spacer DNA sequences and sometimes noncoding intervening sequences within the genes themselves. The functions of such spacers and insertion sequences in multigene families is not known. It is suggested that they play a role in the operation of correction mechanisms (such as unequal crossing over) required for the parallel evolution of tandem genes (160). In addition to tandemly repeated genes, copies of some of these genes can be found on different chromosomes (161). Moreover, all the multigenes of a family may or may not be functional and can be composed of pseudogenes or truncated genes (159). The existence of several multigene families have been shown but the mechanisms by which they have evolved has remained largely speculative. 'Specialized' genes, 'house keeping or constitutive function' genes, 'luxury' genes are some of the concepts of multigenes that have been put forward (162).

Single copy DNA

The most slowly reassociating DNA sequences in eukaryotic genomes are usually assumed to be single copy sequences. There are some plant species like V. faba and V. sativa where truly unique DNA sequences are not present

(75). Based on reassociation kinetics, generally in plants the proportion of single copy DNA varies in the range of 10-65% (Table 1.1).

The unique or single copy DNA sequences are considered to include the sites of classical genetic activity (163-166). However, usually only a very small fraction of unique DNA (1-10%) is found to be transcribed into mRNA (167, 168). Nuclear RNA typically contains 4-10 times excess transcripts, but most of them never leave the nucleus, and, therefore, do not code for proteins (169, 170). Thus, the amount of unique DNA does not correlate with the biological complexity of the organism, any more than the total DNA content does.

GENOME ORGANIZATION IN EUKARYOTES

Satellite DNA and highly repetitive DNA generally occur as tandemly linked simple sequences while the members of the intermediately repetitive class are located throughout the genome. Britten^{and} Kohne (171) were the first to demonstrate the interspersion of intermediately repetitive and unique DNA sequences. Two general types of interspersion patterns have been discerned. The Xenopus pattern or 'short period' interspersion pattern is characterized by repeat units about 300 base pairs long interspersed with single copy sequences about 800 base pairs long (172). The second is the Drosophila pattern or 'long period' interspersion

pattern which is characterized by families of repeat units several kilo base pairs in length separated by tens of thousands of kilobase pairs of single copy sequences (173). The studies of sequence organization in Xenopus (172) sea urchin (174), housefly (175), fishes (176), mammals (177-179) and Dictyostelium (180) demonstrated a wide spread occurrence of the 'short period' interspersion pattern which suggested that this arrangement probably has a functional role. The gene regulation model of Britten and Davidson (181) and its subsequent modification (182, 183) is an attempt to integrate the information available on DNA sequence organization and gene regulation in eukaryotes, where repetitive sequences have been suggested to contain the regulatory elements for the expression of the adjacent structural genes.

Subsequent studies, however, have revealed the occurrence of the 'long period' interspersion pattern in a wide variety of unrelated organisms such as birds (184, 185), nematodes (186) and fungi (187-190).

Data on interspersion patterns in higher plants are still modest. However, the initial characterization of genomes in terms of short term and long term depending on the spacings of repeats flanking single copy segments seems to be an oversimplification as far as plants are concerned. Seventeen plant species, representatives of sixteen genera and eight

families of dicotyledons as well as monocotyledons have been studied so far for their genome organization. The sizes of the interspersed repetitive and single copy sequences and the general patterns of interspersion in these species are summarized in Table 1.6. From the table, it is seen that there is great diversity in the sequence organization patterns of plants. Wheat, rye, maize and tobacco exhibit predominantly a short period interspersion pattern while the rest of the plants have either mixed or long period interspersion patterns. From the table, it is also apparent that the size of interspersed single^{copy} sequences varies in the narrow range of 1000-3000 nucleotide pairs with the exception of Vigna radiata, Gossypium hirsutum and Pisum sativum where other size classes are present. The interspersed repeat sequences, on the other hand, have a variety of size classes such as 50-100 nucleotide pairs, 300-1000 nucleotide pairs, 1250-2000 nucleotide pairs and 4000-4500 nucleotide pairs. Rice and flax are rather unique in exhibiting lack of interspersion at a DNA fragment size of 6400 and 10,000 nucleotide pairs, respectively and the presence of very long repetitive stretches. Apart from the wide variations in the lengths of interspersed repetitive sequences, there is also a broad range of sizes of repetitive DNA sequences that are interspersed with single copy sequences in the same genome like in Vigna

TABLE 1.6 DNA SEQUENCE ORGANIZATION PATTERNS IN PLANTS.

Systematic Name	Haploid DNA content (pg)	Length of the interspersed repetitive sequences (rp)	Length of interspersed single copy sequences (np)	Pattern of interspersed sequences (np)	Reference
1	2	3	4	5	6
<u>Vigna radiata</u>	0.50	300-120, 1200-6700	750, 3950	Intermediate	11
<u>Oryza sativa</u>	0.60	50-100, 6000-6400	-	Absent	191
<u>Linum usitatissimum</u>	0.60	>10,000	-	Long	6
<u>Gossypium hirsutum</u>	0.80	1250	1800, 4000	Intermediate	5
<u>Matthiola incana</u>	1.50	350-400	1050	Short	3
<u>Eleusine coracana</u>	1.60	150-200, 4000-4200	1900	Mixed	192
<u>Phaseolus aureus</u>	1.80	1900	1300	Intermediate	193
<u>Phaseolus vulgaris</u>	1.85	1545	1300	Intermediate	193
<u>Petroselinum sativum</u>	2.00	300-400	1000	Mixed	8
<u>Pennisetum americanum</u>	2.50	4300-4500	1900	Long	14, 16
<u>Glycine max</u>	3.25	300-400, 1500	1235	Short	4, 78
<u>Nicotiana tabacum</u>	3.90	300, 1500	1400	Short	7
<u>Pisum sativum</u>	4.95	300- 400	300-1000	Short	10
<u>Zea mays</u>	7.70	500-1000, >5000	2100	Short	19
<u>Secale cereale</u>	9.00	300-800	1500	Short	194
<u>Alli m cepa</u>	16.75	300	1000	Short	12
<u>Triticum aestivum</u>	18.10	300-400	1500	Short	195

(-) indicates that the precise value has not been cited in the reference.

radiata.

Another interesting aspect that has emerged from the analysis of sequence organization in plants is the dependence of interspersion pattern on the genome size. Plants with a genome size ($2C$) smaller than 5 picograms show predominance of long period or mixed interspersion patterns, while genomes ($2C$) larger than 5 picograms show the predominance of short period interspersion pattern. Thus interspersion patterns are considered to be nucleotypic effect. It is suggested that the great diversity in genome organization pattern may not reflect on differences in the gene regulation mechanisms but is more related to the dynamics of DNA. Much of the DNA serves no strongly sequence-dependent role in gene regulation and is therefore relatively free to evolve by amplification, deletion and base substitutions. When both amplifications and deletions occur during evolution, the portion of the genome subject to these processes can be viewed as 'turning over' on a geological time scale. Frequent amplifications would lead to highly repetitive genomes and when coupled with relatively frequent translocation events, genomes with extensive short period interspersion would result (196). Thus, the pattern of sequence arrangement is determined by the rate of 'turn over' of the genomes.

MODELS OF REPETITIVE DNA IN EVOLUTION

Various models have been presented to account for the accumulation and persistence of repetitive DNA in eukaryotic genomes. These models can be classified according to the level at which natural selection has operated during production of repeats, and whether the action of natural selection is sequence dependent or independent.

The function-dependent models have been suggested particularly by Britten and Davidson (181) and Davidson and Britten (182, 183). These models suggest a sequence-dependent positive selection of the repeats due to their effects on the developmental programmes of the organism. The proposed roles of these repeats include (i) regulation of transcription, (ii) production of evolutionary novelty, (iii) encoding of regions within large RNAs which allows them to function as regulatory activators and (iv) integration of chromosome structure and recognition.

The recent sequence-dependent models include that of 'selfish DNA' proposed by Orgel and Crick (197) and Doolittle and Sapienza (198). However, this model suggests that only those sequences accumulate which can make use of the cell's replication machinery. Phenotypic selection acts only to check the accumulation when it begins to affect severely the working of the genome.

The highly sequence-independent 'ignorant' model of Dover (195) also suggests phenotypic selection only as a check. The model postulates no direct selection of repeats at all and suggests that they are randomly generated side-products of the genomic machinery for DNA metabolism.

The 'nucleotypic' model of Cavalier-Smith (199, 201) is also sequence-independent, but suggests a link between selection based on phenotype and accumulation of moderate repeats. Selection acts only on overall DNA content and this determines the physical volume of the nucleus, and thus indirectly the cell size, cell-cycle time, nucleus-cytoplasm transport and other aspects of nuclear and cell size.

According to Bouchard (202), the genomes are 'predisposed' to produce repeats, and the 'tolerance' of the genome is the rate at which it can accept such events. According to the 'ignorant' model, the production of periodic saltation is an unavoidable byproduct of complex machinery of DNA metabolism of the cell. This may be true to some degree, since Darlington (203) has argued that characteristics which have an overall selective advantage are adopted during evolution, even if they carry small disadvantages. Once such mechanisms are incorporated, they are quite difficult to change or replace. Jain (204) suggests that there may have been

positive selection for saltatory replication as a potential source of evolutionary novelty not provided by point mutations. The tendency towards saltation may vary from one genome to another.

Finally, under the scheme suggested by Bouchard (202), the potential for accumulation of moderate repeats within a given genome will be governed by the extent of predisposition, receipt of certain sequences from external sources and tolerance exhibited by the ancestral lineages. The amount and characteristics of repeats observed at present will be a function of the passive rate of accumulation over time (balanced to an unknown degree by probably passive deletion). In addition, some portion of the accumulated repeats may have become fixed, having possessed or acquired a selectively advantageous function or due to autonomous tendencies (e.g. transposons).

All the aforementioned models of repetitive DNA evolution are not mutually exclusive. Each genome may be subjected to various processes to different degrees and DNA sequence changes accumulate until at a certain level they are compatible with speciation. Speciation is a process of organisms being separated into non-interbreeding units. Both phenotypic and genotypic variations are associated with the differentiation of species and the question arises

as to whether any class of variation can be cited as causative rather than consequential and associative. Since there must be many paths of speciation, a universally applicable answer is unlikely, but we might still meaningfully ask whether there are any particular aspects of DNA sequence variability which are associated with the events of speciation. In order to understand evolution and speciation, one must know the principal sources and kinds of changes in the genome and the mechanisms by which these changes alter gene expression (205).

SCOPE OF THESIS

Among all the group of plants, cereal grains are the most important to mankind. They are used as staple food all over the world. They are widely distributed and occur in different ecological zones. Evidently, such wide distribution and success in evolution has resulted from a range of cytotypes, polyploidy, natural hybridization and lastly by partial apomictic reproduction. A few instances have also been cited where the genotypes have been correlated to specific adaptive features.

Some of the important Gramineae species such as rice, wheat, great millet, pearl millet, finger millet and the minor millets selected for the present study are commonly cultivated as field crops throughout India. Genetical research in cereals and millets in India has achieved significant landmark both in basic and applied research. In rice and wheat, dwarfing genes have brought about a transformation of the yield levels. In maize, pearl millet and sorghum, heterosis as well as composite varieties are the significant achievements. The basic research has been conducted in a variety of aspects including studies of these genomes at the molecular level. Extensive cytogenetic studies have also been carried out with this group of plants from which their taxonomic and

evolutionary relationships may be deduced. In cereals, the temperate grasses like wheat, oat, rye, barley and maize have been extensively characterized at the genomic level with respect to sequence organization and evolution. A few tropical plants like pearl millet, finger millet and rice have also been characterized earlier in our laboratory. Analysis at the molecular level has indicated several distinct differences between the tropical and temperate plants. In order to substantiate these observations, we selected four more tropical plant species, namely barn yard millet (Echinochloa frumentacea), fox tail millet (Setaria italica), little millet (Panicum miliare) and great millet (Sorghum vulgare). The present analysis on the four millet nuclear genomes has been divided into seven main Chapters.

Chapter I summarizes the general concepts about the genome organization and evolution in eukaryotes.

Chapter II describes the isolation of DNA from millets and its characterization. The technique of high resolution thermal denaturation analysis has been used to assess the base composition heterogeneity in millet DNAs.

Chapter III describes the use of non-linear least squares regression analysis to analyze the reassociation kinetics of 0.4 kbp long millet DNAs and thereby determine

the proportion and kinetic complexities of repeated and single copy DNA sequences in millets.

Characterization of isolated repeat fractions by melting and optical reassociation is dealt with in Chapter IV. Nature of repeat families and proportions of 'true' single copy and 'fossil' repeats has been determined by studying the reassociation kinetics of 0.4 kbp long millet DNAs at different incubation temperatures.

Arrangement of repeated and single copy DNA sequences in the four millets is detailed in Chapter V.

In Chapter VI, the sequence homologies among the repeated DNA sequences of millets are studied using the technique of in vitro DNA-DNA hybridization.

The important conclusions of the work, besides being detailed after each chapter, are dealt with in the last Chapter. Speculative aspects of the present study are also discussed.

SOME FEATURES OF PLANTS IN THE PRESENT STUDY

Echinochloa frumentacea (Roxb.) Link. (2n= 36,54)

Syn.: Panicum frumentaceum (Roxb.)

Common names: Japanese barn yard millet, 'Samva' millet.

This is the quickest growing of all the millets and produces a crop in about six weeks. It is grown in The Orient, India, China and Japan as a substitute for rice when the paddy fails. It is grown as a rainfed crop, but may be cultivated in water logged areas. E. frumentacea is said to have originated from E. colona or E. crus-galli and possesses characters intermediate between them. It is grown as forage crop, feed for cage birds, as food, to make beer and as fodder.

It is a robust tufted annual with broad leaves, large spikelets and small, smooth and shining seeds. Inflorescence is a panicle, often tinged purple, bearing upto fifteen lateral branches. Spikelets are awnless. Seeds are greenish yellow in colour, rounded at the base and pointed at the apex.



ECHINOCHLOA FRUMENTACEA
(BARN YARD MILLET)

Setaria italica (L.) Beauv. ($2n = 18$)

Syn.: Panicum italicum (L.) Scribn.

Common names: Fox tail millet, Italian, German, Hungarian
or Siberian millet and 'Kangani' millet.

This millet is widely cultivated in Japan and India. It is grown in regions of low rainfall, from sea level to 2000 m. It can tolerate a wide range of soil conditions from light sands to heavy clays and gives reasonable yield on inferior, infertile soils. It cannot tolerate water-logging and may be susceptible to very long periods of draught. A number of forms and cultivars are recognized. It is probably derived from S. viridis (L.) Beauv. It is used as food, in making beer, as birdfeed and as fodder.

The plant is erect, tufted and an annual grass with prominent jointed slender culms. Leaves are flat, linear or lanceolate. Inflorescence is a spikelike panicle with ciliate lateral branches. Spikelets are subtended by bristles. Seeds are ellipsoid or globose, smooth and pale yellow to dark brown in colour.



SETARIA ITALICA
(FOX TAIL MILLET)

Panicum miliare Lank. ($2n = 36$)

Syn.: Panicum sumatrense Roth. ex. Roem & Schult

Common names: Little millet, 'Shaven' millet.

This minor millet is grown throughout India but is of little importance elsewhere. It can thrive under adverse conditions and can be grown in soils which otherwise produce little or nothing. It can mature into a small crop even in famine years. It is closely related to P. psilopodium Trin. It is used as food and fodder.

It is an annual tufted grass with slender culms and is usually leafy upto the panicle. Leaves are soft, narrow and rather long for the size of the plant. Panicles are oblong with numerous erect, hairy branches which eventually are nodding. The seeds are ellipsoidal or globose, smooth and greenish.



PANICUM MILIARE
(LITTLE MILLET)

Sorghum vulgare Pers. (2n = 20).

Syn.: Sorghum bicolor (L.) Moench, Holchus, Sorghum, L.

Andropogon sorghum (L.) Brot.

Common names: Great millet, Guinea corn, Kafir corn, milo,
Sargo, Kaoliang, durra, mtama, jola, jawa,
jowar or cholam.

Sorghum is the fourth most important world cereal, following wheat, rice and maize. It is a staple food grown in the drier parts of tropical Africa, India and China. Many cultivars are recognized. It is adapted to a wide range of ecological conditions and will produce useful yields of grain under conditions which are unfavourable for most other cereals. It is essentially a plant of hot and warm countries. It can tolerate hot and dry conditions but can also grow in areas with high rainfall in which waterlogging may occur. It is killed by frost. The great merit of sorghum is its drought resistance and xerophytic characters.

Sorghum is an annual, with a single, solid, erect stem. The panicle is also erect with compact racemes. The grain is usually partly covered by the glumes and is pale yellow in colour.



SORGHUM VULGARE
(GREAT MILLET)

REFERENCES

- 1 Dhillon, S.S., Rake, A.V. and Miksche, J.P. (1980).
Plant Physiol., 65: 1121-1127.
- 2 Bhawe, M., Lagu, M. and Ranjekar, P.K. (1984).
Plant Sci.Lett., 33: 127-136.
- 3 Wenzel, W. and Hemleben, V. (1979).
Plant Syst.Evol.Suppl., 2: 29-40.
- 4 Gurley, W.B., Hepburn, A.G. and Key, J.L. (1979).
Biochim.Biophys.Acta, 561: 167-183.
- 5 Walbot, V. and Dure, L.S. (1976).
J.Mol.Biol., 101: 503-536.
- 6 Cullis, C.A. (1981) Biochim.Biophys.Acta,
652: 1-15.
- 7 Zimmerman, J.L. and Goldberg, R.B. (1977).
Chromosoma (Berl.), 59: 227-252.
- 8 Kiper, M. and Herzfeld, F. (1978).
Chromosoma (Berl.), 65: 335-351.
- 9 Seshadri, M. and Ranjekar, P.K. (1979).
Indian J.Biochem.Biophys., 16: 1-15.
- 10 Murray, M.G., Cuellar, R.E. and Thompson, W.F. (1978).
Biochemistry (USA), 17: 5781-5790.
- 11 Murray, M.G., Palmer, J.D., Cuellar, R.E. and
Thompson, W.F. (1979). Biochemistry (USA),
18: 5259-5266.
- 12 Stack, S.M. and Comings, D.E. (1979),
Chromosoma (Berl.), 70: 161-181.
- 13 Deshpande, V.G. and Ranjekar, P.K. (1980).
Hoppe-Seyler's Z.Physiol.Chem., 361: 1223-1233.
- 14 Gupta, V.S. (1981). Structure and sequence
organization of DNA in eukaryotes with special
reference to three Gramineae plant species,
Ph.D. Thesis, Poona University, Poona, India.

- 15 Ranjekar, P.K., Pallotta, D., Lafontaine, J.G. (1976).
Biochim.Biophys.Acta, 425: 30-40.
- 16 Wimpee, C.F. and Rawson, J.R.Y. (1979).
Biochim.Biophys.Acta, 562: 192-206.
- 17 Ranjekar, P.K. Lafontaine, J.G. and Pallotta, D.
(1974). Chromosoma (Berl.), 48: 427-440.
- 18 Smith, D.B. and Flavell, R.B. (1975).
Chromosoma (Berl.), 50: 223-242.
- 19 Hake, S. and Walbot, V. (1980). Chromosoma (Berl.),
79: 251-270.
- 20 Walbot, V. and Dure, L.S. (1976), J.Mol.Biol.,
101: 503-536.
- 21 Bazetoux, S., Jouanin, L. and Huguet, T. (1978).
Nucleic Acids Res., 5: 750-768.
- 22 Wellauer, P.K. and Dawid, I.B. (1974). J.Mol.Biol.,
89: 379-395..
- 23 Jelinek, W., Molloy, G., Salditt, M., Wall, R.,
Sheiness, D. and Darnell, J.E. (Jr.) (1973)
Cold Spring Harbor Symp. Quant.Biol., 38: 891-898.
- 24 Ryskov, A.P. Saunders, G.F., Farashyan, V.R. and
Georgiev, G.P. (1973). Biochim.Biophys.Acta, 312:
152-164.
- 25 Calos, M.P. and Miller, J.H. (1980).
Cell, 20: 579-595.
- 26 Dickson, R.C., Abelson, J., Barnes, W.M. and
Reznikoff, W.S. (1975). Science, 187: 27-35.
- 27 Darlix, J.L. and Horaist, M. (1975).
Nature (Lond.), 256: 288-292.
- 28 Kelly, T.J. and Smith, H.O. (1970).
J.Mol.Biol., 51: 393-409.
- 29 Dunn J.J. and Studier, F.W. (1973). Proc.Natl.
Acad.Sci.(USA), 70: 3296-3300.

- 30 Vogeli, G., Ohkubo, H., Sobel, M.E., Yamada, Y.,
Pastan, I. and de Crombrughe, B. (1981).
Proc.Natl.Acad.Sci.(USA), 78: 5334-5338.
- 31 Skinner, D.M. (1977), Bioscience, 27:790-796.
- 32 Deumling, B., Sinclair, J., Timmis, J.N. and Ingle, J.
(1976). Cytobiologie, 13: 224-232.
- 33 Ohyama, K., Gamborg, O.L. and Miller, R.A.
(1972). Plant Physiol., 50: 319-321.
- 34 Ingle, J., Timmis, J.N. and Sinclair, J. (1975).
Plant Physiol., 55: 496-501.
- 35 Ingle, J., Pearson, G.G. and Sinclair, J. (1973).
Nature (New Biol)., 242:193-197.
- 36 Beridze, T.G. (1975). Biochim.Biophys.Acta,
395 : 274-279.
- 37 Bendich, A.J. and Anderson, R.S. (1974).
Proc.Natl.Acad.Sci. (USA), 71: 1511-1515.
- 38 Ranjekar, P.K. Pallotta, D. and Lafontaine, J.G.
(1978), Can .J.Biochem., 56: 808-815.
- 39 Kadouri, A., Atsmon, D. and Edelman, M. (1975).
Proc.Natl.Acad.Sci.(USA), 72: 2260-2264.
- 40 Duhrssen, E. and Neumann, K. -H.(1980).
Z.Pflanzen physiol., 100: 447-454.
- 41 Aripdzhanov, Sh.A., Beshpal'Ko N.V. and
Ibraginsov, A.P. (1976). Uzbek Biol Zh., 1: 9-10.
- 42 Strubbe, K., van Oostveldt, P., Brockaert, D. and
van Parijs, R. (1982). Planta, 155: 238-243.
- 43 Chilton M-D. (1975). Genetics, 81: 469-483.
- 44 Sehadri, M. and Ranjekar, P.K. (1979).
Indian J.Biochem.Biophys., 16: 1-15.
- 45 Beridze, T.G. and Bragvadze, G.P. (1976).
Molekulyarnaya Biologiya 10: 1279-1289.
- 46 Seshadri, M. and Ranjekar, P.K.
(1980). Hoppe-Seyler's Z. Physiol.Chem., 361:
1041-1048.

- 47 Capesius, I. (1979). Plant Syst.Evol., 133:1-14.
- 48 Deumling, B. and Nagl, W. (1978). Cytobiologie, 16: 412-420.
- 49 Capesius, I. Bierweiler, B. Bachmann, K., Rucker, W. and Nagl, W. (1975). Biochim.Biophys.Acta, 395:67-73.
- 50 Timmis, J.N., Deumling, B. and Ingle, J. (1975). Nature (Lond.), 257: 152-155.
- 51 Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976) Biochim.Biophys.Acta, 520: 103-110.
- 52 Appels, R. Driscoll, C. and Peacock, W.J. (1978). Chromosoma (Berl.), 70: 67-89.
- 53 Capesius, I. (1983). Biochim.Biophys.Acta, 739: 276-280.
- 54 Huguet, T. and Jouanin, L. (1972). Biochem.Biophys. Res.Comm., 46: 1169-1174.
- 55 Peacock, W.J., Lohe, A.R. Gerlach, W.L., Dunsmuir, P., Dennis, E.S. and Appels, R. (1977). Cold Spring Harbor Symp.Quant.Biol., 42: 1121-1135.
- 56 Bendich, A.J. and Taylor, W.C. (1977). Plant Physiol., 59: 604-609.
- 57 Reis, R.S., Timmis, J.N. and Ingle, J. (1981). Biochem. J., 195: 723-734.
- 58 Hemleben, V., Leweke, B., Roth, A. and Stadler, J. (1982). Nucleic Acids Res., 10: 631-644.
- 59 Beridze, T.G. (1979). Biochemistry(USSR), 44: 1546-1552.
- 60 Bragvadze, G.P. (1983). Biochemistry (USSR), 48: 48-55.
- 61 Reisner, A.H., Colless, V.M. and Bucholtz, C. A. (1980). Aust.J.Biol.Sci., 33: 623-632.
- 62 Lee, C.S. (1981). Chromosoma (Berl.), 83: 367-379.
- 63 Nagl, W. and Capesius, I. (1977). Chromosomes Today, 6: 141-152.

- 64 Yunis, J.J. and Yasmineh, W.G. (1971).
Science, 174: 1200-1209.
- 65 Sharma, A.K. (1978). Nucleus, 21: 113-116.
- 66 Fry, K. and Salser, W. (1977). Cell, 12:1069-1084.
- 67 Walker, P.M.B. (1971). Nature (Lond.), 229:
306-308.
- 68 Grisvard, J. and Tuffet-Anghileri, A. (1980).
Nucleic Acids Res., 8: 2843-2858.
- 69 Durante, M., Geri, C., Grisvard, J., Guille, E.,
Parenti, R. and Buiatti, M. (1983). Protoplasma,
114: 114-118.
- 70 Nagl, W. and Rücker, W. (1976). Nucleic Acids Res.,
3: 2033-2039.
- 71 Fuhrmann, B and Nagl, W. (1979). Plant Syst.Evol.,
Suppl. 2: 235-245.
- 72 Cullis, C.A. and Schweizer, P. (1974). Chromosoma
(Berl.), 44: 417-421.
- 73 Chooi, W.Y. (1971). Genetics, 68: 213-230.
- 74 Straus, N.A. (1972). Carnegie Inst.Wash. Yearbook,
71: 257-259.
- 75 Gnucheva, E., Yu, Kir'yanov, G.I., Polykov, V. Yu
and Chenstov Yu S. (1977). Biokhimiya, 42: 2186-2191.
- 76 Schaan, M.E. and Nagl, W. (1979).
Plant Syst.Evol.Suppl., 2: 67-71.
- 77 Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G.
(1978). Can.J.Biochem., 402: 171-178.
- 78 Goldberg, R.B. (1978). Biochem.Genet., 16: 45-68.
- 79 Wilson, J.T., Katterman, F.R.H. and Endrizzi, J.E.
(1976). Biochem.Genet., 14: 1071-1075.
- 80 Nagl, W. and Caperius, I. (1976). Plant Syst.Evol.
126, 221-237.
- 81 Nze-Ekekang, L., Patillon, M., Schäfer, A. and
Kovoor, A. (1974). J.Expt.Bot., 25: 320-329.

- 82 Narayan, R.K.J. and Rees, H. (1977). *Chromosoma Berl.*, 63: 101-107.
- 83 Sivolap, Y.M. and Bonner, J. (1971). *Proc.Natl.Acad.Sci.(USA)*, 68: 387-389.
- 84 Pearson, W.R., Smith, S.L., Wu, J.R. and Bonner, J. (1978). *Plant Physiol.*, 62: 112-115.
- 85 Pearson, W.R., Wu, J.R. and Bonner, J. (1978). *Biochemistry*, 17: 51-59.
- 86 Gaddipatti, J.P.B. (1976). Characterization of chromosomal DNA in higher plants with varying genome size. Ph.D. thesis, Banaras Hindu University, Varanasi, India.
- 87 Ranjekar, P.K., Pallotta, D., and Lafontaine, J.G. (1978). *Biochem.Genet.*, 16: 957-970.
- 88 Bendich, A.J. and Anderson, R.S. (1977). *Biochemistry*, 16: 4655-4663.
- 89 Schaffner, K.H. and Nagl, W. (1979). *Plant Syst. Evol.Suppl.*, 2: 105-110.
- 90 Smyth, D.R. and Stern, H. (1973). *Nature (New Biol.)*, 245: 94-96.
- 91 Flavell, R.B. and Smith, D.B. (1975). *Stadler Symp.*, 7: 47-70.
- 92 Stanziano, P. and Maggini, F. (1976). *ATTI Associazione Genetica Italiana*, 21: 66-69.
- 93 Bendich, A.J. and McCarthy, B.J. (1970). *Genetics*, 65: 545-565.
- 94 Stern, H. and Hotta, Y. (1977). *Philos.Trans.R.Soc. London, Ser.B*, 277: 277-294.
- 95 Stern, H. and Hotta, Y. (1978). *Annu.Rev.Plant Physiol.*, 29: 415-436.
- 96 Jelinek, W.R. and Haynes, S.R. (1983). *Cold Spring Harbor Symp. Quant.Biol.*, 47: 1123-1130.
- 97 Soriano, P. Meunier-Rotival, M. and Bernardi, G. (1983). *Proc.Natl.Acad.Sci. (USA)*, 80: 1816-1820.

- 98 Daniels, G.R. and Deininger, P.L. (1983).
Nucleic Acids Res., 11: 7595-7610.
- 99 Daniels, G.R., Fox, G.M., Loewensteiner, D.,
Schmid, C.W. and Deininger, P.L. (1983).
Nucleic Acids Res., 11: 7579-7593.
- 100 Fritsch, E.F., Shen, C.K.J., Lawn, R.M. and
Maniatis, T. (1981). Cold Spring Harbor Symp.
Quant.Biol., 45: 761-775.
- 101 Gebhard, W., Meitinger, T., Hochtl, J. and
Zachau, H.G. (1982). J.Mol.Biol., 157: 453-471.
- 102 Gebhard, W. and Zachau, H.G. (1983), J.Mol.Biol.,
170: 255-270.
- 103 Shimizu, Y., Yoshida, K., Ren, C-S, Fujinaga, K.
Rajagopalan, S. and Chinnadurai, G. (1983).
Nature (Lond.), 302: 587-590.
- 104 Shafit-Zagardo, B., Brown, F.L., Maio, J.J. and
Adams, J.W. (1982). Gene, 20: 397-407.
- 105 Gillespie, D., Adams, J.W., Costanzi, C. and
Jo Caranfa, M. (1982). Gene, 20: 409-414.
- 106 Shafit-Zagardo, B., Brown, F.L., Zavodny, P.J.
and Maio, J.J. (1983). Nature (Lond.), 304:
277-280.
- 107 Lerman, M.I., Thayer, R.E. and Singer, M.F.
(1983). Proc.Natl.Acad.Sci. (USA), 80: 3966-3970.
- 108 Wu, J.C. and Manuelidis, L. (1980), J.Mol.Biol.,
142: 363-386.
- 109 Brown, S.D.M. and Dover, G. (1981), J.Mol.Biol.,
150: 441-446.
- 110 Dubnick, M., Chou, J., Petes, T.D. and Farber, R.A.
(1983), J.Mol.Evol., 19: 115-121.
- 111 Heller, R. and Arnheim, N. (1980). Nucleic Acids Res.,
8: 5031-5042.
- 112 Manuelidis, L. (1980). Nucleic Acids Res., 8:
3247-3258.
- 113 Manuelidis, L. (1982). Nucleic Acids Res., 10:
3211-3219.

- 114 Manuelidis, L. and Biro, P.A. (1982).
Nucleic Acids Res., 10: 3221-3239.
- 115 Meunier-Rotival, M., Soriano, P., Cuny, G.,
Strauss, F. and Bernardi G. (1981).
Proc.Natl.Acad.Sci.(USA), 79: 355-359.
- 116 Fox, G.M., Hess, J.F., Shen, C-K.J. and Schmid, C.W.
(1983). Cold Spring Harbor Symp. Quant.Biol.,
47: 1131-1134.
- 117 Taylor, J.H. (1984). Molec. and Cell Biol.,
61: 99-109.
- 118 Kraye, A.S., Markuseva, T.V., Kramerov, D.A.,
Ryskov A.P., Skryalin, K.G., Bayer, A.A. and
Georgiev, G.P. (1982). Nucleic Acids Res.,
10: 7461-7475.
- 119 McClintock, B. (1950). Proc.Natl.Acad.Sci.USA,
36: 344-355.
- 120 McClintock, B. (1951). Cold Spring Harbor Symp.,
Quant. Biol., 16: 13-47.
- 121 Burr, B. and Burr, F.A. (1981). Genetics, 98:
143-156.
- 122 Courage-Tebbe, U., Doring, H.P., Federoff, N.
and Starlinger, P. (1983). Cell, 34: 383-393.
- 123 Echt, C.S. and Schwartz, D. (1981). Genetics, 99: 275-284. 119-131
- 124 Shure, M. Wessler, S. and Federoff, N. (1983),
Cell, 35: 225-233.
- 125 Federoff, N., Wessler, S. and Shure, M. (1983).
Cell, 35: 235-242.
- 126 Reddy, A.R. and Peterson, P.A. (1983).
Molec.Gen.Genet., 192: 21-31
- 127 Kermicle, J.L. (1980). Science, 208: 1457-1459.
- 128 Friedemann, P. and Peterson, P.A. (1982).
Molec.Gen.Genet., 187: 19-29.
- 129 Shepherd, N.S. Schwarz-Sommer, Z., vel Spalve, J.B.,
Gupta, M., Weinand, U. and Saedler, H. (1984).
Nature (Lond.) 307, 185-187.

- 130 Goldberg, R.B., Hoschek, G. and Vodkin, L.O.
(1983). *Cell*, 33: 465-475.
- 131 Vodkin, L.O., Rhodes, P.R. and Goldberg, R.B.
(1983). *Cell*, 34: 1023-1031.
- 132 Hinton, D.M. and Musso, R.E. (1983).
J.Mol.Biol., 169: 53-81.
- 133 Jaurin, B. and Normark, S. (1983). *Cell* 32:
809-816.
- 134 Simons, R.W., Hoopes, B.C., McClure, W.R.
and Kleckner, N. (1983). *Cell*, 34: 673-682.
- 135 Cohen, S.N. and Shapiro, J.A. (1980).
Scientific American, 242: 40-49.
- 136 Crick, F. (1979). *Science*, 204: 264-271.
- 137 Nerbert, D.W. (1979). *Molec.Cell Biochem.*,
27: 27-46.
- 138 Shapiro, J.A. and Cordell, B. (1982).
Biol.Cell, 43: 31-54.
- 139 Flavell, R.B., O'Dell, M. and Hutchinson, J.
(1981). *Cold Spring Harbor Symp. Quant.Biol.*,
45: 501-508.
- 140 Hemleben, V., Bachmann, K. and Price, H.J. (1978).
Experientia, 34: 1452-1453.
- 141 Gerlach, W.L. and Bedbrook, J.R. (1979).
Nucleic Acids Res., 7: 1869-1885.
- 142 Carrozza, M.L., Cremonini, R. and Giorgi, L.
(1979). *ATTI Associazione Genetica Italiana*
Vol. XXIV pg. 87.
- 143 Ingle, J. (1979). In: *Molecular Biology of Plants*.
Rubenstein, I., Phillips, R., Green, C.E. and
Gengenbach, B.G. eds. pp.139-164. Academic Press,
New York.
- 144 Varsanyi-Breiner, A., Gusella, J.F., Keys, C.
and Housman, D.E. (1979). *Gene*, 7: 317-334.
- 145 Appels, R., Gerlach, W.L. Dennis, E.S.,
Swift, H. and Peacock, W.J. (1980).
Chromosoma (Berl.) 78: 293-311.

- 146 Cullis, C.A. and Goldsbrough, P.B. (1980).
In: The Plant Genome. Davies, D.R. and Hopwood,
D.A. eds. pp.91-97. The John Innes Charity, Norwich.
- 147 Marazia, T., Barsanti, P. and Maggini, F. (1980).
Biochem.Genet., 18: 509-517.
- 148 Maggini, F. and Carmona, M.J. (1981).
Protoplasma, 108: 163-171.
- 149 Appels, R. and Dvorak, J. (1982). Theor.Appl.
Genet., 63:361-365.
- 150 Oono, C. and Sugiura, M. (1980). Chromosoma (Berl.),
76: 85-89.
- 151 Langridge, P. and Feix, G. (1983).
Cell, 34: 1015-1022.
- 152 Rubenstein, I. (1982). In: Maize for Biological
Research. Sheridan, W.F. ed. pp.189-195.
Plant Molecular Biology Association. Charlottesville,
VA 22905.
- 153 Viotti, A., Sala, E., Marotta, R., Alberi, P.,
Balducci, C. and Soave, C. (1979).
Eur.J.Biochem., 102: 211-222.
- 154 Schuler, M.A., Schmitt, E.S. and Beachy, R.N.
(1982). Nucleic Acids Res., 10: 8225-8244.
- 155 Hall, T.C., Slightom, J.L., Ersland, D.R.,
Murray, M.G., Hoffman, L.M., Adang, M.J.,
Brown, J.W.S., Ma, Y., Matthews, J.A., Cramer, J.H.,
Barker, R.F. Sutton, D.W., and Kemp, J.D. (1983).
In: Structure and Function of Plant Genomes.
Ciferri, O. and Dure III, L. eds. pp.123-142.
Plenum Publishing Corporation.
- 156 Nagao, R.T., Shah, D.M., Eckenrode, V.K.
and Meagher, R.B. (1981) DNA, 1: 1-10.
- 157 Shah, D.M., Hightower, R.C. and Meagher, R.B. (1983).
J.Mol.Appl.Genetc., 2: 111-126.
- 158 Baulcombe, D. and Verma, D.P.S. (1978).
Nucleic Acids Res., 5: 4141-4153.
- 159 Brisson, N. and Verma, D.P.S. (1982).
Proc.Natl. Acad.Sci. (USA), 79: 4055-4059..

- 160 Smith, G.P. (1976). Science, 191:528-535.
- 161 Soave, C., Reggiani, R., DiFonzo, N. and Salamini, F. (1982). Biochem. Genet., 20: 1027-1038.
- 162 Lewin, B. (1983). In: Genes. pp.292-299 Wiley, J. and Sons. Inc., New York.
- 163 Suzuki, Y., Gage, L.P. and Brown, D.D. (1972). J.Mol.Biol., 70: 637-649.
- 164 Bishop, J.O. and Rosbash, M. (1973). Nature(New Biol), 241: 204-207.
- 165 Harrison, P.R. Birnie, G.D., Hell, A., Humphries, S., Young, B.D. and Paul, J. (1979). J.Mol.Biol., 84: 539-554.
- 166 Harris, S.E., Means, A.R., Mitchell, W.M. and O'Malley, B.W. (1973). Proc.Natl.Acad.Sci.(USA), 70: 3776-3780.
- 167 Kleiman, L., Birnie, G.D., Young, B.D. and Paul, J. (1977). Biochemistry, 16: 1218-1223.
- 168 Goldberg, R.B., Hoschek, G., Kamalay, J.C. and Timberlake, W.E. (1978). Cell, 14: 123-131.
- 169 Galau, G.A., Britten, R.J. and Davidson, E.H. (1974). Cell, 2: 9-20.
170. Hough, B.R., Smith, M.J., Britten, R.J. and Davidson, E.H. (1975). Cell, 5: 291-299.
- 171 Britten, R.J. and Kohne, D.E. (1968). Science, 161: 529-540.
- 172 Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973). J.Mol.Biol., 77: 1-23.
- 173 Manning, J.E., Schmid, C.W. and Davidson, N. (1975). Cell, 4: 141-155.
- 174 Graham, D.E., Neufeld, B.R., Davidson, E.H. and Britten, R.J. (1974). Cell 1: 127-137.
- 175 Efstratiadis, A., Crain, W.R. Britten, R.J., Davidson, E.H. and Kafatos, F.C. (1976). Proc.Natl.Acad.Sci.(USA),73: 2289-2293.

- 176 Olmo, E. Stingo, V., Odierna, G. and Capriglione, T. (1982). *Experientia*, 38: 339-340.
- 177 Mayfield, J.E., Mc Kenna, J.F. and Lessa, B.S. (1980). *Chromosoma (Berl.)*, 76: 271-294.
- 178 Avvedimento, V.E., Acquaviva, A.M. and Varrone, S. (1976). *Nucleic Acids Res.*, 3: 2491-2505.
- 179 Schmid, C.W. and Deininger, P.L. (1975). *Cell*, 6: 345-358.
- 180 Firtel, R.A. and Kindle, K. (1975). *Cell*, 5: 401-411.
- 181 Britten, R.J. and Davidson, E.H. (1969). *Science*, 165: 349-357.
- 182 Davidson, E.H. and Britten, R.J. (1975). *Quart.Rev.Biol.*, 48: 565-613.
- 183 Davidson, E.H. and Britten, R.J. (1979). *Science*, 204: 1052-1059.
- 184 Epplen, J.T., Leipoldt, M., Engel, W. and Schmidtke, J. (1978). *Chromosoma (Berl.)*, 69: 307-321.
- 185 Wagenmann, M. Epplen, J.T., Bachmann, K., Engel, W. and Schmidtke, J. (1981). *Experientia*, 37: 1274-1276.
- 186 Schachat, F., O'Connor, D.J. and Epstein, H.F. (1978). *Biochim.Biophys.Acta*, 520: 688-692.
- 187 Ullrich, R.C., Kohorn, B.D. and Specht, C.A. (1980). *Chromosoma (Berl.)*, 81: 371-378.
- 188 Harshey, R.M., Jayaram, M. and Chamberlin, M.E. (1979). *Chromosoma (Berl.)*, 73: 143-151.
- 189 Hudspeth, M.E.S., Timberlake, W.E. and Goldberg, R.B. (1977). *Proc.Natl.Acad.Sci. (USA)*, 74: 4332-4336.
- 190 Krumlauf, R. and Marzluf, G.A. (1979). *Biochemistry*, 18: 3705-3713.
- 191 Gupta, V.S. Gadre, S.R. and Ranjekar, P.K. (1981). *Biochim.Biophys.Acta*, 656: 147-154.
- 192 Gupta, V.S. and Ranjekar, P.K. (1981). *J.Biosci.*, 3: 417-430.

- 193 Seshadri, M. and Ranjekar, P.K. (1980).
Biochim.Biophys.Acta, 610: 211-220.
- 194 Smith, D.B. and Flavell, R.B. (1977).
Biochim.Biophys. Acta, 474: 82-97.
- 195 Flavell, R.B. and Smith, D.B. (1976). Heredity,
37: 231-252.
- 196 Thompson, W.F. and Murray, M.G. (1980).
In: The Plant Genome. Davies, D.R. and Hopwood,
D.A., eds.pp.31-45. The John Innes Charity, Norwich.
- 197 Orgel, L.E. and Crick, F.H.C. (1980).
Nature (Lond.), 284: 604-607.
- 198 Doolittle, W.F. and Sapienza, C. (1980).
Nature (Lond.), 284: 601-603.
- 199 Dover, G. (1980). Nature (Lond.), 285: 618-620.
- 200 Cavalier-Smith, T. (1978). J.Cell.Sci.,
34: 247-278.
- 201 Cavalier-Smith, T. (1980). Nature (Lond.),
285: 617-618.
- 202 Bouchard, R.A. (1982) In: International Review
of Cytology, Bourne, G.H. and Danielli, J.F. Eds.
Vol. 32, pp.113-193. Academic Press, New York.
- 203 Darlington, C.D. (1971). Nature (Lond.),
234: 521-525.
- 204 Jain, H.K. (1980). Nature (Lond.), 288:647-648.
- 205 Peacock, W.J., Dennis, E.S. and Gerlach, W.L.
(1982). In: Mechanisms of speciation,
pp. 123-142, Alan, R. Liss, Inc., 150
5th Avenue, New York, N.Y. 10011.

CHAPTER II

HIGH RESOLUTION THERMAL DENATURATION
ANALYSIS OF MILLET DNAS

SUMMARY

Thermal denaturation of the DNAs of four millet species were carried out to determine the T_m , G + C content, hyperchromicity and transition width. The T_m values fall in the range of 85.3 - 85.7°C and the G + C content, as determined from the T_m , is 43.87 - 45.09%. The transition width varies in the range of 7.67 - 8.03°C indicating a moderate sequence divergence and hyperchromicities average between 22-25%. High resolution thermal denaturation analysis indicates the presence of 11-14 components with distinct base composition. Species specific differences are observed in the G + C rich side of the derivative melting profile. Though barn yard millet, fox tail millet and little millet show two distinct peaks in the region between 88.4 - 91.0°C, their proportions and positions are apparently different. Great millet shows no distinct peaks in the region of 87 - 90.5°C, but has a distinct peak at 85.3°C which is absent in the other millets. These species specific differences that are observed in the G + C rich sequences may be responsible for the speciation events in the four millet species.

INTRODUCTION

The physical structure of DNA is considered to have evolved in such a way that the genetic information is probably protected by an array of stable base pairs inside the double helical folding of the two chains. Since the DNA double helix is unfolded while reading out the genetic code, the unfolding process is directly related to genetic function.

Melting of DNA indicates the separation of the two strands initially held together by hydrogen bonds and hydrophobic interactions. The DNA melting profile in general is a two dimensional display obtained by plotting a change in some measurable property of DNA like alteration in viscosity, optical rotation, buoyant density, light scattering and extinction coefficient against an external variable responsible for melting of DNA such^{as}/solvents or concentration of ligand molecules. The thermal melting profile of DNA is primarily characterized by the following parameters, namely, the melting temperature (T_m) which is defined as the temperature at which 50% of the DNA is denatured (1), the transition width ΔT_a^b or σ 2/3rd defined as the difference between the temperatures at which a and b % melting occurs where a = 16.67% and b = 18.33% (2) and hyperchromicity which is expressed in terms of percent increase in absorbance after complete

denaturation (3,4). Under constant conditions of ionic strength and pH, a linear relationship exists between the melting temperature and base composition of DNA expressed in terms of G + C.

The most convenient and reliable method for studying the melting of DNA is to monitor the change in the UV absorbance by elevating the temperature, under precise environmental conditions. Since absorbance by nucleotide bases in the wavelength region around 260 nm is due to $\pi-\pi^*$ electronic transition in purine and pyrimidine bases, the intensity of absorption increases by 20-40% when the hydrogen bonds between the bases are disrupted. A linear relationship exists between this hyperchromicity and the extent of disruption of base pairs.

Eukaryotic DNAs are generally heterogeneous. However, this heterogeneity is not immediately apparent from a simple curve relating percent hyperchromicity and temperature. It can be readily detected by plotting 'derived' melting curves, where the ratio of hyperchromicity over a small rise in temperature (e.g. 0.1°C) to the total hyperchromicity is plotted against temperature. Such studies have revealed an ordered and non-random distribution of base sequences in most organisms. Three kinds of derivative melting profiles are generally recognised: bacteriophage and (a) polyphasic melting of short, homogeneous viral DNAs

consisting of unique sequences (5 - 9); (b) smooth, monophasic melting of bacterial and large viral DNAs (6,10) and (c) the broad, polyphasic melting of higher eukaryotes (10-15).

Since the melting profiles provide a histogram of the local G + C content arranged over several hundred base pairs, it provides a convenient method to examine the homology and/or to explore the taxonomic relationship among DNAs like deletion mutants of lambda phage (16), kinetoplast DNAs of hemoflagellates (17) and a series of related bacteriophages of Bacillus subtilis (9). Species and family specific melting curves have been obtained in some higher plants (12, 18). Fine melting studies have been shown to have advantages over the other techniques of genetic distance evaluation, like restriction mapping and DNA-DNA hybridization (9). In addition, the technique has apparently a greater resolution compared to techniques like density gradient centrifugation and electron microscopy employed for studying distribution in base composition (10,19).

This Chapter describes the thermal denaturation characteristics and high resolution melting analysis of DNAs of barn yard millet, fox tail millet, little millet and great millet, ^{to} assess the heterogeneity in base composition of these millet DNAs and the extent of similarities and

differences among them.

MATERIALS AND METHODS

Seed material

Seeds of barn yard millet (Echinochloa frumentacea var. Madira) were obtained from Vivekananda Parvatiya Krishi Anusandhan Shala, ICAR, Almora, Uttar Pradesh. Italian fox tail millet (Setaria italica) seeds were received from IDRC project, Cotton Research Station, Nandyal, Andhra Pradesh. Little millet (Panicum miliare) seeds were procured from Regional Research Station, OUAT, Semiliguda, Sunabeda, Orissa State. Seeds of great millet (Sorghum vulgare var. Maldandi) and wheat (Triticum aestivum) were purchased locally.

Chemicals

All the chemicals used throughout the work were of Analytical Reagent (AR) or Guaranteed Reagent (GR) Grade, and were obtained from British Drug House (BDH), Sarabhai Chemicals or E. Merck. Fine chemicals/reagents such as RNase A, Tri^zma base [tris(hydroxymethyl)aminomethane] and agarose were obtained from Sigma Chemical Co., MO, USA. Molecular weight markers (Lambda phage DNA digested with Hind III and phage ϕ x174 RF DNA digested with Hae III) and restriction enzymes were obtained from Boehringer Mannheim, West Germany or Bethesda Research Laboratories, USA. Calf

thymus DNA was also obtained commercially from Sigma Chemical Co., MO, USA.

Treatment and germination of seeds

Seeds, washed thoroughly with soap water and distilled water, were allowed to remain in alcohol (50%) for 10 min. They were then washed and treated with 0.001% mercuric chloride for 10 min. to reduce the bacterial contamination. The seeds were then kept soaked in distilled water for 5-6 h. The swollen seeds were kept for germination on wet cotton in the dark. Growth was maintained under constant temperature (25°C) and humidity (90%). Shoots, about 2-3" long, (grown for 6-7 days) were harvested an inch above the surface to minimize bacterial contamination. They were frozen immediately in liquid nitrogen and stored at -20°C till further use.

DNA isolation

The extraction procedure was a combination of that of Marmur (20) and Ranjekar et al. (21). Plant tissue was frozen and powdered in liquid nitrogen and homogenized in a Remi blender at high speed in buffer I (0.5 M sucrose, 0.05 M Tris, 0.05 M maleic acid and 0.003 M $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, pH 6.0) containing 0.1% Triton X-100 for five minutes. The homogenate was filtered through two layers of cheese cloth or muslin cloth and the filtrate centrifuged at 1000 x g for 30 min. The crude nuclear pellet was washed twice with buffer I,

once with buffer II (1.0 M sucrose, 0.05 M Tris, 0.05 M maleic acid and 0.003 M $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, pH 6.0) containing 0.1% Triton X-100 and finally washed with buffer I. Each washing involved resuspending the pellet in the buffer and centrifugation at 1000 x g for 20 min. The washed pellet was suspended in saline-EDTA (0.15 M NaCl, 0.1M sodium salt of EDTA, pH 8.0) and the nuclei were lysed by the addition of sodium dodecyl sulphate to a final concentration of 2% followed by incubation at 62°C for 30 min. with intermittent gentle shaking. The lysate was then cooled to 4°C and the proteins were denatured by the addition of sodium perchlorate (effective concentration of 1M) followed by deproteinization with chloroform/isoamyl alcohol (24:1 v/v). The latter step was repeated till a clear interphase was obtained. The aqueous layer was chilled and the nucleic acids were precipitated with 1.5 volumes of chilled ethanol. The fibres were centrifuged down, dried free of ethanol, dissolved in 1 x SSC (0.15 M NaCl, 0.015 M trisodium citrate, pH 7.0) or T_{10}E_1 (10 mM Tris-HCl, pH 7.5, 1 mM sodium salt of EDTA) and treated with DNase free RNase (50 µg/ml) at 37°C for 1 h. The DNA solution was then deproteinized with chloroform/isoamylalcohol (24:1 v/v) mixture and DNA precipitated from the aqueous layer with 1.5 volume of chilled ethanol. Using the procedure outlined above, the yields of DNA were in the order of 1-2 mg per kg of shoot and the molecular weight of

the DNAs was 1-2 kbp.

To obtain better yields of higher molecular weight DNA, the nuclear pellet after lysis with sodium dodecyl sulphate, was incubated with self digested protease at 37°C for 2 h. (50-100 µg/ml), followed by two deproteinization steps using freshly distilled phenol saturated with Tris-HCl (10 mM, pH 7.5) containing m-cresol and 8-hydroxyquinoline and finally with chloroform/isoamyl alcohol (24:1, v/v). DNA was precipitated with chilled ethanol. Yields were of the order of 3-4 mg per kg of shoots and the molecular weight of the DNA was > 20 kilobase pairs. Wheat DNA was isolated as above. DNA from E.coli B NCIM 2089 was isolated according to Marmur's procedure (20).

Criteria of purity of DNA preparations

(i) All the DNA preparations were checked routinely for RNA, protein and polysaccharide contamination by orcinol test (22), Lowry's assay (23) and paper chromatography of sugars obtained by the hydrolysis of polysaccharides (24) respectively. The DNA preparations had less than 1% contamination of RNA and proteins.

Detection of polysaccharides

10-15 µg of DNA precipitate was dissolved in 1N HCl and hydrolysed in a boiling water bath for 30-35 min. All the operations were carried out in a 1.5 ml capacity Eppendorf

Figure 2.1 SEPARATION OF SUGARS BY PAPER CHROMATOGRAPHY

Lane a Authentic D-ribose sugar (control)

Lane b Barn yard millet DNA hydrolysate

Lane c Fox tail millet DNA hydrolysate

Lane d Authentic D-glucose sugar (control).

The translucent spot in lane b and c
is that of deoxyribose sugar obtained
on acid hydrolysis of DNA.

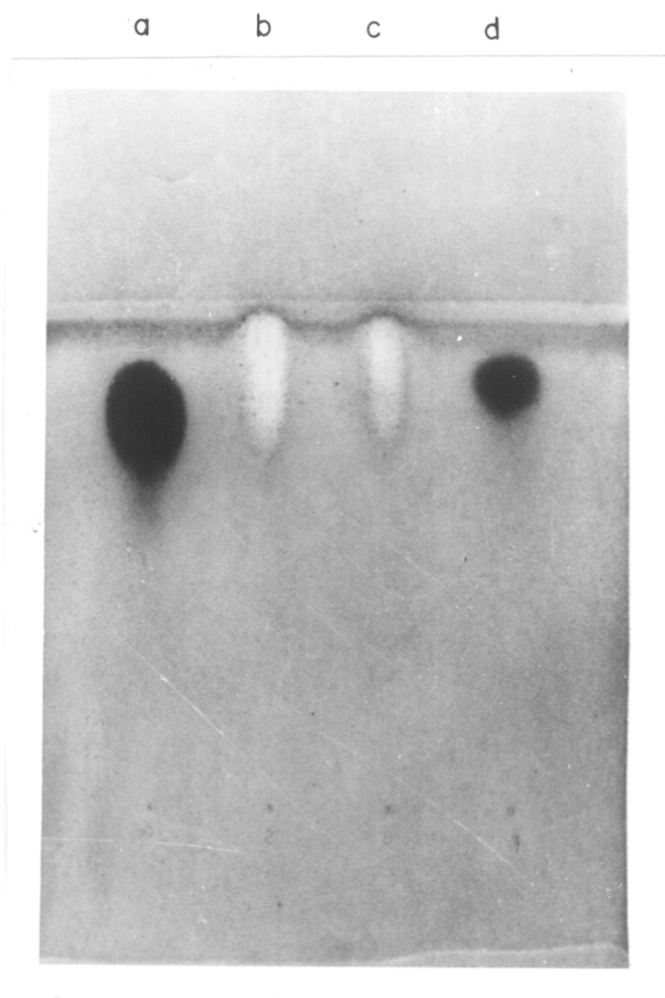


FIGURE 2 · 1

microfuge vial. After hydrolysis, a pinch of animal charcoal was added and the vial was shaken. The solution was then centrifuged and the supernatant discarded. Only sugars get adsorbed on the charcoal. The adsorbed sugars were eluted from charcoal by suspending the charcoal pellet in ethanol and collecting the supernatant by centrifugation. The ethanol containing the sugars was allowed to evaporate at room temperature to reduce the volume to 5-10 μ l. It was then chromatographed on Whatman filter paper No.1 (ascending) in N-propanol: acetic acid: water (7:2:1) system. The chamber and the paper were saturated before the run. The chromatograph was then air-dried and developed in the following sequence: (i) dipping in a saturated solution of silver nitrate in acetone and drying, (ii) dipping in 0.5% NaOH in 90% ethanol and drying and (iii) finally dipping in 0.5% sodium thiosulphate solution.

Authentic sugars like glucose and ribose were spotted as standards. The standard sugars give dark brown spots while the DNA preparations showed only deoxyribose sugar (Figure 2.1). The preparations had no detectable polysaccharide contamination.

(ii) Wavelength scanning

Wavelength scanning of the DNAs was carried out in the range of 220 to 300 nm. The absorbance of the DNA preparations at 300 nm was always less than 0.1 absorbance units.

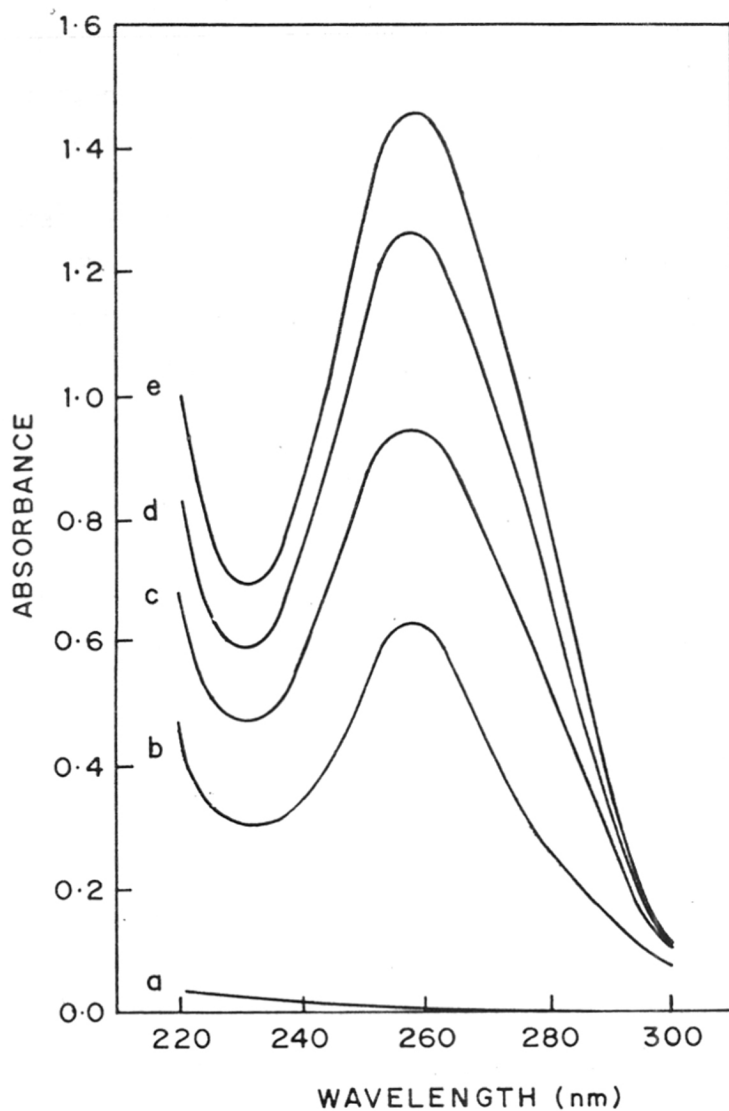


Figure 2.2 ABSORPTION SPECTRA OF MILLET DNAs.
 The spectra were obtained in the range of 220 to 300 nm. The DNA was dissolved in $T_{10}E_1$ (Tris-HCl 10 mM pH 7.5, sodium salt of EDTA 1 mM).

- a Blank $T_{10}E_1$ buffer
- b Barn yard millet DNA
- c Fox tail millet DNA
- d Little millet DNA
- e Great millet DNA

DNA preparations exhibited an optical density ratio of $A_{280}/A_{260} = 0.55$ and $A_{230}/A_{260} = 0.45$ (Figure 2.2).

(iii) Thermal denaturation

The DNAs showed approximately 25% hyperchromicity after denaturation. The presence of sharp melting profiles and absence of a 'foot' or shallow absorbance rise before the actual start of melting was indicative of the absence of single-stranded DNA and RNA (25).

DNA sizing

The average size of DNAs was determined by agarose slab gel electrophoresis using 1% agarose gels in Tris-(hydroxymethyl)-aminomethane (Tris) (40 mM) - acetic acid (20 mM) -EDTA (2 mM) at pH 8.1. Gels were stained with ethidium bromide (5 $\mu\text{g}/\text{ml}$) for 30 min. DNA bands were visualised using ultra-violet transilluminator and photographed using a Pentax SII camera fitted with a 50 mm lens. A 400 ASA, 35 mm negative film was used. Lambda phage DNA digested with Hind III and phage $\phi\times 174$ RF DNA digested with Hae III served as molecular weight markers. The DNA preparations had a molecular weight greater than 20 kilobase pairs (Figure 2.3).

Thermal denaturation

Thermal denaturation was carried out using a Gilford spectrophotometer (model No.250) equipped with thermoprogrammer (Model 2527), analog multiplexer (Model 6046) and automatic

Figure 2.3 AGAROSE SLAB GEL ELECTROPHORESIS OF
GREAT MILLET DNA WITH DNA MOLECULAR
WEIGHT MARKERS.

a Lambda DNA digested with Hind III
(23.3 kbp, 9.5 kbp, 6.4 kbp, 4.2 kbp,
2.2 kbp and 1.8 kbp).

b -

c Unsheared native great millet DNA

d Sheared (0.4 kbp) great millet DNA

e Phage ϕ x174 RF DNA digested with
Hae III (1.32 kbp, 1.078 kbp, 0.872 kbp,
0.606 kbp, 0.310 kbp).

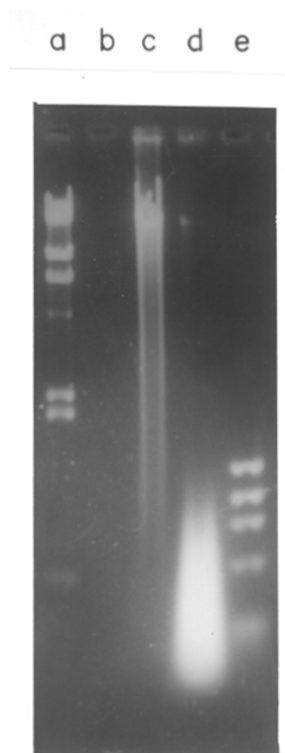


FIGURE 2.3

reference compensator (21). Thermal cuvettes were filled with 0.3 ml DNA solution in 0.12M sodium phosphate buffer, pH 6.8 (concentration 25 - 50 $\mu\text{g/ml}$) and absorbance was recorded at room temperature. Cuvettes were filled in such a way that the air space was less than 10% of the solution volume and evaporation was prevented using teflon stoppers. The temperature was then raised to 100°C at the heat rate of 1°C/min and the absorbance change occurring during the heating process from 62°C to 100°C was monitored, either with a recorder (Model 6051) or a programmed printer. The hyperchromicity of the DNA samples was calculated using the following formula:

$$H = \frac{A_{260}(100^\circ\text{C}) - A_{260}(62^\circ\text{C})}{A_{260}(100^\circ\text{C})} \quad (3,4)$$

where H is the hyperchromicity and A_{260} is the absorbance of the DNA sample at 260 nm. The total hyperchromicity was normalized to 100% and the graph of percent hyperchromicity versus temperature was plotted. The melting temperature (T_m) of the DNAs was calculated as the temperature at which half of the total increase in absorbance is obtained. The G + C content of DNAs was determined using the formula:

$$\%G + C = (T_m - 69.3) 2.44 \quad (26)$$

The transition interval (ΔT) was calculated as T_b °C (83.33% hyperchromicity) - T_a °C (16.66% hyperchromicity) (2,11). DNAs from E. coli, calf thymus and wheat were used as standards.

High resolution thermal denaturation analysis

For such analysis, DNAs were denatured at a heat rate of 0.25°C/min and absorbance values were recorded at 0.1°C increment, from 62°C to 100°C. Thus the experimental curve consisted of 300 - 400 points depending upon the melting range of DNA. Noise levels were reduced with a large slit width (1 mm) together with a high initial DNA concentration of 50-60 µg/ml. Data smoothing was done by Fourier Transform Method (19) and differentiation using the central difference formula. The derivative curves were obtained by plotting the temperature derivative of the absorbance of DNA against temperature (dA/dT vs T). To check the reproducibility of the procedure, at least five DNA preparations each in duplicate, were melted and the resulting curves were analyzed separately. Fitting 10 points at a time each indicating 0.1°C increment represents a temperature rise of 1°C which is expected to resolve individual regions melting below and above the average T_m of the DNA. The limit of resolution of this analysis was 0.001 absorbance and peaks separated by interval as narrow

as 0.4°C (approx. 1% G + C) could be detected reproducibly.

RESULTS

Yield and purity of DNA preparations

Getting good, high molecular weight DNA in sufficient amounts was an extremely difficult task in all the four millets. Though several extraction procedures were tried (27, 28), the yield was poor due to the fibrous nature of millet tissues. As a result, it was difficult to make a homogenous suspension of the ground tissue in the crushing medium. In order to overcome this, the millet tissue frozen in liquid nitrogen was first powdered. Another factor which may be responsible for the low yields of DNA could be the low DNA content in these species.

DNA was specifically extracted from nuclear pellets, instead of total tissue homogenates, to minimize the possibility of contamination by other organelle DNAs. Further, precautions were taken against bacterial contamination by carrying out thorough surface sterilization of seeds prior to germination and using only shoots for DNA extraction. The purity of the isolated samples were checked as described in Materials and Methods.

Thermal denaturation of native DNAs

Native high molecular weight DNAs (> 20 kbp) of barn yard millet, fox tail millet, little millet and great

TABLE 2.1 - THERMAL DENATURATION PROPERTIES OF NATIVE DNAs

Species	No. of experi- ments	Tm ^a Mean (°C) ± SD	G + C content ^b Mean (%) ± SD	Hyperchromicity ^c Mean (%) ± SD	ΔT °C ^d Mean (°C) ± SD
<u>E. coli</u>	6	90.83 ± 0.96	52.54 ± 2.34	21.30 ± 0.81	4.60 ± 2.31
<u>Calf thymus</u>	6	86.18 ± 0.41	41.19 ± 1.00	24.89 ± 1.68	7.36 ± 0.33
<u>Wheat</u>	6	85.57 ± 0.12	39.69 ± 0.28	24.34 ± 0.29	10.67 ± 0.23
<u>Barn yard millet</u>	9	87.48 ± 0.25	44.36 ± 0.61	23.63 ± 1.73	7.90 ± 0.27
<u>Fox tail millet</u>	10	87.78 ± 0.42	45.09 ± 1.02	23.46 ± 1.66	7.67 ± 0.29
<u>Little millet</u>	7	87.28 ± 0.35	43.87 ± 0.86	25.16 ± 0.88	8.03 ± 0.31
<u>Great millet</u>	8	87.65 ± 0.37	44.77 ± 0.91	22.02 ± 2.72	7.50 ± 0.39

^aTemperature that corresponds to 50% hyperchromicity

^bG + C content has been calculated using the formula: %G + C content = (Tm - 69.3) 2.44 (26)

^cCalculated using the formula: $H = \frac{A_{260}(100^\circ\text{C}) - A_{260}(62^\circ\text{C})}{A_{260}(100^\circ\text{C})}$ (3,4)

$A_{260}(100^\circ\text{C})$

^dCalculated using the equation: ΔT = Temperature that corresponds to 83.33% hyperchromicity temperature that corresponds to 16.67% hyperchromicity.

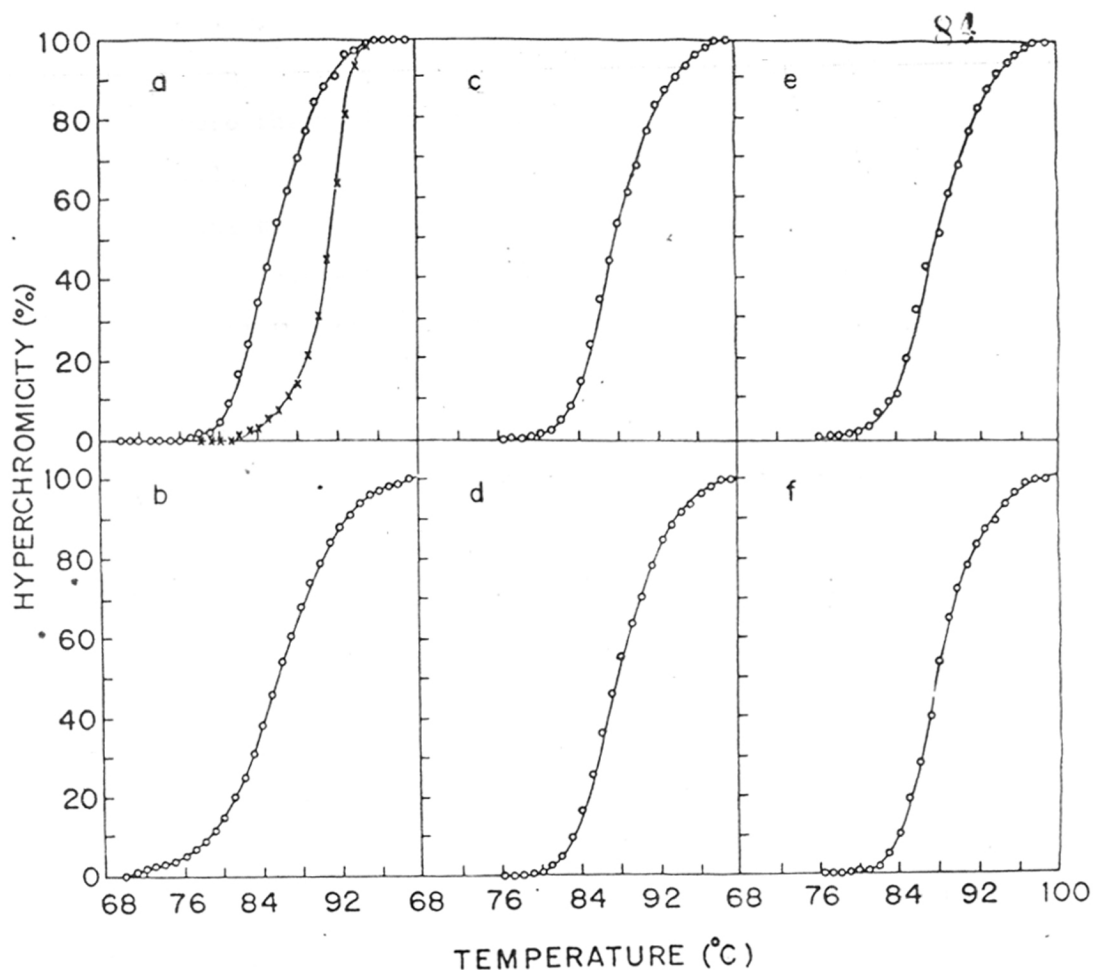


Figure 2.4 THERMAL DENATURATION PROFILES OF HIGH MOLECULAR WEIGHT DNAs (> 20 kbp)

- a Calf thymus (O—O) and E.coli (X—X) DNA
- b Barn yard millet DNA
- c Fox tail millet DNA
- d Wheat DNA
- e Little millet DNA
- f Great millet DNA

The melting of DNA was carried out in 0.12 M sodium phosphate buffer, pH 6.8 at a heat rate of 1°C/min.

millet were thermally melted at a heat rate of 1°C/min (Table 2.1). The thermal denaturation properties of E. coli, calf thymus and wheat DNAs compare well with the values reported in literature (6,10,11). The melting profiles of all the four millet DNAs are smooth and monophasic (Figure 2.4) and their melting temperatures are in the range of 87.3 to 87.8°C. The hyperchromicity values average between 22-25% and the transition width ($\Delta T_{1/2}$) in the range of 7.7 - 8.0°C. From the T_m values, the G + C content is estimated to be 43.9 - 45.1% (Table 2.1).

High resolution thermal denaturation analysis

To get a detailed information about the primary make up of the DNA and its base composition heterogeneity the DNAs were denatured at a slow heat rate of 0.25°C/min and the melting curves thus obtained were differentiated using the first derivative analysis. The derivative melting profiles (Figure 2.5) of the four millet DNAs reveal several interesting features. They show several similarities as well as distinct differences. The derivative profiles are asymmetric and skewed towards the G + C rich side. The melting starts at about 74°C and is complete by 99°C. In all, there are 11-14 components with distinct base composition. They have been arbitrarily divided into three groups. The first group includes maxima peaks near the T_m i.e. 84-88°C. This region shows 2-4 peaks and the peak of

Figure 2.5

THE FIRST DERIVATIVE MELTING PROFILES OF HIGH
MOLECULAR WEIGHT DNAs (> 20 kbp)

- | | |
|---|--|
| a | Calf thymus (————) and <u>E. coli</u> (·—·—·—) DNA |
| b | Wheat DNA |
| c | Barn yard millet DNA |
| d | Fox tail millet DNA |
| e | Little millet DNA |
| f | Great millet DNA |

The melting of DNA was carried out in 0.12 M sodium phosphate buffer, pH 6.8 at a heat rate of 0.25°C/min and the absorbance values were recorded at 0.1°C increment. The DNA concentrations were 50-60 µg/ml.

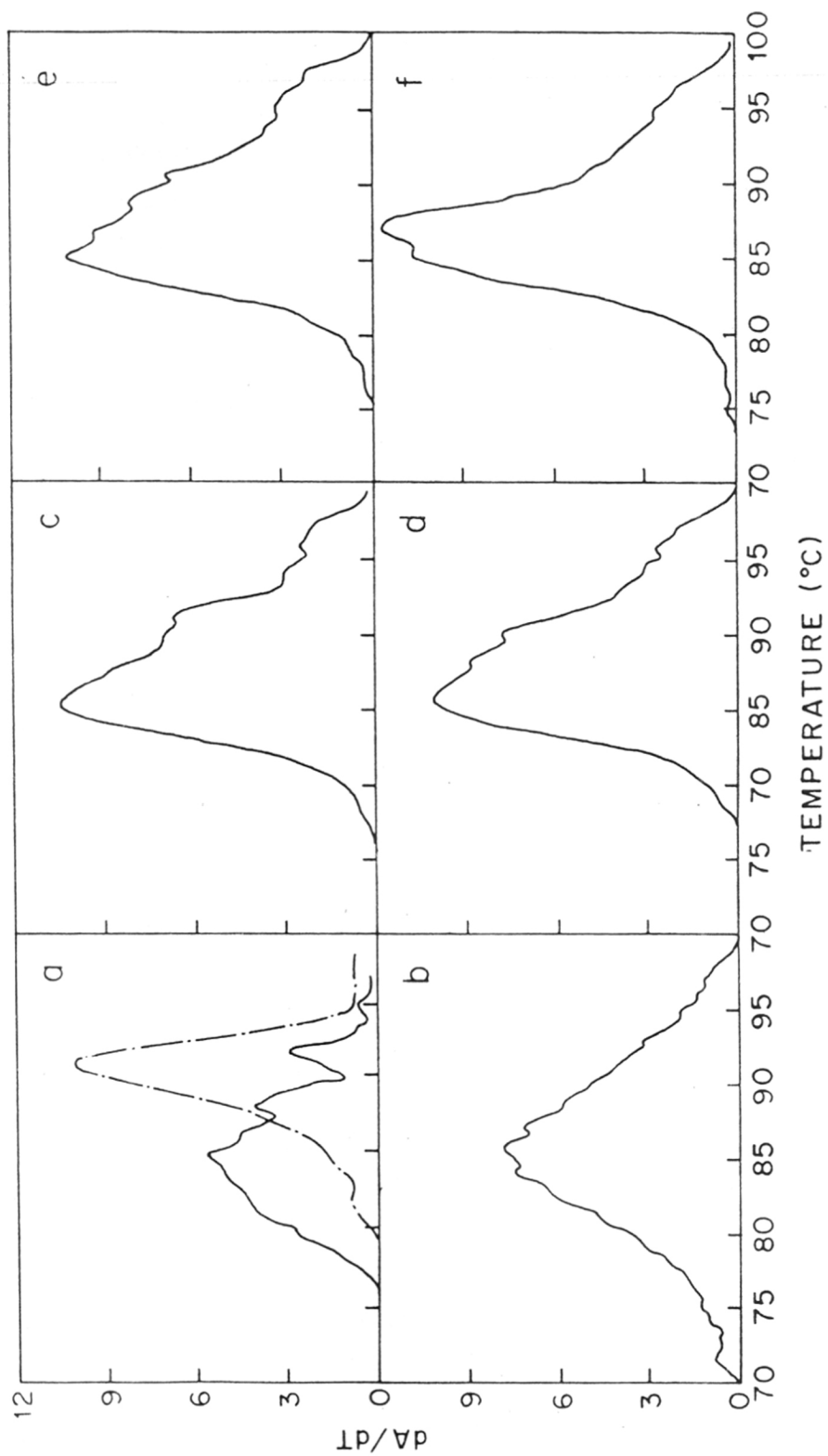


FIGURE 2.5

the total derivative profile lies in this region in all the four cases. This maxima is situated in the narrow temperature range of 85.5 - 85.7°C indicating that a significant portion of DNA has similar average melting temperature in the four millets. In the same temperature range, only great millet shows an additional prominent peak at 85.3°C and little millet at 86.6°C. The second group comprises the early melting A + T rich components and these are distinguishable only as shoulders and include components melting in the temperature range of 76-84°C. It consists of 2-4 components. The third group represents late melting G + C rich components melting at temperatures higher than 88°C. The base composition heterogeneity is specifically pronounced on this side of the profile and much of the species specific differences in these millets are observed in this region. The heterogeneity covers a temperature span of 13°C. Barn yard millet exhibits two prominent peaks in the region between 89-91°C of apparently the same proportion. Fox tail millet shows two distinct components at 88.4°C and 90.2°C and little millet at 88.9°C and 90.7°C. The proportions of these two components are, however, different. It is interesting to note that great millet does not show any distinct melting component in the region between 87-90.5°C. Finally all the millets ^{show} 3-5 distinct late melting components after 92.5°C.

A qualitative comparison of the derivative melting profiles of barn yard millet, fox tail millet, little millet and great millet DNAs was made by plotting a graph indicating the positions of the peaks and determining the number of peaks shared among the four millet species (Figure 2.6). A temperature span of 0 to 0.6°C was considered to estimate the number of shared components and these data are summarized in Table 2.2. In a temperature range of 0.6°C (1.46% G + C), four components are shared by all the four species. This number reduces to 3 and 2 at a temperature difference of 0.5°C (1.2% G + C) and 0.4°C (1% G + C) respectively. No peak occupies an identical position in all the four millets. Barn yard millet, fox tail millet and little millet which belong to the tribe Paniceae show an additional peak shared between them in the G + C side of the profile.

DISCUSSION

Precautions taken during experimentation

Since fine melting analysis is a very sensitive method, several precautions were taken to avoid artifacts. High molecular weight DNA preparations (> 20 kbp), free from detectable levels of polysaccharides, were used. The absence of a 'foot' or shallow absorbance rise before the actual start of the DNA melting indicated the absence of

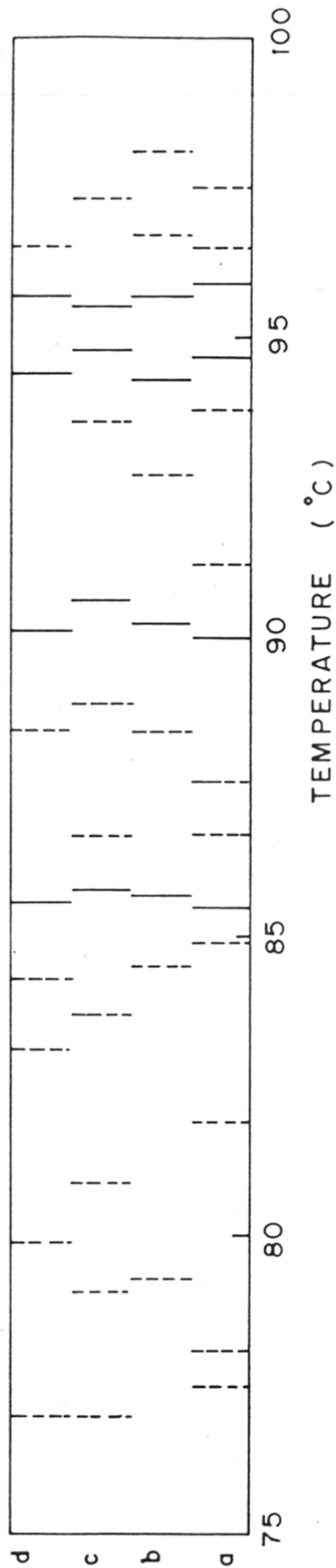


Figure 2.6 THE POSITIONS OF MELTING COMPONENTS OBSERVED IN THE DERIVATIVE MELTING PROFILES OF THE FOUR MILLETS DNAs. THE SOLID LINES INDICATE THE PEAKS SHARED BY ALL THE FOUR SPECIES IN THE TEMPERATURE RANGE OF 0.6°C

- a Barn yard millet DNA
- b Fox tail millet DNA
- c Little millet DNA
- d Great millet DNA

TABLE 2.2 - NUMBER OF MELTING COMPONENTS SHARED BY THE FOUR MILLET SPECIES

Temperature range (°C)	0.6	0.5	0.4	0.3	0.2	0.1	0.0
G + C content range (%)	1.46	1.22	0.98	0.73	0.49	0.24	0.00
No. of components shared by all four species	4	3	2	1	0	0	0
No. of components shared by millets of tribe Paniceae	5	3	2	1	0	0	0

single stranded DNA and RNA (25). Experimental conditions such as salt concentration (0.18 M Na^+) and DNA concentration ($50\text{-}60 \text{ }\mu\text{g/ml}$) were kept constant. An average of eight to ten experiments were carried out with the DNAs of each plant species and the number and position of all the components in every preparation were reproducible in all the cases. Finally, the first derivative analysis of DNAs of *E. coli* calf thymus and wheat, which served as controls, compared well with the reported values (6,10,11).

Comparison of DNA melting temperatures of Gramineae plant species.

The millets under investigation showed an average melting temperature in the range of $87.3 - 87.8^\circ\text{C}$. Earlier reports from our laboratory have shown the Tms of finger millet, pearl millet and rice as 85.0°C , 88.6°C and 85.0°C respectively (29). Other members of the family Gramineae, namely, wheat, oat, rye, barley and maize have Tms in the range of $85 - 87^\circ\text{C}$ (30-32). Thus, it appears that the millets, in general, have higher melting temperatures than the other grasses. Finger millet appears to be an exception.

Comparison of high resolution thermal denaturation profile of millets

Each of the four millet species has a polyphasic derivative melting profile (Figure 2.5). Such profiles in eukaryotes are considered to be at least in part due to the

presence of satellite and repetitive families. The bacteriophage DNAs are made of unique sequences and show the presence of 'thermalites' melting in very narrow intervals (33). As the genome size increases, the overlap between the neighbouring sequences increases and this results in smooth, monophasic melting profiles as observed in bacterial DNAs which are also composed of only unique sequences (6). This should be true for all higher plant and animal DNAs if they contained only unique sequences. The multiphasic derivative profiles that we observe are probably due to the presence of families of repeated DNA sequences in these genomes. Each repetitive DNA family should in general have a distinct base composition (10) and hence melt as a 'block'. Such 'blocks' are resolvable using the high resolution thermal denaturation technique, even if they are interspersed with other DNA sequences. Considering that the present technique allows only the qualitative determination of such repeated families it appears that the repeated sequences in these millets do differ to some extent.

A qualitative comparison of the derivative melting profiles of the four millet genomes indicates that the number of shared peaks reduces as the ^{temperature} range of comparison is narrowed down. Not a single melting component occupies an identical position in all the four millets. Several

distinct species-specific characters are observed as described in the results. Moreover, the G + C rich sequences melt over a broader range of temperature (13°C) as compared to that of the A + T rich sequences, indicating that the former are more heterogenous in base composition than the latter. Great millet differs from the other millets in having an additional prominent peak at 85.3°C and no peaks in the region between $87-90.5^{\circ}\text{C}$. Barn yard millet, fox tail millet and little millet show two prominent peaks in the region between $88.4 - 91^{\circ}\text{C}$ but their proportions and positions, however, are different. Little millet, in addition, shows a peak at 86.6°C . These differences may be assumed to be responsible for the speciation events in these millets.

In the present study, we have only made a qualitative comparison of the derivative melting profiles of the four millet DNAs. It is necessary to quantitate the results before one can comment on species relatedness. Hence preliminary attempts were made to resolve the melting profiles into individual Gaussian components and to quantitate the proportions of different melting components shared among the different millets. However, due to the large number of peaks and the tremendous overlap of areas it has not been possible to develop a programme suitable for this resolution. Programmes developed by Dr. R.D. Blake

(Department of Biochemistry, University of Maine, Orono, USA), C.E. Cuellar (Carnegie Institution of Washington) and A. Wada (Department of Physics, Faculty of Science, University of Tokyo, Japan) have been useful to resolve derivative profiles with 5-6 melting maxima. It was also suggested (personal correspondence) that it is not possible (and statistically incorrect) to resolve profiles of higher organisms characterized with a large number of peaks. Attempts are being made to present an overlap of derivative melting profiles and manually calculate the area of shared and unshared peaks in different millet species.

REFERENCES

- 1 Marmur, J. and Doty, P. (1962). J. Mol. Biol.,
5: 109-118.
- 2 Mahler, H.R. and Dutton, G. (1964). J. Mol. Biol.,
10: 157-175.
- 3 Zimmerman, J.L. and Goldberg, R.B. (1977). Chromosoma
(Berl.), 59: 227-252.
- 4 Wimpee, C.F. and Rawson, J.R.Y. (1979). Biochim.
Biophys. Acta, 562: 192-206.
- 5 Vizard, D.L. and Ansevin, A.T. (1976). Biochemistry,
15: 741-749.
- 6 Blake, R.D. and Lefoley, S.G. (1978). Biochim. Biophys.
Acta, 518: 233-246.
- 7 Tachibana, H., Wada, A., Gotoh, O., and Takanami, M
(1978). Biochim. Biophys. Acta, 517: 319-328.
- 8 Lyubchenko, Yu, L., Frank-Kamenetskii, M.D.,
Vologodskii, A.V., Lazurkin, Yu, S. and Gause, G.G.
Jr. (1976). Biopolymers, 15: 1019-1036.
- 9 Ansevin, A.T., Vizard, D.L., Mandel, M.L. and
Dean, D.H. (1983). Biochim. Biophys. Acta, 740:
127-133.
- 10 Mayfield, J.E. (1977). Biochim. Biophys. Acta, 477:
97-101.
- 11 Huguet, T. and Jouanin, L. (1972). Biochim. Biophys.
Acta, 262: 431-440.
- 12 Pivec, L., Horska, K., Vitek, A. and Deskocil, J.
(1974). Biochim. Biophys. Acta, 340: 199-206.
- 13 Szecsi, A. (1979) Acta Microbiol. Acad.Sci. Hung.,
26: 337-347.
- 14 Szecsi, A. and Dobrovolsky, A. (1980). Nucleic
Acids Res., 8: 2517-2525.
- 15 Guttman, T., Vitek, A. and Pivec, L (1977). Nucleic
Acids Res. 4: 285-297.
- 16 Yabuki, S., Gotoh, O. and Wada, A. (1975). Biochim.
Biophys. Acta, 395: 258-273.

- 17 Steinert, M. and Van Assel, S. (1974). *Biochim. Biophys. Res. Commun.*, 61: 1249-1255.
- 18 Bhawe, M., Lagu, M., Gadre, S.R. and Ranjekar, P.K. (1984). *Indian J. Biochem. Biophys.*, 21: 81-84.
- 19 Cole, R.H. and Winsor, P. (1982). In: Fourier, Hadamard and Hilbert Transforms in Chemistry, pp.183-206, Marshall, A.G. ed., Plenum Publishing Corp., New York.
- 20 Marmur, J. (1961). *J. Mol. Biol.*, 3: 208-218.
- 21 Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976). *Biochim. Biophys. Acta*, 425: 30-40.
- 22 Volkin, E. and Cohn, W.E. (1954). In: *Methods of Biochemical Analysis*. Glick, D. ed. Vol.1, pp.287-303, Interscience Publishers Inc., New York.
- 23 Lowry, O.H., Rosenbrough, N.J., Farr, A.L. and Randall, R.L. (1951). *J. Biol. Chem.*, 193: 265-275.
- 24 Darby, G.K., Jones, A.S., Kennedy, J.F. and Walker, R.T. (1970). *J. Bacteriol.*, 103: 159-165.
- 25 Walbot, V and Goldberg, R. (1979). In: *Nucleic Acids in Plants*, Hall, T.C. and Davies, J.W. eds. Vol.1, pp.3-40, CRC Press, Florida.
- 26 Mandel, M. and Marmur, J. (1968). In: *Methods in Enzymology*. Grossman, L. and Moldave, K. Eds. Vol. 12B, pp.195-206, Academic Press, New York.
- 27 Bendich, A.J. and Bolton, E.T. (1967). *Plant Physiol.* 42: 959-967.
- 28 Flavell, R.B., Bennett, M.D., Smith, J.B. and Smith, D.B. (1974). *Biochem. Genet.*, 12: 257-269.
- 29 Deshpande, V.G. and Ranjekar, P.K. (1980). *Hoppe-Seyler's Z. Physiol. Chem.*, 361: 1223-1233.
- 30 Bendich, A.J. and McCarthy, B.J. (1970). *Genetics*, 65: 545-565.
- 31 Ranjekar, P.K., Lafontaine, J.G. and Pallotta, D. (1974). *Chromosoma (Berl.)*, 48: 427-440.
- 32 Hake, S. and Walbot, V. (1980). *Chromosoma (Berl.)*, 79: 251-270.
- 33 Wada, A., Yabuki, S. and Husimi, Y. (1980). In: *CRC Critical Reviews in Biochem.*, 9: 87-144.

CHAPTER III

NON-LINEAR LEAST SQUARES REGRESSION
ANALYSIS OF MILLET DNAS

SUMMARY

Employing the non-linear regression analysis method, reassociation kinetics of the DNAs of four millet plant species have been analyzed to determine (i) the number of DNA components reassociating with second order kinetics; (ii) proportion of repeated and single copy DNA sequences; (iii) frequency of repetition and kinetic complexity of different classes of DNA. The repeated DNA sequences account for 26 - 52% of the total DNA. In fox tail millet, the repetitive DNA content of 30% is the lowest reported so far in the family Gramineae. The proportion of unresolvable rapidly reassociating DNA is 5.4 - 6.4% in barn yard millet and great millet and 14.7 - 17.9% in fox tail millet and little millet. The fast repetitive component is present only in barn yard millet and great millet. Its kinetic complexity and reiteration frequency are different in the two species. The slow repetitive component present in all the four millet species exhibits a low reiteration frequency of 91 - 288 and a kinetic complexity of 2.84×10^4 - 3.40×10^6 nucleotide pairs. The single copy DNA sequences have a kinetic complexity of 3.87×10^7 - 6.81×10^8 nucleotide pairs.

The species specific variations in proportions, frequencies of reiteration and kinetic complexities of different repetitive classes of DNA may contribute to the speciation events in millet genomes.

Based on the literature data in other Gramineae plant species and the present data on millet genomes, it is observed that the contents of repetitive as well as single copy DNA sequences increase with the increase in nuclear DNA content. However, ~~above~~ the nuclear DNA content of about 6 picograms, the content of single copy DNA remains more or less constant while that of repetitive DNA continues to increase. It, therefore, appears that repeated DNA contributes significantly to the genome size of these plant species.

INTRODUCTION

The technique of reassociation kinetics has proved to be a very powerful tool in elucidating the primary make up of a genome. These studies involve separation of the two complementary strands of DNA and measurement of the progress of strand reassociation as a function of DNA concentration and time - defined as Cot . The reassociation of DNA can be monitored by various methods such as hypochromic shift in UV absorbance (optical reassociation), hydroxyapatite column chromatography and S1 nuclease assay

(1 - 3). Using the above techniques, a curve can be plotted relating percent reassociation of denatured DNA to different Cot values. Such Cot curves can be resolved into discrete second order components, with distinct rate constants employing the non-linear least squares regression analysis. The fraction of the genome composed of repetitive and single copy sequences can be estimated. Parameters like proportion, kinetic complexities and reiteration frequencies of various repetitive components can also be calculated.

In the family Gramineae, the DNAs of only a few temperate grasses have been characterized with respect to the proportion of repetitive DNA content in their genomes. Wheat, oat, rye and barley are reported to have 80% repetitive DNA (4 - 7). Similar observations were also made in Avena sativa, Poa trivialis and Poa annua (8). The maize genome, however, revealed the presence of 60% repeated DNA sequences (9).

All the aforementioned Gramineae plant species have a 1C nuclear DNA content in the range of 4 - 21.5 pg. The work was initiated in our laboratory involving the characterization of plant genomes in the family Gramineae with a 1C nuclear DNA content less than 5 pg. The occurrence of a rather low proportion of repetitive DNA has been shown in finger millet, pearl millet and rice

(49-54%) (10). The 1C nuclear DNA content of these species lie in the range of 0.6 - 2.5 pg. These low DNA containing species have also revealed novel DNA sequence organization patterns (11-13). Based on these data, it was proposed that nuclear DNA content could play an important role in the structure and organization of DNA sequences. In order to verify the proposed hypothesis, characterization of four additional Gramineae plant species with 1C nuclear DNA content of less than 5 picograms was carried out. This Chapter describes the reassociation kinetics of the nuclear DNAs of four millet species namely barn yard millet, fox tail millet, little millet and great millet in order to determine (i) how many DNA components, each reassociating with second order kinetics are present in the total DNA; (ii) what is the proportion of total repetitive DNA; (iii) what is the frequency of repetition and kinetic complexity of different repetitive DNA components.

MATERIALS AND METHODS

DNA isolation

DNAs were isolated from the seedlings of barn yard millet, fox tail millet, little millet and great millet by the procedure outlined in Chapter II.

DNA content determination

The DNA content in all the four millets was determined according to the standard procedure of Feulgen cytophotometry (14).

DNA shearing and sizing

Native high molecular weight DNAs, dialyzed against 0.12 M sodium phosphate buffer (pH 6.8) were fragmented using a Sonic Oscillator (Bronwell Model, Biosonik III, 250W, 20 KHz) fitted with a half inch probe (Green) to a modal length of 0.4 kbp. Prior to sonication, nitrogen gas was bubbled through the DNA solution for 5-10 mins. (15). Approximately 30 ml of DNA solution was sonicated at a maximum intensity for 3 mins; with 10 mins. interval between each successive one minute pulse of sonication. Care was taken to see that the temperature of the solution did not rise above 4°C during the sonication.

The average size of the DNA fragments after shearing was determined by agarose slab gel electrophoresis as described in Chapter II and was found to be approximately 0.4 kbp.

Hydroxyapatite preparation

Hydroxyapatite was prepared essentially according to the procedure of Tiselius et al. (16) and stored at 4°C with a few drops of chloroform. The hydroxyapatite

preparations usually formed columns with good flow characteristics. The extent of DNA recovery from each batch of hydroxyapatite was checked by loading known quantities of DNA. Recovery was always more than 95%.

DNA reassociation kinetics

The major steps in a typical reassociation experiment involve (i) strand separation by heat denaturation, and (ii) reassociation of complementary strands being initiated by rapidly cooling the sample to a predetermined appropriate temperature. A proper choice of incubation temperature (62°C , $T_m - 25^{\circ}\text{C}$) and salt concentration (0.12M sodium phosphate buffer, pH 6.8; Na^+ 0.18M) provides optimum conditions for reassociation. Two techniques were used to measure the progress of reassociation:

(i) Optical reassociation: Since strand separation produces a hyperchromic shift in absorbance, which is reversed as base pairs reform, reassociation may be followed by monitoring the decrease in absorbance (also known as hypochromic shift) as a function of time.

DNA samples, 0.4 kbp long in 0.12 M sodium phosphate buffer (pH 6.8), were first denatured in a Gilford 250 spectrophotometer equipped with a thermoprogrammer (Model No.2527), analogue multiplexer (Model No.6046) and automatic reference compensator. Thermal cuvettes were filled with

approximately 0.3 ml of DNA solution (25-50 μ g of DNA) and temperature of the solution was raised to 100°C. The absorbance change occurring during the heating process from 62° to 100°C was monitored at 260 nm. The hyperchromicity of the DNA sample was calculated as described in Chapter II. The denatured samples were maintained at 100°C for 4-5 min. to ensure complete denaturation and then quickly cooled to 62°C by the circulation of cold water. The drop in temperature to 62°C was achieved in 60-90 seconds. The change in the absorbance of the DNA was continuously recorded with time. For the determination of zero time or the start of reassociation, the mid point (80°C) between 100° and 62°C was considered (2). The Cot value at different time intervals is defined as half the product of absorbance of dissociated DNA x hours of incubation i.e. $Cot = \frac{A}{2} \times t$ (h) (17, 18). Thus a Cot value of one is equivalent to incubation of a DNA solution of concentration 83 μ g/ml (approx. 2 O.D.) for one hour (19). The percentage reassociation at different Cot values was calculated using the formula:

$$\% R = \frac{A_{260} \text{ of denatured DNA} - A_{260} \text{ at a given time}}{A_{260} \text{ of denatured DNA} - A_{260} \text{ of native DNA at } 62^{\circ}\text{C}} \times 100$$

Reassociation was monitored optically in the Cot range of 1×10^{-3} to 1×10^1 M.s. approximately.

(ii) Hydroxyapatite column chromatography: The hydroxyapatite slurry was poured into a double jacketed glass column (15 x 2.5 cms) to a bed volume of approximately 15 ml. The column was equilibrated with 0.12 M sodium phosphate buffer (pH 6.8) and maintained at 62°C with a thermostatic circulating water bath.

The hydroxyapatite chromatography was essentially according to Kohne and Britten (18) and Britten et al. (2). Sonicated DNA samples ranging in concentration from 30 $\mu\text{g/ml}$ for low Cot values to 300 $\mu\text{g/ml}$ for higher Cot values in 0.12 M sodium phosphate buffer (pH 6.8) were sealed in boiling tubes and denatured by boiling for 10 minutes at 100°C. These were then incubated to specific Cot values ^{pre-equilibrated} at 62°C. The samples were loaded on columns/ with 0.12M sodium phosphate buffer (pH 6.8) and maintained at 62°C. Single stranded material was eluted with 0.12 M sodium phosphate buffer (pH 6.8) and double stranded DNA with 0.4 M sodium phosphate buffer (pH 6.8). The relative amounts of single and double stranded DNA were determined by measuring absorbance of the fractions at 260 nm on a Shimadzu double beam spectrophotometer (Model UV-210 A) assuming that one absorbance unit (A_{260}) is equal

to 50 μg of DNA/ml. The extent of reassociation was calculated using the formula:

$$\% R = \frac{\mu\text{gms of DNA eluted with } 0.4 \text{ M phosphate buffer (pH 6.8)}}{\text{total } \mu\text{gms of DNA eluted from the HAP column}} \times 100 \dots (19)$$

In experiments at high Cot values ($> 1 \times 10^2$ M.s.), the DNA samples were incubated in 0.36 M sodium phosphate buffer (pH 6.8) in order to accelerate the rate of reassociation and thus reduce incubation time. Samples reassociated in higher salt concentrations were diluted to 0.12 M sodium phosphate buffer (pH 6.8) with double distilled water prior to loading onto the hydroxyapatite column. A corresponding correction was applied to the reassociation rate using the factor 4,4778 (2). The fraction remaining single stranded were plotted as a function of Cot on a semilogarithmic paper. The data points obtained optically and by hydroxyapatite column chromatography were fitted together for the construction of the Cot curve.

Reassociation kinetics of E. coli DNA, mouse liver DNA and wheat DNA studied under identical conditions served as internal controls.

Nonlinear least squares regression analysis of reassociation kinetics data

The reassociation of a eukaryotic DNA commonly

occurs over a range of Cot values spanning upto eight orders of magnitude or even more. Since the individual components of the reaction usually are not well separated, a computer fitted programme was used to generate a set of Cot curves (whose sum fits the experimental data) to resolve them. Individual curves represent single kinetic components and are fitted to the second order equation by least squares analysis of the data and thus can be modelled with two or three ideal components. These components may not exist as discrete entities in the genome but are merely simplified mathematical descriptions of the data. The programme is usually run on the basis of finding the minimum number of curves required to fit the data (20).

Since DNA reassociation follows second order kinetics, it can be defined using the equation $dC/dt = -KC^2$ where C is the concentration of single stranded sequences and t is the incubation time (2). This yields the equation $C/C_0 = 1/1 + K Cot$ where C_0 is the initial concentration of single stranded DNA in moles of nucleotides per liter i.e. $C = C_0$ when $t = 0$. Least squares analysis of the experimental data was carried out by employing a standard optimization subroutine STEPIT (written by J.P. Chandler and distributed by QCPE, Indiana

University, Bloomington, Indiana, USA). The computer fitted curves are drawn as a solid line through the data points and the dotted lines represent kinetics of individual components if they existed alone. The RMS error is given by the equation:

$$\text{Error} = \sum \left[\left(\frac{C}{C_0} \right)_{\text{expt}} - \left(\frac{C}{C_0} \right)_{\text{fit}} \right] \quad \dots\dots 1$$

and was minimized by a direct search procedure. The form for $\frac{C}{C_0}$ was assumed to be

$$\frac{C}{C_0} = T + \sum \frac{F_i}{1 + K_i C_0 t} \quad \dots\dots 2$$

where $\frac{C}{C_0}$ is the fraction denatured,

T the fraction of the genome that fails to reassociate

F_i the fraction of the genome reassociating in the i^{th} component

K_i the reassociation rate for the i^{th} component.

The parameters T, F_i and K_i are all allowed to free float to minimize the error in equation 1.

The computer used for such curve fitting procedures was ICL 1900 series.

RESULTS

DNA content

The haploid nuclear DNA content was shown to be 2.70, 0.80, 1.00 and 4.60 pg in barn yard millet, fox tail millet, little millet and great millet respectively.

Reassociation kinetics

Precautions taken during reassociation kinetics studies

(i) Purity of DNA

Since the presence of contaminating proteins, carbohydrates and organelle DNA can interfere in the reassociation kinetics studies, care was taken to remove such contaminating factors (Chapter II). Samples were routinely checked for their nativity to ensure the absence of any single stranded material.

(ii) Fragment length of DNA

In the present study, the reassociation was monitored by two methods. In the optical method, we observe the time dependent decrease in absorbance at 260 nm of DNA. Here the hypochromicity is assumed to vary linearly with the restored base pairing (2). This method thus measures the actual amount of DNA helix that is formed. The other method is to observe the amount of DNA which binds to hydroxyapatite as a function of incubation period. However, in this method, measurement of reassociation can lead to an overestimate

of double stranded DNA even though DNA fragments may be paired along only some part of their length. The use of very small fragments of DNA, therefore, can minimize the problem and provide a better estimate of the extent of reassociation at a given Cot value. Hence, DNA fragments with an average size of 0.4 kbp were used. Recovery of DNA from hydroxyapatite at this fragment length was 100% indicating the absence of hyperpolymer formation.

(iii) Salt, temperature and concentration of DNA

The rate of reassociation is dependent on the concentration of the complementary DNA strands in solution, temperature of incubation and salt concentration. Extreme care was taken to maintain the proper conditions of reassociation. Since some of the above conditions are interdependent, DNA concentrations were kept in the range of 30 - 300 $\mu\text{g/ml}$ to avoid viscosity effects on the rate of reassociation (24). Salt concentration was maintained at 0.18 M Na^+ which is the optimum for reassociation. The optimum reassociation rate lies at a broad plateau at about 25°C below the melting temperature of DNA. We, therefore, carried out reassociations at 62°C since the T_m values of the millet DNAs average between 87-88°C (Chapter II).

Figure 3.1 REASSOCIATION KINETICS OF 0.4 kbp LONG MILLET DNAs AND OF *E.coli*, WHEAT AND MOUSE LIVER DNAs. THE LATTER THREE DNAs SERVED AS CONTROLS. THE DATA POINTS ARE OBTAINED BY BOTH OPTICAL AND HYDROXYAPATITE REASSOCIATION METHODS. THE SOLID LINE THROUGH THE DATA POINTS REPRESENTS THE BEST LEAST SQUARES FIT, ALLOWING ALL THE PARAMETERS TO FREE FLOAT. THE LOWER DASHED CURVES REPRESENT THE PREDICTED REASSOCIATION KINETICS OF PURE COMPONENTS.

- a Wheat DNA (O—O), RMS error = 0.032
and *E.coli* DNA (X—X), RMS error = 0.042.
- b Mouse liver DNA, RMS error = 0.043.
- c Barn yard millet DNA, RMS error = 0.043
- d Fox tail millet DNA, RMS error = 0.043
- e Little millet DNA, RMS error = 0.041
- f Great millet DNA, RMS error = 0.027

All the DNA reassociation experiments were carried out at 62°C and the DNA concentrations used were 20 - 50 $\mu\text{g/ml}$ upto a Cot value of 5.0 M.s. and 200 - 300 $\mu\text{g/ml}$ in experiments higher than a Cot value of 5.0 M.s. Reassociation experiments at Cot values greater than 100 M.s. were carried out in 0.36 M sodium phosphate buffer, pH 6.8 and the corresponding correction to the rate was applied (2).

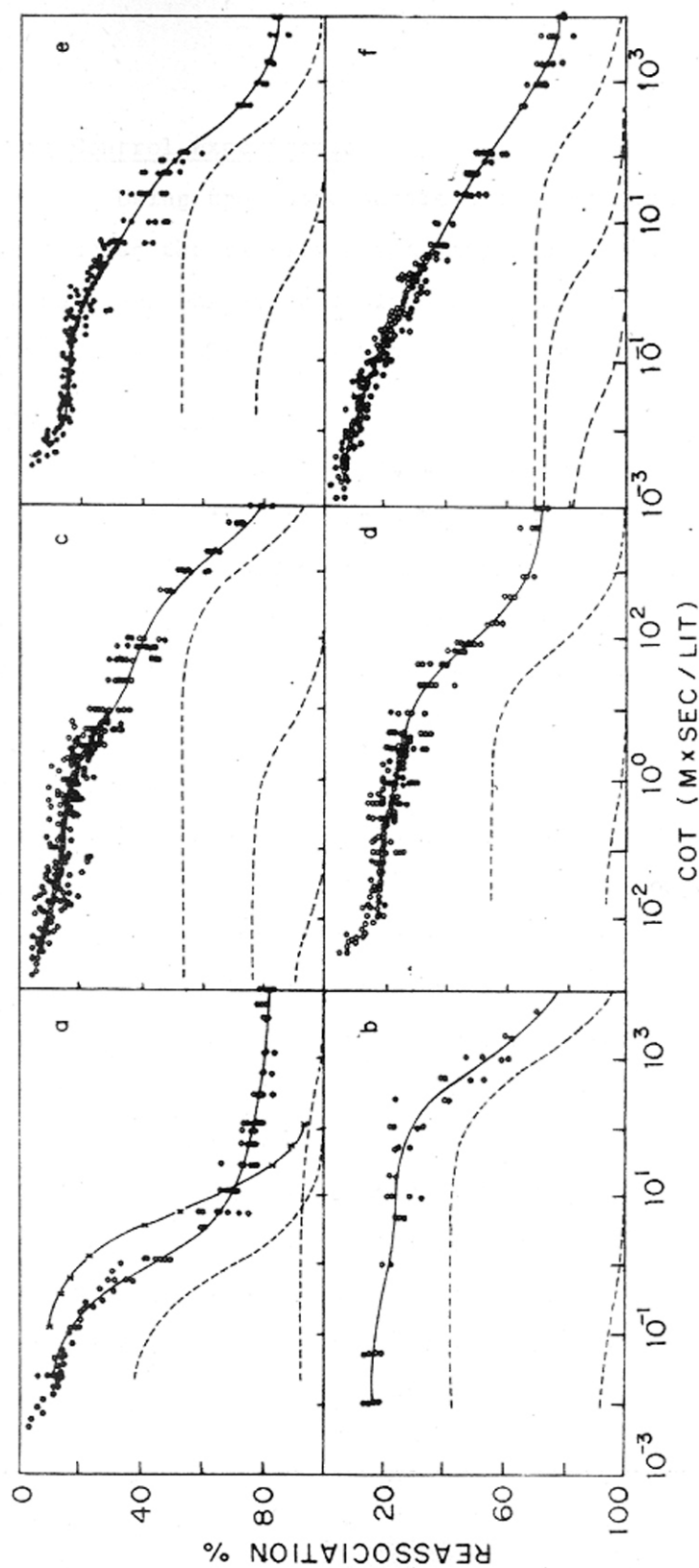


FIGURE 3.1

(iv) Control experiments

Using the above parameters for optimum reassociation, we studied the reassociation kinetics of wheat, E. coli, and mouse liver DNA as controls. Our results with these DNAs compare well with those reported in the literature (4,8,22-24).

Present analysis

Non-linear least squares regression analysis of the reassociation data obtained by optical measurements and hydroxyapatite column chromatography helped in determining the number of components reassociating with second order kinetics in the four millet DNAs. The results of such analysis are summarized in Table 3.1, and depicted in Figure 3.1.

Barn yard millet

The Cot curve of barn yard millet DNA could be modelled into three distinct second order components differing in their reiteration frequencies. The first component, comprising 6% of the genome, reassociates instantaneously ($Cot < 1.4 \times 10^{-3}$ M.s.) and cannot be resolved with the techniques employed in the present investigation. The next is a very fast reassociating component which accounts for 13% of the genome with a kinetic complexity of 8.75×10^2 nucleotide pairs and a high frequency of reiteration of 2×10^5 . The slow component comprising 24% of the genome has a kinetic complexity of 1.23×10^6 nucleotide pairs

and a repetition frequency of 2.9×10^2 . All the repeated sequences (42%) in barn yard millet reassociate by a Cot value of 7.5×10^1 M.s. Though approximately 45% of the genome reassociates with a rate constant of 1.25×10^{-3} indicating the presence of mainly single copy sequences, about 14% of the genome fails to reassociate at a Cot value of 1×10^4 M.s.

Fox tail millet

Besides the rapidly reassociating repetitive component (Cot $< 1.8 \times 10^{-2}$ M.s.) which constitutes 18% of the total genome, the Cot curve of fox tail millet shows one repetitive and one single copy component. The repetitive component represents 8% of the genome with an estimated kinetic complexity of 2.8×10^4 nucleotide pairs and a reiteration frequency of 2.4×10^2 . All the repeats (26%) of fox tail millet reassociate by a Cot value of 5×10^0 M.s. Single copy DNA reassociates with a rate constant of 2.23×10^{-2} . As much as 28% of the genome fails to reassociate at a Cot value of 1×10^4 M.s.

Little millet

The Cot curve of little millet is similar to that of fox tail millet and can be modelled into one repetitive and one single copy component in addition to the rapidly reassociating repetitive component. The latter comprises 14.7% of the genome and includes DNA sequences reassociating

TABLE 3.1 - KINETIC ANALYSIS OF MILLET DNAs

1.	2	3	4	5	6	7
<u>Barn yard millet</u>						
Very fast						
Cot < 1.40×10^{-3}	0.054	-	-	-	-	-
Fast						
Cot 1.40×10^{-3} to 5.00×10^{-1}	0.123	8.36×10^{-3}	1.03×10^{-3}	9.71×10^2	8.75×10^2	2.05×10^5
Slow						
Cot 5.00×10^{-1} to 1.00×10^2	0.244	5.93×10^0	1.45×10^0	6.90×10^{-1}	1.23×10^6	2.88×10^2
Single copy						
Cot 1.00×10^2 to 1.00×10^4	0.469	1.71×10^3	8.02×10^2	1.25×10^{-3}	6.81×10^8	1.00×10^0
Fraction unassociated	0.137	-	-	-	-	-
Cot > 1.00×10^4						
<u>Fox tail millet</u>						
Very fast						
Cot < 1.80×10^{-2}	0.179	-	-	-	-	-
Slow						
Cot 1.80×10^{-2} to 5.00×10^0	0.082	4.07×10^{-1}	3.34×10^{-2}	2.99×10^1	2.84×10^4	2.40×10^2

113

.....Contd.

TABLE 3.1 (Contd.)

1	2	3	4	5	6	7
Single copy Cot 5.00×10^4 to 1.00×10^5	0.455	9.85×10^1	4.48×10^1	2.23×10^{-2}	3.82×10^7	1.00×10^0
Fraction unreassociated Cot $> 1.00 \times 10^4$ Little millet	0.289	-	-	-	-	-
Very fast Cot $< 1.80 \times 10^{-2}$ Slow	0.147	-	-	-	-	-
Cot 1.80×10^{-2} to 2.50×10^1	0.235	2.05×10^0	4.81×10^{-1}	2.08×10^0	4.08×10^5	9.61×10^1
Single copy Cot 2.50×10^4 to 1.00×10^5	0.478	1.97×10^2	9.41×10^1	1.06×10^{-2}	7.99×10^7	1.00×10^0
Fraction unreassociated Cot $> 1.00 \times 10^4$	0.143	-	-	-	-	-
Great millet Very fast Cot $< 1.00 \times 10^{-3}$	0.0635	-	-	-	-	-
Fast Cot 1.00×10^{-3} to 5.00×10^{-1}	0.1736	5.17×10^{-2}	8.97×10^{-3}	1.12×10^2	6.94×10^3	7.62×10^3
Slow Cot 5.00×10^{-1} to 2.00×10^1	0.2741	3.92×10^0	1.08×10^0	9.26×10^{-1}	3.40×10^6	9.17×10^1

.....Contd.

TABLE 3.1 (Contd.)

1	2	3	4	5	6	7
Single copy						
Cot 2.00×10^4 ¹ to 1.00×10^4	0.3103	3.59×10^2	1.11×10^2	8.97×10^{-3}	3.05×10^8	1.00×10^0
Unreassociated						
Cot $> 1.00 \times 10^4$	0.1823	-	-	-	-	-

^aValues obtained from the respective reassociation curves (Fig. 3.1)

^bCot 1/2 values observed x fraction of genome

^cK pure = (Cot 1/2 pure)⁻¹

^dValues obtained from the standard relationship for E. coli DNA where Cot 1/2 is 5.3 M.s (under our experimental conditions) and genome size is 4.5×10^6 nucleotide pairs.

^eCot 1/2 values of unique DNA divided by Cot 1/2 values of the given fraction.

by a Cot value of 1.8×10^{-2} M.s. The slow repetitive component has a kinetic complexity of 4.08×10^5 nucleotide pairs and a reiteration frequency of 9.6×10^1 and represents 23.5% of the genome. All the repeated DNA sequences (38%) reassociate by a Cot value of 1.5×10^1 M.s. The single copy component reassociates with a rate constant of 1.06×10^{-2} and approximately 14% of the genome fails to reassociate at a Cot of 1×10^4 M.s.

Great millet

In the great millet genome, 6% of the DNA reassociates very rapidly by Cot 1×10^{-3} M.s. The fast repetitive component, comprising 17.4% of the genome, has a kinetic complexity of 6.9×10^3 nucleotide pairs and a reiteration frequency of 7.62×10^3 . The kinetic complexity and frequency of repetition of the slow repetitive component is 3.40×10^6 ^{nucleotide pairs} and 9.17×10^1 respectively and it constitutes 27.4% of the genome. In great millet, the repeat sequences (52%) reassociate by Cot 2.0×10^1 M.s. Approximately 31% of the genome contains sequences which reassociate with a rate constant of 8.97×10^{-3} indicating predominantly single copy sequences. The remaining 18% of the genome fails to reassociate at a Cot value of 1×10^4 M.s.

Sequences reassociating by Cot 1.5×10^{-3} M.s. in case of barn yard millet and great millet and those reassociating by Cot 1.8×10^{-2} M.s. in case of fox tail

millet and little millet may consist of one or more classes of highly repeated DNA sequences and/or fold back sequences which reassociate instantaneously (25). The fraction of the millet genome which fails to reassociate at a Cot of 1×10^4 M.s. has been attributed to a number of reasons such as degradation resulting from long incubations, failure of some sequences to form stable duplexes at the criterion used for these experiments and the presence of sequences requiring a longer time to reassociate (26, 27).

DISCUSSION

Comparison of DNA reassociation kinetics

A comparison of DNA reassociation kinetics of the four millet species reveals both similarities and species specific differences. The slow repetitive DNA component present in all the four millet species accounts for 23.5 - 27.4% of the total genome and includes DNA sequences with a kinetic complexity in the range of 2.84×10^4 - 3.40×10^6 nucleotide pairs and reiteration frequencies in the range of 91 - 288. Fox tail millet is an exception where the proportion of the slow repetitive component is only 8%. The single copy DNA sequences constitute 45.5 - 47.8% of the total genome and have a kinetic complexity of 3.87×10^7 - 6.81×10^8 nucleotide pairs in all the millets except great millet DNA where the proportion of single copy DNA

sequences constitute 31% of the total genome.

Significant differences are, however, observed in the kinetic parameters of the rapidly reassociating and fast reassociating DNA components. The proportion of unresolvable rapidly reassociating DNA component, for example, is 5.4 - 6.4% in case of barn yard millet and great millet and 14.7 - 17.9% in fox tail millet and little millet. Barn yard millet and great millet also differ from fox tail millet and little millet in the presence of a fast reassociating DNA component. This component however, has a kinetic complexity and a reiteration frequency of 8.75×10^2 nucleotide pairs and 2×10^5 respectively in barn yard millet and 6.94×10^3 nucleotide pairs and 7.62×10^3 respectively in great millet. Thus the fast component of barn yard millet shows a higher reiteration frequency and a lower complexity than the corresponding fast component in great millet.

The differences observed in the kinetic parameters in the four millet genomes may be the factors responsible for speciation. From the present data, the rapidly reassociating and fast reassociating DNA sequences appear to play an important role in these events in the millet genomes.

Salient features of fox tail millet

Fox tail millet has the lowest DNA content (0.80 pg)

among the millets under investigation. The proportion of repetitive DNA is 26% and is the lowest reported so far in the family Gramineae. Furthermore, this genome is characteristic in having the lowest proportion (8%) of the slow repetitive DNA component among the four millet species. Finally, it has the largest proportion of single copy DNA which fails to reassociate even at a Cot value of 1×10^4 M.s. The presence of low proportion of repetitive DNA and a small nuclear genome could be a functional or any other constraint imposed by selective or random forces.

Relationship between kinetic and cytophotometric estimates of genome size

Apart from furnishing information about the proportion of repetitive and single copy components, the DNA reassociation kinetics can also be used to estimate the haploid genome size from the Cot 1/2 values of unique or non-repetitive DNA fractions. These values are 0.75 pg, 0.042 pg, 0.088 pg & 0.33 pg for barn yard millet, fox tail millet, little millet and great millet respectively (Table 3.2). These kinetic estimates of haploid genome size are lower than those obtained by cytophotometry. In plants, genome size estimates obtained by cytophotometry and reassociation kinetics are often not comparable. In tobacco (28), for example, the DNA content of 1.65 pg obtained from reassociation kinetics

TABLE 3.2 - GENOME SIZE ESTIMATE BASED ON CYTOPHOTOMETRY AND REASSOCIATION KINETIC MEASUREMENT IN FOUR MILLET SPECIES

Species	Kinetic ^a complexity in nucleotide pairs	DNA ^b content (1C pg)	DNA ^c content (1C pg)
<u>Barn yard millet</u>	6.81×10^8	0.750	2.70
<u>Fox tail millet</u>	3.82×10^7	0.042	0.80
<u>Little millet</u>	7.99×10^7	0.088	1.00
<u>Great millet</u>	3.05×10^8	0.330	4.60

^aFrom Table 3.1

^bValues obtained from kinetic complexity using the relationship of
1 pg = 9.12×10^8 nucleotide pairs.

^cObtained by cytophotometry

differs significantly from that of the cytophotometric estimate of 10 pg. Similar discrepancies have also been observed in other species such as cotton (25), broad bean (29), pea (30) and pearl millet (31). The observed discrepancies in the genome size estimates can be explained on the basis that truly unique sequences are not present in these species. For example, the genome of Vicia faba (29) contains 60 identical copies of the so-called unique DNA sequences whereas Vicia sativa contains only four identical copies. Certain inaccuracies in the Cot 1/2 determination of unique DNA may also lead to erroneous estimates of genome size. According to Kemp and Merlo (32) and Merlo and Kemp (33), factors capable of accelerating the rate of reassociation may be present in plant DNA preparations.

Relationship between the proportion of repetitive DNA
and single copy/^{DNA}with nuclear DNA content

From our data on the repetitive DNA content determinations of four additional Gramineae species, it is possible to comment on the relationships between the amount of repetitive DNA and nuclear DNA content in Gramineae. Flavell et al. (34) have shown that plant species with DNA content above 5 pg have more repetitive DNA than those with DNA content less than 5 pg. It can be seen from

TABLE 3.3 - RELATIONSHIP BETWEEN GENOME SIZE AND PROPORTIONS OF REPETITIVE AND UNIQUE DNA

Species	Genome size pg	Total repetitive DNA		Unique DNA		Reference
		%	pg	%	pg	
<u>Oryza sativa</u>	0.60	52.0	0.312	48.0	0.288	Deshpande & Ranjekar, (10)
<u>Sataria italica</u>	0.80	30.0	0.240	70.0	0.560	Present data
<u>Panicum miliare</u>	1.00	38.0	0.380	62.0	0.620	Present data
<u>Eleusine coracana</u>	1.60	49.0	0.784	51.0	0.816	Deshpande & Ranjekar, (10)
<u>Pennisetum americanum</u>	2.45	54.0	1.323	46.0	1.127	Deshpande & Ranjekar, (10)
<u>Echinochloa frumentacea</u>	2.70	42.0	1.134	62.0	1.674	Present data
<u>Sorghum vulgare</u>	4.60	52.0	2.392	48.0	2.208	Present data
<u>Hordeum vulgare</u>	5.45	78.0	4.251	22.0	1.199	Ranjekar et al., (35)
<u>Zea mays</u>	7.70	58.0	4.466	42.0	3.232	Hake and Walbot, (9)
<u>Secale cereale</u>	9.00	73.5	5.850	26.5	3.285	Ranjekar et al., (35)
<u>Triticum aestivum</u>	17.30	80.0	13.840	20.0	3.460	Ranjekar et al., (7)

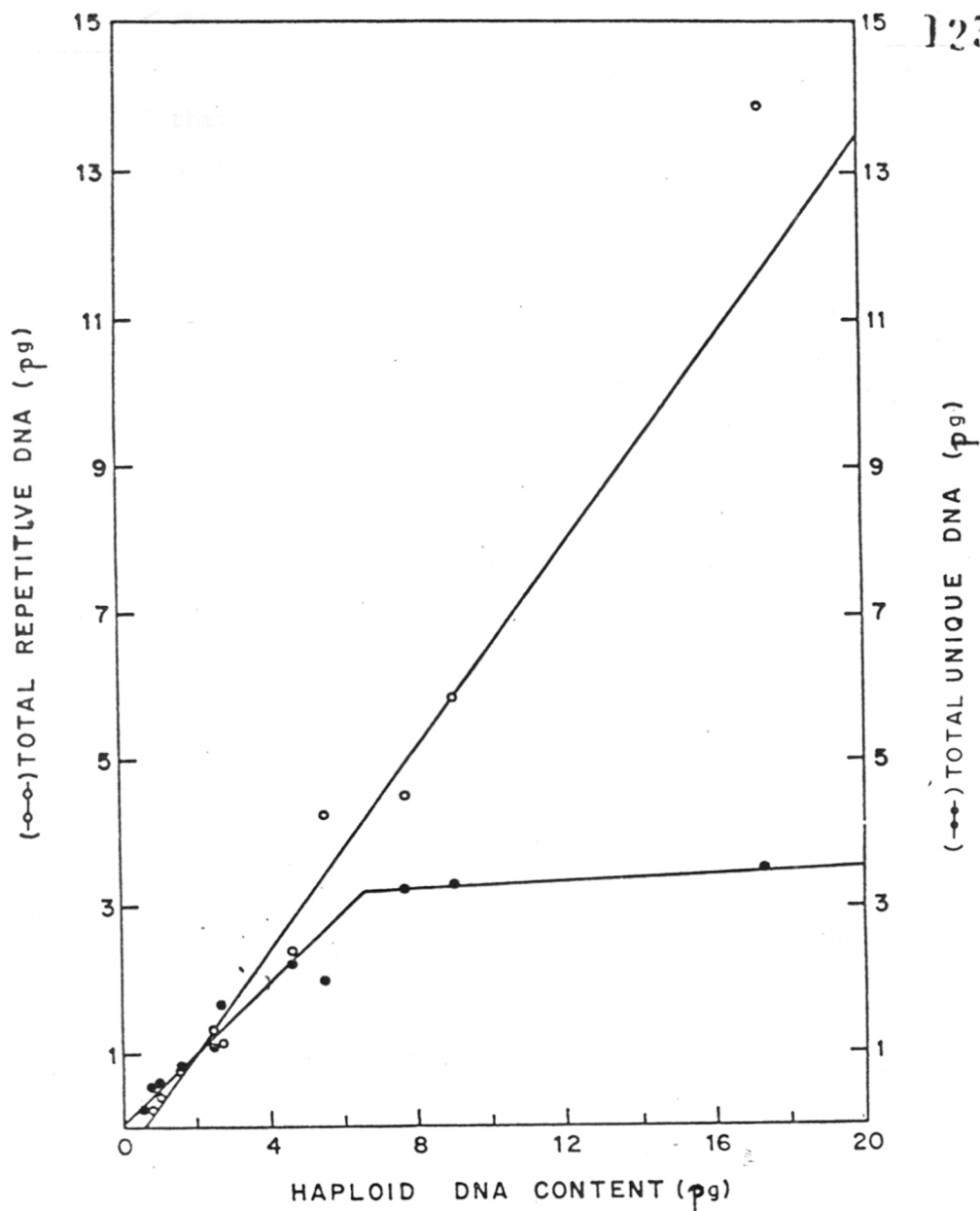


Figure 3.2 RELATIONSHIP BETWEEN NUCLEAR DNA CONTENT (pg) AND REPETITIVE DNA (pg) (O—O) AND SINGLE COPY DNA (pg) (●—●) CONTENT IN PLANT SPECIES BELONGING TO THE FAMILY GRAMINEAE. THE DATA POINTS USED ARE FROM TABLE 3.3.

Fig. 3.2 that the proportion of repetitive DNA varies linearly with DNA content. This is true for the single copy DNA also but the slopes of the two lines are different. Above a nuclear DNA content of about 6 picograms, the content of single copy DNA remains more or less constant while that of repetitive DNA continues to increase. Hence it appears that repeated DNA contributes significantly to the genome size of these plant species.

REFERENCES

- 1 Thompson, W.F. and Murray, M.G. (1981). In: *The Biochemistry of Plants. A comprehensive Treatise* Vol.6, Proteins and Nucleic acids. Marcus, A. ed. pp.1 - 81, Academic Press, N.Y.
- 2 Britten, R.J., Graham, D.E. and Neufeld, B.R. (1974). In *Methods in Enzymol.*, Grossman, G.L., Moldave K. eds. Vol.29E, pp 363 - 418, Academic Press, New York.
- 3 Britten, R.J., Graham, D.E., Eden, F.C., Painchaud, D.M. and Davidson, E.H. (1976). *J. Mol. Evol.*, 9: 1-23.
- 4 Smith, D.B. and Flavell, R.B. (1974) *Biochem. Genet.*, 12: 243-256.
- 5 Ranjekar, P.K., Lafontaine, J.G. and Pallotta, D. (1974). *Chromosoma (Berl.)*, 48: 427-440.
- 6 Smith, D.B. and Flavell, R.B. (1975). *Chromosoma (Berl.)* 50: 223-242.
- 7 Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976). *Biochim. Biophys. Acta*, 425: 30-40.
- 8 Flavell, R.B. and Smith, D.B. (1975). *Stadler Symp.*, 7: 47-70.
- 9 Hake, S. and Walbot, V. (1980). *Chromosoma (Berl.)*, 79: 251-270.
- 10 Deshpande, V.G. and Ranjekar, P.K. (1980). *Hoppe-Seyler's Z. Physiol. Chem.*, 361: 1223-1233.
- 11 Gupta, V.S. and Ranjekar, P.K. (1981). *J. Biosci.*, 3: 417-430.
- 12 Gupta, V.S. and Ranjekar, P.K. (1982). *Indian J. Biochem. Biophys.*, 19: 167-170.
- 13 Gupta, V.S., Gadre, S.R. and Ranjekar, P.K. (1981). *Biochim. Biophys. Acta*, 656: 147-154.
- 14 McLeish, J. and Sunderland, N. (1961). *Exp. Cell. Res.*, 24: 527-540.
- 15 Walker, P.M.B. and McLaren, A. (1965). *J. Mol. Biol.*, 12: 394-409.

- 16 Tiselius, A., Hjerten, S. and Levin, O. (1956). Arch. Biochem. Biophys., 65: 132-155.
- 17 Britten, R.J. and Kohne, D.E. (1968). Science, 161: 529-540.
- 18 Kohne, D.E. and Britten, R.J. (1971). In 'Procedures in Nucleic Acids Research'; Cantoni, G.C. and Davies, D.R. eds. Vol. 2, pp.500-512. Harper and Row, New York.
- 19 Ranjekar, P.K. and Murthy, M.R.V. (1973). J. Neurochem., 20: 1257-1264.
- 20 Pearson, W.R., Davidson, E.H. and Britten, R.J. (1977). Nucleic Acids Research, 4: 1727-1735.
- 21 Wetmur, J.G. and Davidson, N. (1968). J. Mol. Biol., 31: 349-370.
- 22 Cairns, J. (1963). Cold Spring Harbor Symp. Quant. Biol., 28: 43-46.
- 23 McConaughy, B.L. and McCarthy, B.J. (1970). Biochem. Genet., 4: 425-446.
- 24 Pays, E. and Ronsse, A. (1975). Biochem. Biophys. Res. Commun., 62: 862-867.
- 25 Walbot, V. and Dure, L.S. (1976). J. Mol. Biol., 101: 503-536.
- 26 Cullis, C.A. (1981). Biochim. Biophys. Acta, 652: 1-15.
- 27 Pearson, W.R., Wu, J.R. and Bonner, J. (1978). Biochemistry, 17: 51-59.
- 28 Zimmerman, J.L. and Goldberg, R.B. (1977). Chromosoma (Berl.), 59: 227-252.
- 29 Gnucheva, E. Yu., Kir'Yanov, G.I., Palyakov, V. Yu and Chentsov, Yu. S. (1977). Biokhimiya, 42: 2186-2191.
- 30 Pearson, G.G., Smith, S.L., Wu, J.R. and Bonner, J. (1978). Plant Physiol., 62: 112-115.
- 31 Wimpee, C.F. and Rawson, J.R.Y. (1979). Biochem. Biophys. Acta, 562: 192-206.
- 32 Kemp, J.D. and Merlo, D.J. (1975). Biochem. Biophys. Res. Commun., 67: 1522-1526.

- 33 Merlo, D.J., and Kemp, J.D. (1976). Plant Physiol. 58: 100-106.
- 34 Flavell, R.B., Bennett, M.D., Smith, J.B. and Smith, D.B. (1974). Biochem. Genet., 12: 257-269.
- 35 Ranjekar, P.K., Pallota, D. and Lafontaine, J.G. (1978). Biochem. Genet., 16: 957-970.

CHAPTER IV

CHARACTERIZATION OF REPEATED DNA
SEQUENCES IN MILLETS

SUMMARY

Isolated repeat fractions of the four millet genomes were characterized by studying their thermal denaturation and optical reassociation properties. The thermal denaturation profiles of all the repetitive DNAs, with the exception of great millet, are biphasic with the presence of a minor high melting component (T_m : 96-100°C) accounting for about 3-8% of the total genome. This component occurs mainly in the rapidly reassociating (Cot 0.1 M.s.) DNA fraction. Total repeated DNA sequences show a moderate sequence divergence of 6-8% in all the four millets. The intermediately repetitive DNA shows a greater sequence divergence than highly repetitive DNA indicating that the former is older in origin.

Mini Cot curves reveal no additional sub-class of repetitive DNA other than those observed in the reassociation of the total 0.4 kbp^{long}/millet DNAs. Great millet DNA is the only exception where an additional fast reassociating DNA component is observed in the highly repetitive DNA fraction.

Reassociation of 0.4 kbp long DNAs of barn yard millet, little millet and great millet were studied at different incubation temperatures, ranging from $T_m - 30^\circ\text{C}$ to $T_m - 5^\circ\text{C}$; allowing the fractionation of repeats according

to their relatedness. In all the three species, the proportion of all the repetitive fractions and their reassociation rates decrease with an increase in the incubation temperature. Since only short and perfect duplexes, with low copy number, reassociate at increased stringency of reassociation, the repeat families in these millets appear to be heterogeneous in nature. Furthermore, comparison of the proportions of 'fossil repeats' and 'true single copy' sequences suggests that the genome of great millet is 'turning over' at a much faster rate than that of barn yard millet and little millet.

INTRODUCTION

Analysis of DNA reassociation kinetics yields a valuable information with respect to the structure and organization of eukaryotic genomes. Most of the reassociation studies are performed at a temperature about 25°C below the melting temperature of native DNA, a criterion at which optimal reassociation rates are generally achieved. However, the repetitive DNA populations often show extreme kinetic heterogeneity (1,2), and different degrees of sequence divergence resulting in a frequency/divergence continuum rather than belonging to discrete classes. One way to study this heterogeneity and sequence divergence is to analyse

isolated repeat fractions by reassociation kinetics and melting, which often reveal features that are normally obscured during the study of total DNAs. It is also possible to fractionate the repetitive sequences according to their relatedness, by studying their reassociation kinetics at different temperatures. When the reassociation temperature is reduced below that of a standard criterion such as $T_m - 25^{\circ}\text{C}$, to a more permissive criterion, more mismatched duplexes are allowed to form to give an apparently higher proportion of repeated sequences. In contrast, at higher reassociation temperatures, only the very well matched sequences (probably conserved sequences) reassociate. Thus, the terms repetitive and single copy sequences are largely operational definitions (3) and can be defined as sequences that are present in many and one copy, respectively, at standard criterion. The term 'fossil repeats' is used for those sequences which reassociate with single copy kinetics at standard criterion but with repetitive kinetics at lower temperatures. 'True single copy' sequences are those which follow single copy kinetics at both standard and permissive criteria (3). By fractionating the genome into 'fossil' repeats and 'true single copy', it is possible to predict about the rate of 'turn over' including events like amplification and deletion during evolution (4).

The nature of families of repetitive DNA sequences can also be determined by studying the reassociation kinetics at different criteria of temperature stringency. A family is defined as a group of sequences which have sufficient similarity to reassociate with one another at a given criterion of temperature stringency and salt conditions. Studies at different temperatures reveal variations in (i) the number of repetitive families in a repetitive pool and their relatedness and (ii) the number of different members in a repeat family and their relatedness. Two types of families generally recognized are (i) a homogeneous family which contains member sequences which are related to the same extent and (ii) a heterogeneous family which contains member sequences of varying similarity ranging from nearly perfect replicas to sequences barely similar enough to reassociate at standard criterion of reassociation (5,6).

This Chapter describes the following: (i) Melting behaviour of different repeat fractions isolated by hydroxyapatite column chromatography to estimate the sequence divergence among repeated DNA sequences. (ii) Mini Cot curves of total repetitive DNA fractions of the four millets and those of highly repetitive and intermediately repetitive fractions of barn yard millet and great millet to identify the presence or absence of repeat families with discrete

repetition frequency. (iii) Reassociation of sonicated, 0.4 kbp long DNAs under different incubation temperatures to determine the nature of repeated DNA sequences (~~homogeneous~~ or heterogeneous).

MATERIALS AND METHODS

Chemicals

Restriction enzymes Msp I, Hpa II and Hae III were obtained from Bethesda Research Laboratory, USA. Methanol (AR Grade) was procured from Loba Chemie, Bombay, India and pTT (dithiothreitol) from Sigma Chemical Co., MO, USA.

Preparation of DNA

Extraction of DNA from the seedling of barn yard millet, fox tail millet, little millet and great millet, shearing of DNAs to an average size of 0.4 kbp, and sizing of DNAs were carried out as detailed in Chapters II and III.

Isolation of repetitive DNA fractions

As shown earlier, the reassociation curves of 0.4 kbp long barn yard millet DNA and great millet DNA could be modelled into two repetitive and one single copy component and that of fox tail millet and little millet DNA into one repetitive and one single copy component (Chapter III). Hence kinetic fractions enriched in highly repetitive and

intermediately repetitive sequences were prepared in case of barn yard millet and great millet. Total repetitive DNA sequences were isolated in case of all the four millets. In case of barn yard millet, the highly repetitive DNA included sequences which were bound to hydroxyapatite at Cot 0.5 M.s. while DNA sequences which were bound to hydroxyapatite at Cot 75 M.s. but not at Cot 0.5 M.s. were considered to be intermediately repetitive. In case of great millet, sequences reassociating by Cot 0.5 M.s. and those reassociating between Cot 0.5 and Cot 20 M.s. were considered highly repetitive and intermediately repetitive respectively. Total repetitive DNA in case of barn yard millet, fox tail millet, little millet and great millet included sequences reassociating by Cot 75 M.s., Cot 5 M.s., Cot 15 M.s. and Cot 20 M.s., respectively. Isolations of all the repeat fractions were done by hydroxyapatite column chromatography (Chapter III).

Thermal denaturation of native 0.4 kbp long DNAs and of repetitive DNA fractions

Thermal denaturation of various DNA samples was carried out in a Gilford 250 spectrophotometer as detailed in Chapter II. Denaturation was studied in 0.12 M sodium phosphate buffer (pH 6.8) at a heat rate of 1°C per min. Total hyperchromicity was normalized to 100% and the graph

of percentage hyperchromicity versus temperature was plotted. Calculations of T_m , hyperchromicity and base mismatch were done according to the standard procedures (7-9). In barn yard millet, fox tail millet and little millet, the repeat fractions were melted in 0.012 M sodium phosphate buffer (pH 6.8) and the melting temperatures were corrected for salt concentration effect (10).

Mini Cot curves

Reassociation of repetitive DNA fractions, isolated at 62°C, were monitored optically in the Gilford 250 spectrophotometer. They were dialyzed extensively to 0.12 M sodium phosphate buffer (pH 6.8) prior to optical reassociation studies. DNA concentrations were such that the absorbance after melting still remained within the linear absorbance range of the spectrophotometer. Optical reassociation was carried out as detailed in Chapter III.

Reassociation kinetics studies at different conditions of temperature stringency

To fractionate populations of repeated sequences by their extent of relatedness, reassociation of 0.4 kbp long barn yard millet, little millet and great millet were followed optically at different incubation temperatures in the range of 55°C - 75°C. The experiments were monitored as described earlier except that the incubation temperatures were 55°C,

68°C and 75°C in addition to 62°C. Zero time varied according to the incubation temperature and was the midpoint between 100°C and the incubation temperature. The DNA samples were melted after reassociation to determine the thermal stability of duplexes formed during reassociation.

All optical reassociation data were evaluated using a computer aided non-linear regression procedure as described earlier in Chapter III.

Digestion of total millet DNAs with restriction endo-nucleases

Digestion of total millet DNAs with restriction enzymes was carried out according to the conditions recommended by the manufacturer. The digestions were generally performed in a total volume of 10 - 15 μ l at 37°C for 1-2 h with a two fold excess of enzyme. While assay buffers for the enzymes Msp I and Hpa II comprised 6 mM KCl, 10 mM Tris-HCl (pH 7.4), 10 mM MgCl₂ and 1 mM DTT, that of Hae III comprised 10 mM NaCl, 6 mM Tris-HCl (pH 7.4), 6 mM MgCl₂ and 6 mM 2-mercaptoethanol. The digested DNAs were subjected to a single deproteinization step using CHCl₃/IAA (24:1 v/v).

High performance liquid chromatography (HPLC)

Base composition analysis of millet DNAs was carried out using HPLC to determine the presence of 5-methyl cytosine. DNA samples (1.5 - 2 mg) were hydrolysed with 70% perchloric

acid in a boiling water bath for 1-2 h, chilled and excess perchloric acid was neutralized with 10 M KOH. The aqueous layer containing free bases was analyzed. The Bondapak reverse phase column C₁₈ μ was used on a Waters HPLC system (Model ALC-GPC-244). The solvent used was water: methanol: acetic acid (94:4:0.25 v/v/v). The eluate was monitored at 254 nm. Authentic mixtures of free bases and acid hydrolysed calf thymus DNA were run as controls. The flow rate of the column was kept at 0.9 ml/min and the chart speed was 1 cm/min. The sensitivity of the instrument was kept constant to allow a better comparison.

RESULTS

Melting analysis of 0.4 kbp long millet DNAs and isolated repeat fractions

As seen in Fig.4.1, all the native 0.4 kbp long millet DNAs show smooth melting profiles with average melting temperatures varying in the narrow temperature range of 85.9°C to 86.9°C and hyperchromicities in the range of 23 - 25%. Repeat DNA fractions (Cot 0.1 and limit repetitive Cot), on the other hand exhibit biphasic melting profiles except in the case of great millet DNA. The melting temperature of the high melting fraction varies from 96-100°C while that of the low melting fraction is in the range of

78-84°C. In great millet, the total repetitive DNA has a melting temperature of 79.6°C. The average hyperchromicity of all the repetitive DNA fractions varies in the range of 12-17%.

The proportion of the high melting component is 10-17% in the limit Cot fractions (Table 4.1). Its proportion in the Cot 0.1 DNA is, however, significantly high and is in the range of 37-45%. The DNA sequences reassociating in the Cot range of 0.1 to 10, however, do not reveal the presence of this high melting DNA and have a melting temperature in the range of 77.4 - 79.4°C (Figure 4.1). The high melting DNA sequences thus appear to be confined to the rapidly reassociating fraction and represent 3-8% of the total DNA in barn yard millet, fox tail millet and little millet.

The highly repetitive and intermediately repetitive DNA fractions which are present only in barn yard millet and great millet were isolated and characterized. In barn yard millet, the highly repetitive DNA has a biphasic melting profile. The T_m values of the low melting and high melting DNA fractions are 78.7°C and 98°C respectively (Table 4.1 and Figure 4.2a). The intermediately repetitive fraction (Cot 1.0 - Cot 75 M.s.) has a T_m of 79.4°C. In great millet, the highly repetitive DNA (Cot < 0.5 M.s.) shows a melting

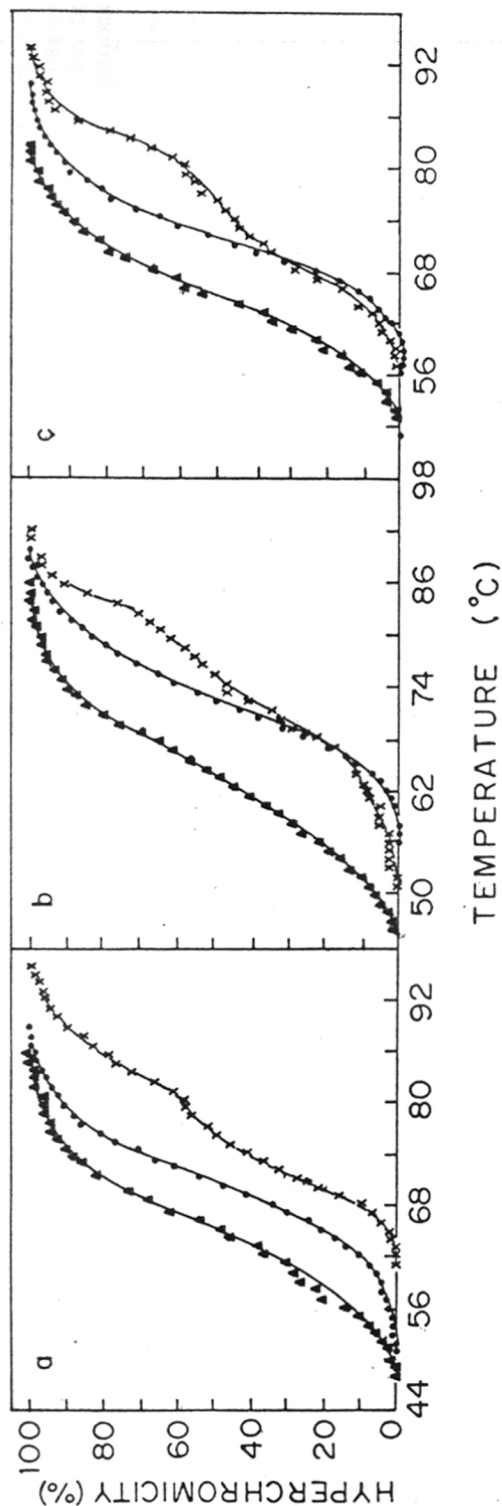


Figure 4.1 THERMAL DENATURATION PROFILES OF AVERAGE 0.4 kbp LONG DNA (●—●), COT 0.1 DNA (X—X) AND COT 0.1 - 1.0 M.s. DNA (▲—▲).

- a Barn yard millet
- b Fox tail millet
- c Little millet.

Melting of Cot fractions was done in 0.012 M sodium phosphate buffer, pH 6.8.
Melting temperatures were corrected for salt concentration effect (10).

TABLE 4.1 - MELTING ANALYSIS ON 0.4 kbp LONG DNAs AND REPETITIVE DNA FRACTIONS

Species and Cot fraction	No. of expts.	Proportion ^a		T _m ^b	Hyperchro- ^c		Extent of reassociation	Proportion ^e of the high melting DNA component in the total genome (%)	
		(%)	(%)	(°C)	(%)	(%)	(%)	(%)	(%)
I	2	3	4	5	6	7			
<u>Barn yard millet</u>									
Native 0.4 kbp long DNA	8	-		85.95±0.37	24.90±0.37	-	-	-	-
Cot 0.1	4	54.6±4.3 45.4±4.3		84.70±1.60 99.10±2.50	16.50±0.30	15.0	6.8		
Cot 0.5 (highly repetitive)		63.2±1.8 36.8±1.8		78.70±0.30 97.95±0.35	16.00±1.00	17.0	6.3		
Cot 0.1 - 10	3	-		79.40±0.50	12.08±0.08	3.5	-		
Cot 0.1 - 75 (Intermediately repetitive)	3	-		79.40±1.00		27.0	-		
Cot 75.0 (total repetitive)	5	88.4±3.0 11.6±3.0		78.90±0.80 96.10±1.90	14.40±2.00	42.0	4.9		
<u>Fox tail millet</u>									
Native 0.4 kbp long DNA	8	-		86.80±0.35	25.01±1.90	-	-		
Cot 0.1	3	63.0±3.0 37.0±3.0		86.60±0.20 100.30±0.00	16.90±0.30	18.0	6.7		
Cot 0.1 - 10	3	-		77.40±0.60	13.21±0.21	6.0	-		
Cot 50 (total repetitive)	4	83.5±3.5 16.5±3.5		80.40±1.00 96.00±2.40	14.30±1.40	30.0	5.0		

TABLE 4.1 (Contd.)

1	2	3	4	5	6	7
<u>Little millet</u>						
Native 0.4 kbp long DNA	10	-	86.40±0.20	23.00±1.20	-	-
Cot 0.1	4	54.8±7.4 45.2±7.4	83.40±1.40 97.80±0.50	13.80±0.30 12.45±0.39	16.0	7.2
Cot 0.1 - 10	3	-	78.90±0.30		6.5	-
Cot 15.0 (total repetitive)	10	90.3±3.2 9.7±3.2	79.30±1.00 96.40±1.70	15.50±2.70	8.0	3.7
<u>Great millet</u>						
Native 0.4 kbp long DNA	8	-	86.9±0.90	23.80±1.00	-	-
Cot 0.5 (highly repetitive)	5	-	80.06±0.84	17.30±0.80	25.0	-
Cot 0.5 - 20 (Intermediately repetitive)	3	-	77.25±0.45	20.5±1.20	27.00	-
Cot 20 (total repetitive)	5	-	79.55±0.25	18.25±0.70	52.00	-

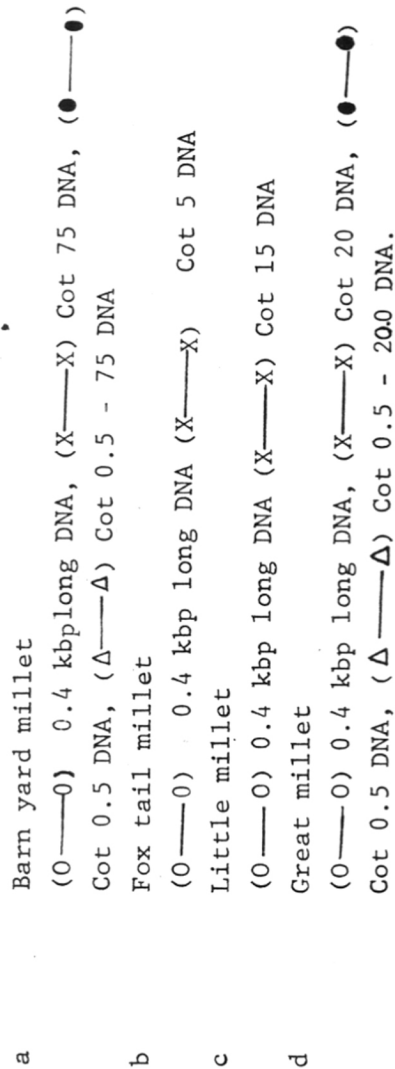
a, b Obtained from Figures 4.1 & 4.2

c Calculated as described in caption to Table 2.1

d Obtained from Fig. 3.1

e Obtained as $e = d \times a/100$.

Figure 4.2 THERMAL DENATURATION PROFILES OF ENRICHED DNA FRACTIONS



Melting of enriched fractions was done in 0.12 M sodium phosphate buffer, pH 6.8.

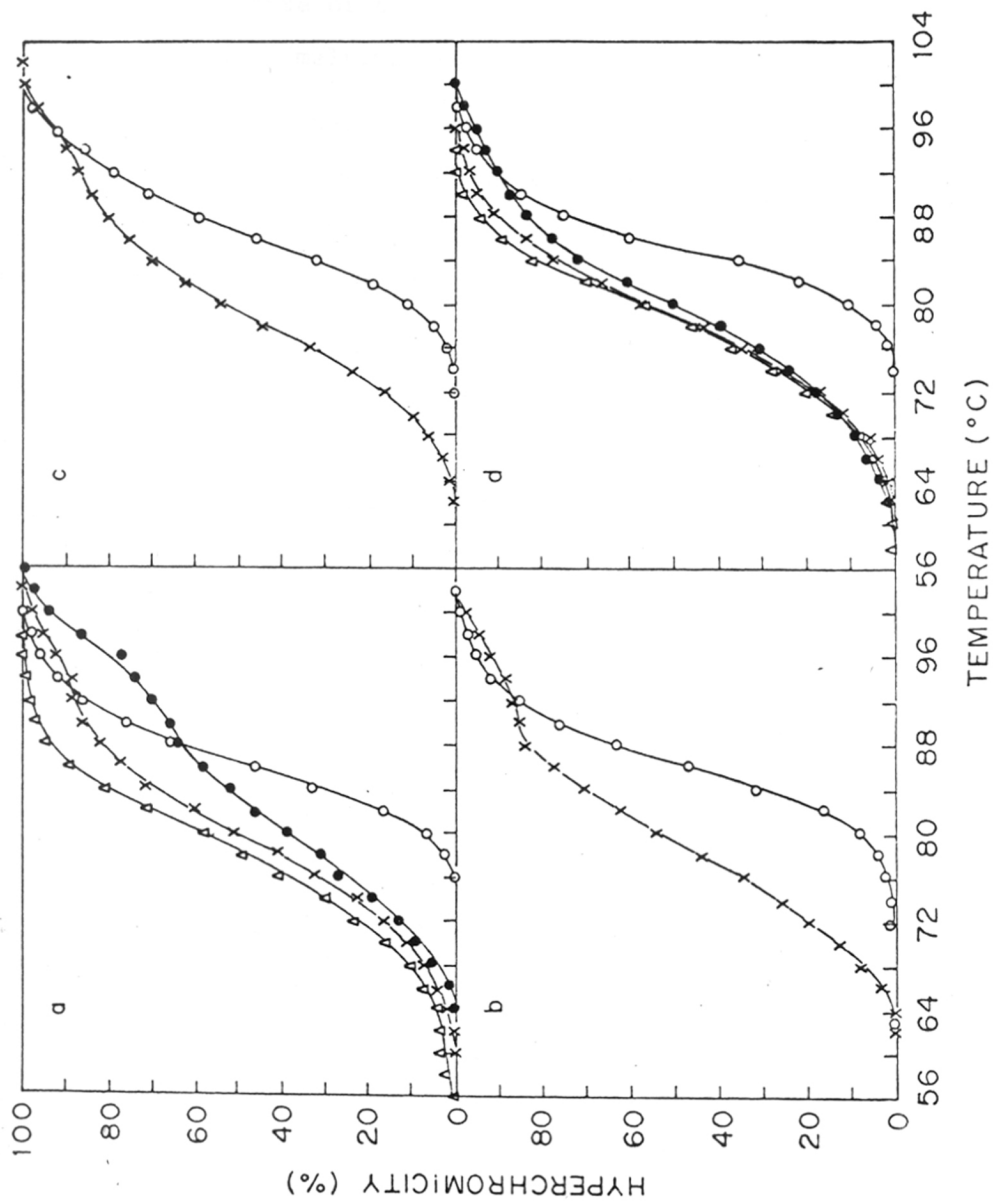


FIGURE 4.2

temperature of 80.6°C and the intermediately repetitive DNA has a melting temperature (Figure 4.2d) of 77.25°C.

In barn yard millet, fox tail millet and little millet, the T_m values of the low melting fraction of the total repetitive DNAs were compared to those of corresponding native, sheared DNAs. From the lowering in T_m of these repeat fractions, the % base mismatch is estimated to be 6.4-7.4% in all the four millet species. The intermediately repetitive component isolated in case of barn yard millet and great millet show a lowering in T_m of 6.6°C and 10.0°C respectively compared to that of their native DNAs. Hence the sequence divergence of the intermediately repetitive DNA fraction in these millets varies in the range of 6.55 to 9.65%. The highly repetitive DNA in these two plant species shows a base mismatch of 6.84 to 7.25% (Table 4.2).

Analysis of mini Cot curves

The total repetitive DNA fractions were isolated in all the four species and their optical reassociation was studied to assess their kinetic heterogeneity. Since the two repetitive components - one highly repetitive and one intermediately repetitive - were identified in barn yard millet and great millet, these were also isolated and similarly studied. The results are summarized in Table 4.3.

TABLE 4.2 - SEQUENCE DIVERGENCE AMONG REPEATED DNA SEQUENCES

Species and repetitive fraction	T _m ^a (°C)	Base mismatch ^b (%)
<u>Barn yard millet</u>		
0.4 kbp long DNA	85.95 ± 0.37	-
Highly repetitive DNA	LMF 78.70 ± 0.30	7.25
	HMF 97.95 ± 0.35	
Intermediately repetitive DNA	79.40 ± 1.00	6.55
Total repetitive DNA	LMF 78.90 ± 0.80	7.05
	HMF 96.10 ± 1.90	
<u>Fox tail millet</u>		
0.4 kbp long DNA	86.80 ± 0.35	-
Total repetitive DNA	LMF 80.40 ± 1.00	6.40
	HMF 96.00 ± 2.40	
<u>Little millet</u>		
0.4 kbp long DNA	86.40 ± 0.20	-
Total repetitive DNA	LMF 79.30 ± 1.00	7.10
	HMF 96.40 ± 1.70	
<u>Great millet</u>		
0.4 kbp long DNA	86.90 ± 0.90	-
Highly repetitive DNA	80.06 ± 0.84	6.84
Intermediately repetitive DNA	77.25 ± 0.45	9.65
Total repetitive DNA	79.55 ± 0.25	7.35

^aObtained from Fig. 4.2

^bCalculated by using the relationship that 1°C lowering in T_m corresponds to 1% base mismatch. LMF and HMF stand for low melting fraction and high melting fraction respectively.

In a barn yard millet, the total repetitive DNA shows two kinetic components. The fast component has a rate constant of 1.09×10^3 and a repetition frequency of 7.22×10^5 and the slow component has a rate constant of 6.25×10^{-1} and a repetition frequency of 4.91×10^2 (Figure 4.3c). The highly repetitive DNA is modelled into a single kinetic component with a rate constant of 1.03×10^2 and reiteration frequency of 9.45×10^5 (Figure 4.3a). The intermediately repetitive DNA is also resolved into a single second order component with a rate constant of 1.65×10^{-1} and repetition frequency of 1.06×10^3 (Figure 4.3b).

In fox tail millet and little millet, the total repetitive DNA is resolved into a single second order component and this component has a rate constant and reiteration frequency of 7.69×10^0 and 3.11×10^2 respectively in the former and 4.50×10^0 and 4.47×10^2 respectively in the latter (Figure 4.4a,b).

In great millet, the total repetitive DNA is resolved into two second order components with rates of 2.85×10^2 and 9.71×10^{-1} and reiteration frequencies of 2.62×10^3 and 2.56×10^2 respectively (Figure 4.5c). The highly repetitive DNA is resolved into two second order kinetic components (Figure 4.5a). The faster of the two components,

TABLE 4.3 - KINETIC ANALYSIS OF MINI COT CURVES

DNA	Fraction ^a	Cot 1/2 ^a values observed (M.s.)	Cot 1/2 ^b values pure (M.s.)	K pure ^c	Frequency ^d of repeti- tion per 1C genome	Kinetic ^e complexity in base pairs
1	2	3	4	5	6	7
<u>Barn yard millet</u>						
<u>Cot 0.5 M.s.</u>						
Very fast Cot < 9.4 x 10 ⁻⁴	0.1855	-	-	-	-	-
Fast Cot 9.4 x 10 ⁻⁴ to 5.0 x 10 ⁰	0.5345	1.81x10 ⁻²	9.67x10 ⁻³	1.03x10 ²	9.45x10 ⁵	8.21x10 ⁹
<u>Unreassociated</u>						
Cot > 5.0 x 10 ⁰	0.2800	-	-	-	-	-
<u>Cot 5 to 75 M.s.</u>						
<u>Very fast</u>						
Cot < 4.2 x 10 ⁻³	0.064	-	-	-	-	-
<u>Fast + Slow</u>						
Cot 4.2 x 10 ⁻³ to 5.0 x 10 ⁰	0.3735	1.62x10 ⁰	6.05x10 ⁻¹	1.65x10 ⁻¹	1.06x10 ³	5.14x10 ⁵
<u>Unreassociated</u>						
Cot > 5.0 x 10 ⁰	0.5611	-	-	-	-	-
<u>Cot 75 M.s.</u>						
<u>Very fast</u>						
Cot < 4.4 x 10 ⁻⁴	0.0458	-	-	-	-	-
<u>Fast</u>						
Cot 4.4 x 10 ⁻⁴ to 5.0 x 10 ⁻¹	0.3858	2.37x10 ⁻³	9.14x10 ⁻⁴	1.09x10 ³	7.22x10 ⁵	7.76x10 ⁰

.....Contd.

TABLE 4.3 (Contd.)

1	2	3	4	5	6	7
Slow Cot 5.0×10^{-1} to 5.0×10^0	0.4603	3.48×10^0	1.60×10^0	6.25×10^{-1}	4.91×10^2	1.36×10^6
Unreassociated Cot $> 5.0 \times 10^0$	0.1682	-	-	-	-	-
<u>Fox tail millet</u>						
<u>Cot 5 M.s.</u>						
Very fast Cot $< 1.21 \times 10^{-2}$	0.2980	-	-	-	-	-
Fast and slow Cot 1.21×10^{-2} to 5.00×10^0	0.4091	3.17×10^{-1}	1.30×10^{-1}	7.69×10^0	3.11×10^2	1.10×10^5
Unreassociated Cot $> 5.00 \times 10^0$	0.3018	-	-	-	-	-
<u>Little millet</u>						
<u>Cot 15 M.s.</u>						
Very fast Cot $< 1.21 \times 10^{-2}$	0.1037	-	-	-	-	-
Fast and slow Cot 1.21×10^{-2} to 1.50×10^1	0.5000	4.41×10^{-1}	2.21×10^{-1}	4.50×10^0	4.47×10^2	1.88×10^5
Unreassociated Cot $> 1.50 \times 10^1$	0.3995	-	-	-	-	-
<u>Great millet</u>						
<u>Cot 0.5 M.s.</u>						
Very fast Cot $< 2.3 \times 10^{-3}$	0.0000	-	-	-	-	-
Fast Cot 2.3×10^{-3} to 4.0×10^{-2}	0.3700	4.84×10^{-3}	1.77×10^{-3}	5.65×10^2	7.42×10^4	1.48×10^3

.....Contd.

TABLE 4.3 (Contd.)

1	2	3	4	5	6	7
Slow Cot 4.0×10^{-2} to 3.0×10^0	0.6400	3.61×10^{-1}	2.31×10^{-1}	4.33×10^0	9.95×10^2	1.93×10^5
Unreassociated Cot $> 3.0 \times 10^0$	0.0000	-	-	-	-	-
Cot 0.5 - 20 M.s.						
Very fast Cot $< 2.0 \times 10^{-3}$	0.0910	-	-	-	-	-
Fast + Slow Cot 2.0×10^{-3} to 2.0×10	0.6600	4.08×10^{-1}	2.69×10^{-1}	3.72×10^0	8.80×10^2	2.24×10^5
Unreassociated Cot $> 2.0 \times 10^1$	0.2490	-	-	-	-	-
Cot 20 M.s.						
Very fast Cot $< 1.6 \times 10^{-3}$	0.0385	-	-	-	-	-
Fast Cot 1.6×10^{-3} to 1.13×10^{-1}	0.2600	1.37×10^{-2}	3.51×10^{-3}	2.85×10^2	2.62×10^3	2.93×10^3
Slow Cot 1.13×10^{-1} to 2.00×10^1	0.7300	1.40×10^0	1.03×10^0	9.71×10^{-1}	2.56×10^2	8.58×10^5
Unreassociated Cot $> 2.00 \times 10^1$	0.0000	-	-	-	-	-

^aValues obtained from Figure 4.3, 4.4 and 4.5 respectively.^bCot 1/2 values observed x fraction of the genome. ^cK pure = (Cot 1/2 pure)⁻¹.^dCot 1/2 of unique DNA (obtained from Fig.3.1 of Chapter III) divided by Cot 1 1/2 values of the given number fraction.^eValues obtained from the standard relationship for *E. coli* where the Cot 1/2 is 5.3 M.s. (under our experimental conditions) and genome size of 4.5×10^6 nucleotide pairs.

Figure 4.3

MINI COT CURVES OF ISOLATED REPEAT FRACTIONS
IN BARN YARD MILLET DNA

- a Highly repetitive DNA reassociating by Cot 0.5
- b Intermediately repetitive DNA reassociating
between Cot 0.5 - Cot 75 M.s.
- c Total repetitive DNA reassociating by Cot 75 M.s.

Enriched fractions, isolated by hydroxyapatite column chromatography, were melted and their optical reassociation was followed in 0.12 M sodium phosphate buffer, pH 6.8.

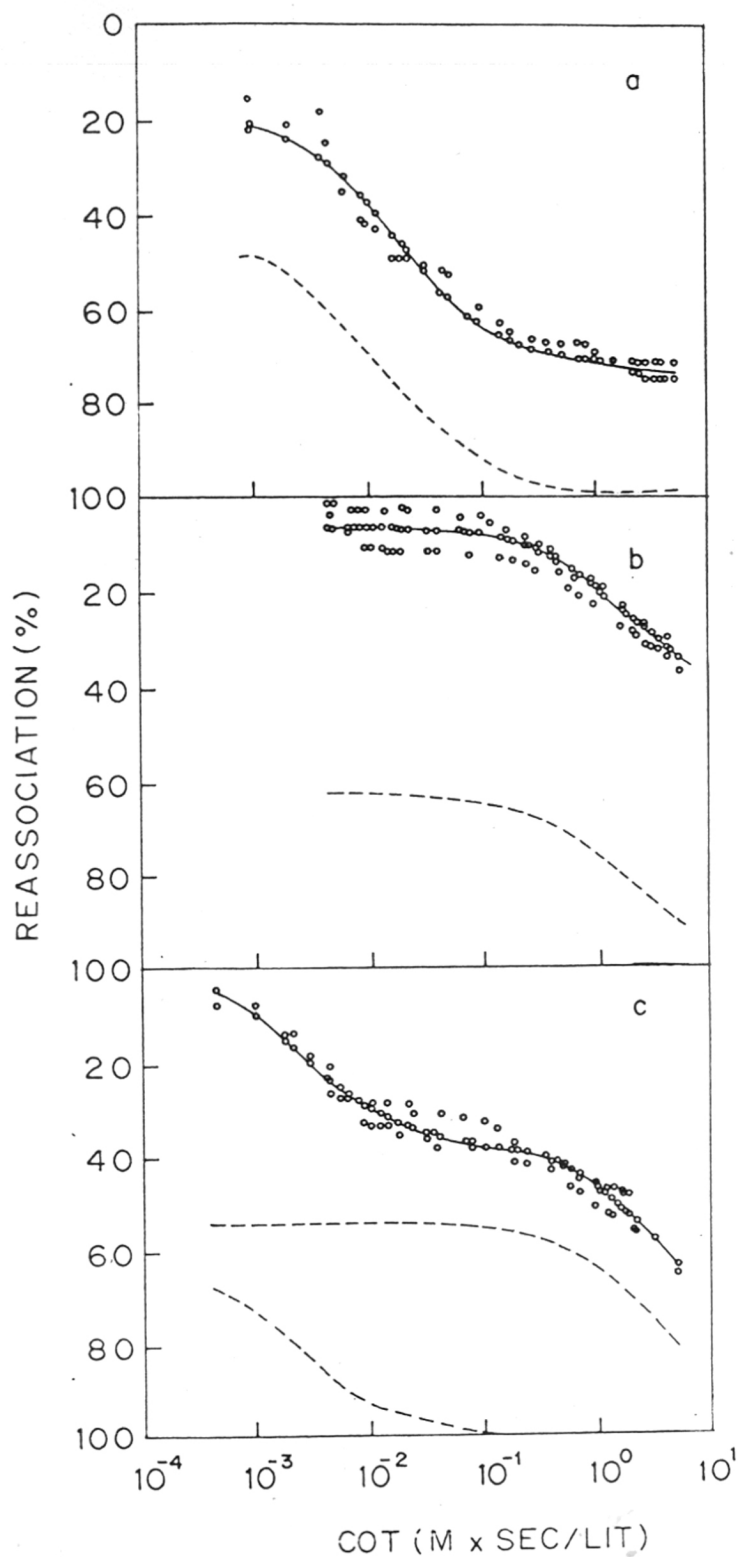


FIGURE 4.3

Figure 4.4

MINI COT CURVES OF ISOLATED TOTAL REPETITIVE
DNA FRACTIONS

- a Fox tail millet reassociating by Cot 5 M.s
 - b Little millet reassociating by Cot 15 M.s
- All the relevant details are as described in
the caption to Figure 4.3.

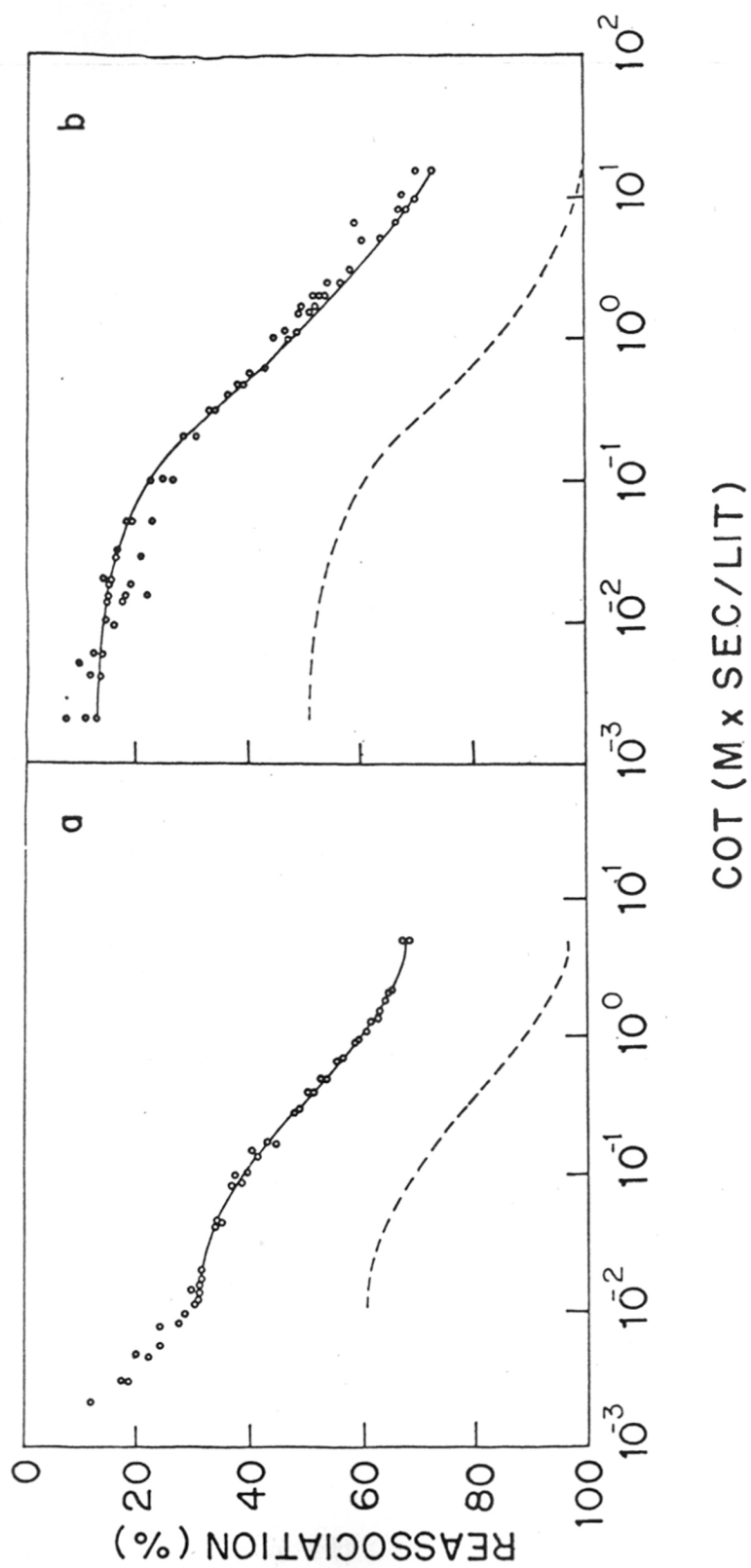
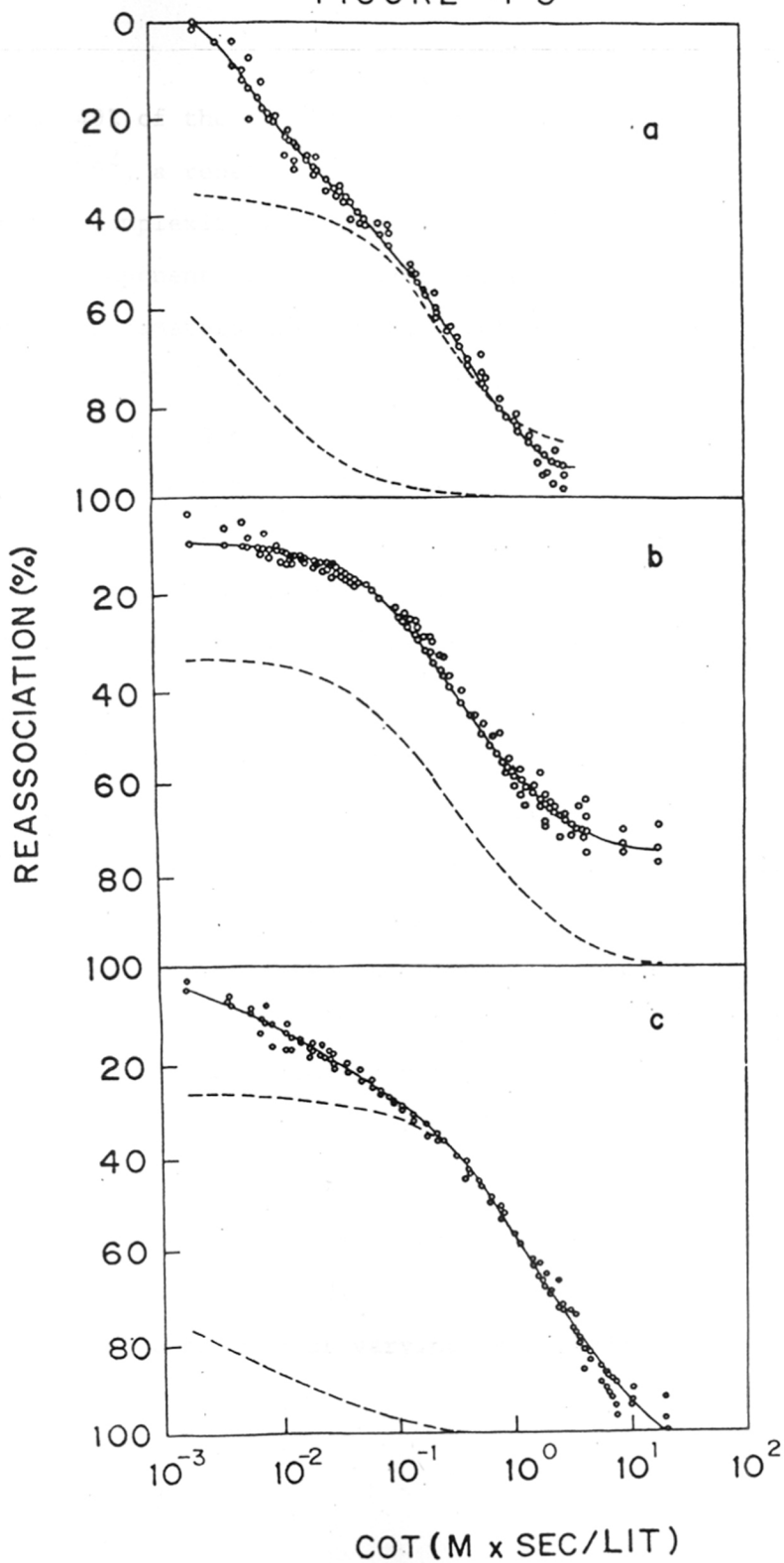


FIGURE 4.4

Figure 4.5 MINI COT CURVES OF ISOLATED REPETITIVE
DNA FRACTIONS IN GREAT MILLET.

- a Highly repetitive DNA reassociating by
Cot 0.5 M.s.
- b Intermediately repetitive DNA reassocia-
ting between Cot 0.5 - Cot 20 M.s.
- c Total repetitive DNA reassociating by
Cot 20 M.s.



almost 40% of the total fraction, has a rate constant of 5.65×10^2 , a repetition frequency of 7.4×10^4 and a kinetic complexity of 1.5×10^3 nucleotide pairs. The second component of this highly repetitive fraction exhibits kinetic parameters that are typical of intermediately repetitive DNA. It has a rate constant of 4.33×10^0 and reiteration frequency of 9.95×10^2 . The intermediately repetitive class is modelled as a single kinetic component with a rate constant of 3.73×10^0 and reiteration frequency of 8.80×10^2 indicating that it has kinetics similar to the slow component of the highly repetitive DNA (Figure 4.5b).

From the above data, it is clear that no subclass of repetitive DNA, other than those observed in the reassociation of the total 0.4 kbp long millet DNAs are observed (Chapter III). The only exception is great millet DNA where an additional fast reassociating component is seen in the highly repetitive fraction. It is clear from the analysis that the two classes of repetitive DNA namely highly repetitive and intermediately repetitive DNA are not completely distinct and a continuum of Cot 1/2 values can be found between the two main values.

The kinetics of a repetitive component are affected to some extent when studied in the presence of other repetitive components of varying reiteration frequency or

in the presence of single copy DNA. The isolated components, therefore, show an increase in the rate of reassociation and copy number as against the values obtained for the same component in a total kinetic curve. This^{is}/well illustrated in the study of mini Cot curves of all the millets except barn yard millet where a decrease in rate is observed even though the copy number shows an increase (Table 4.3).

Analysis of reassociation kinetics at different incubation temperatures

As detailed in the introduction of this Chapter, reassociation of DNA at different stringency criteria is useful for fractionating populations of repeated DNA sequences and their extent of relatedness. In the present work, barn yard millet, little millet and great millet DNAs have been analyzed for their reassociation behaviour at 55°C, 68°C and 75°C in addition to 62°C (standard criterion). The reassociations were studied optically and all the curves were evaluated using a computer aided non-linear regression method. Though in a few instances, the reassociation curves could be modelled into two components, all the data were fitted as a single component to facilitate a better comparison of their rate constants. The Cot 1/2 of single copy DNA in all the reassociations

TABLE 4.4 - MELTING ANALYSIS OF DUPLICES FORMED AT DIFFERENT CRITERIA OF TEMPERATURE STRINGENCY

Species	T _m ^a (°C)	T _p ^b	T _i - T _p ^c	Rate relative ^d to optimum rate
<u>Barn yard millet</u>				
55°C	73.9	79.95	-24.95	1.00
62°C	79.7	82.85	-20.85	1.00
68°C	81.2	83.60	-15.60	0.86
75°C	83.6	84.80	- 9.80	0.60
<u>Little millet</u>				
55°C	75.7	81.05	-26.05	1.00
62°C	80.4	83.40	-21.40	0.98
68°C	80.7	83.55	-15.55	0.88
75°C	84.1	85.25	-10.25	0.61
<u>Great millet</u>				
55°C	76.5	81.7	-26.70	0.97
62°C	79.4	83.15	-21.15	0.98
68°C	82.0	84.45	-16.45	0.87
75°C	83.5	85.20	-10.20	0.61

^aT_m of duplexes formed after reassociation at different criteria of temperature stringency (Fig. 4.7).

^bT_p = (T_n + T_m)/2. T_n is the melting temperature of 0.4 kbp long millet DNAs. This value is 85.95°C, 86.40°C and 86.90°C for barn yard millet, little millet and

great millet DNAs respectively (Table 4.1 & 4.2).

c_{Ti} = Incubation temperature

d_{Taken} from Fig.8 of Bonner et al. (11).

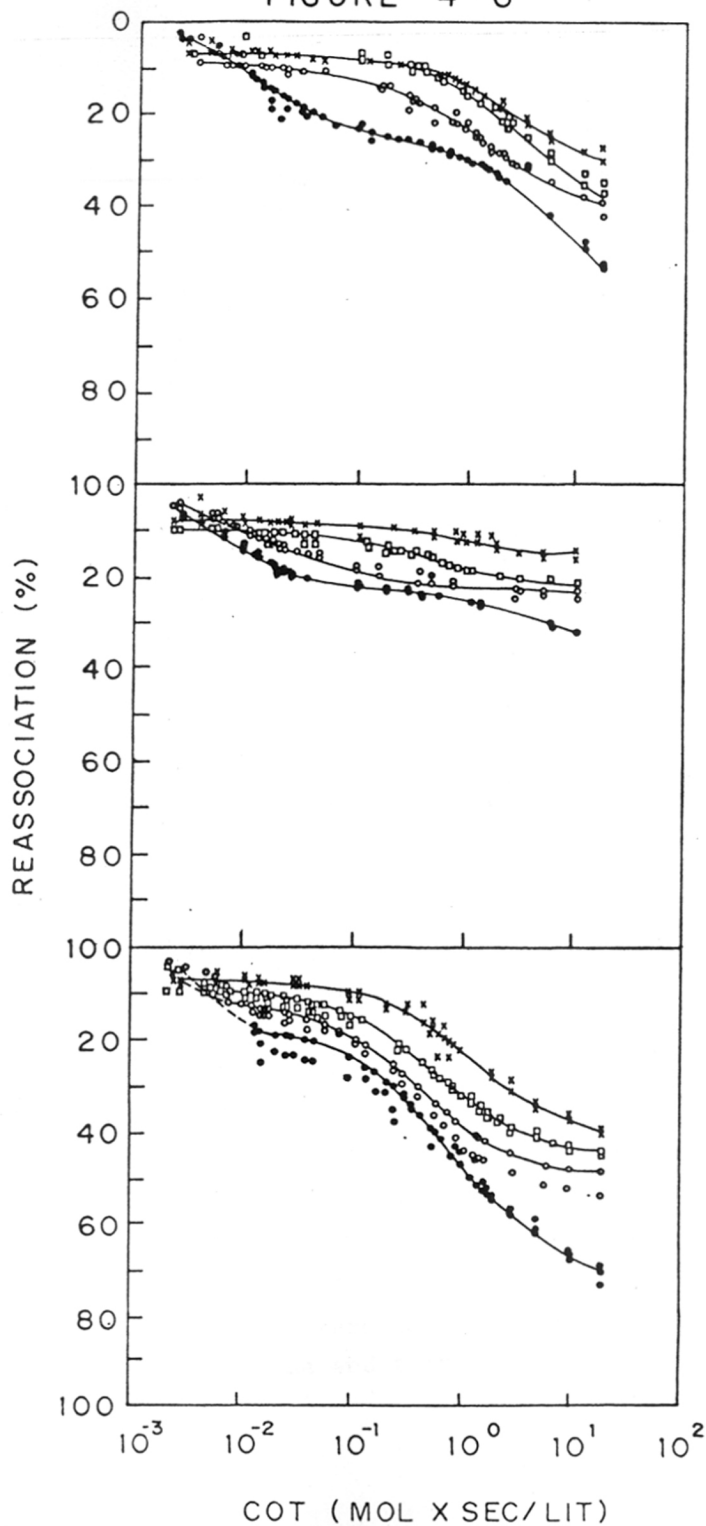
Figure 4.6 OPTICAL REASSOCIATION OF TOTAL SONICATED
0.4 kbp LONG MILLET DNAs AT DIFFERENT
TEMPERATURES OF REASSOCIATION

- a Barn yard millet
- b Little millet
- c Great millet

Reassociations were carried out as detailed in the text. To facilitate a better comparison of rates all the curves were modelled as a single component reassociating with second order kinetics. The solid lines through the data points represent the best least squares solutions allowing all parameters to free float.

- (● — ●) Reassociation at 55°C
- (○ — ○) Reassociation at 62°C
- (□ — □) Reassociation at 68°C
- (X — X) Reassociation at 75°C

FIGURE 4 · 6



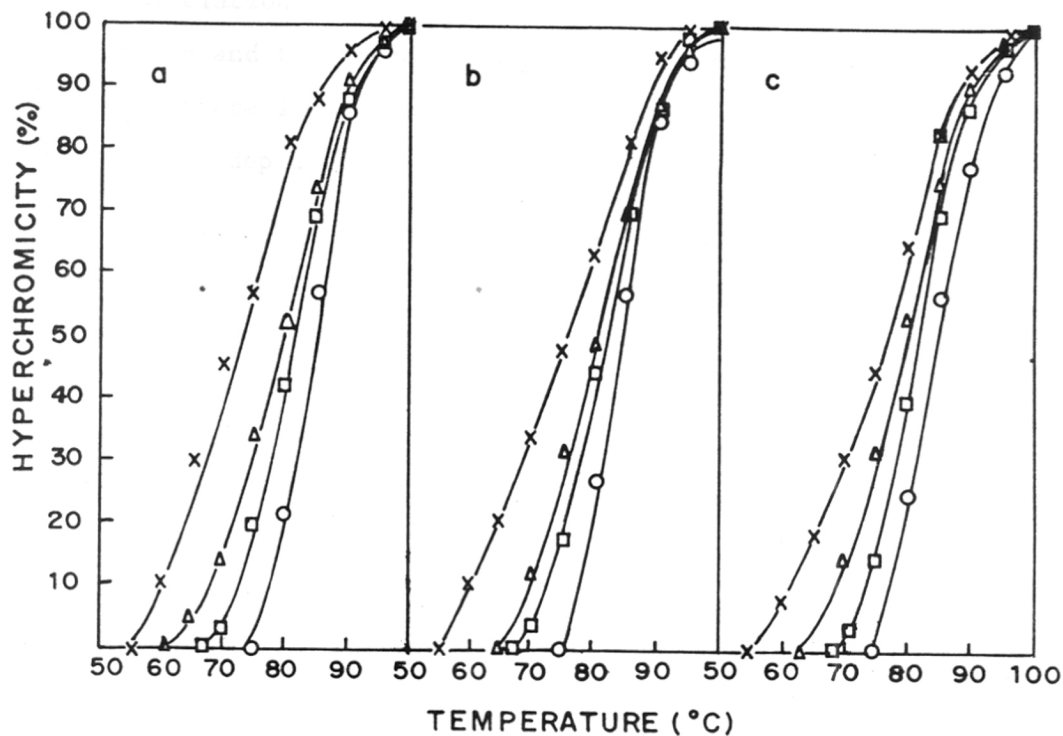


Figure 4.7 THERMAL DENATURATION PROFILES OF DNA DUPLEXES FORMED WHEN REASSOCIATION IS CARRIED OUT AT DIFFERENT INCUBATION TEMPERATURES

- a Barn yard millet
- b Little millet
- c Great millet

(X—X) DNA duplexes obtained at 55°C

(Δ—Δ) DNA duplexes obtained at 62°C

(□—□) DNA duplexes obtained at 68°C

(O—O) DNA duplexes obtained at 75°C

The DNAs were reassociated to the required Cot value and then remelted.

were considered to be the same as that at 62°C (Chapter III). The reassociation rate was corrected for the effect of temperature and the decreased extent of reassociation (11).
 and 4.5
 The results are summarized in Table 4.4 /and reassociation profiles are depicted in Figure 4.6.

In all the three millet species, the fraction recorded as 'repetitive' reduces as conditions of reassociation are made more stringent. The extent of reduction is, however, different in the three millet species. In barn yard millet, the proportion of the very fast reassociating component remains in the range of 5.5 - 9.9% at all temperatures of reassociation. However, the slow reassociating component decreases gradually from 23.9% to 8.1% as the temperature is raised from 55°C to 75°C. In little millet, the very fast component decreases almost two-fold (from 13.5% to 6.9%) as the temperature of reassociation is raised from 55°C to 75°C. The slow component shows a decrease of about 8% within this temperature rise. In great millet, the very fast component shows a three-fold decrease. At the limit Cot value of 20 M.s., the decrease in reassociation is from 51.2% to 34.7%. Thus, in all the three millets, the fraction scored as repetitive decreases as temperature of reassociation is increased.

Apart from the changes in the proportions of the

TABLE 4.5 - KINETIC ANALYSIS OF 0.4 kbp LONG MILLET DNAs AT DIFFERENT CRITERIA OF TEMPERATURE STRINGENCY

	1	2	3	4	5	6	7	8	9
		Frac- ^a tion	K obser- ^b ved	Cot 1/2 ^c observed	Cot 1/2 ^d adjusted	Cot 1/2 ^e pure	K pure ^f	Kinetic ^g complex- ity in base pairs	Frequency ^h of repeti- tion per IC genome
				M.s.	M.s.				
<u>Barn yard millet</u>									
<u>55°C (RMS error 0.014)</u>									
Cot 2.70x10 ⁻³		0.0684	-	-	-	-	-	-	-
Cot 2.70x10 ⁻³ to 1.00x10 ¹		0.2385	9.877x10 ¹	1.013x10	1.013x10 ⁻²	2.416x10 ⁻³	4.139x10 ²	2.05x10 ³	1.688x10 ⁵
Cot 1.00x10 ¹		0.7433	-	-	-	-	-	-	-
<u>62°C (RMS error 0.067)</u>									
Cot 2.65x10 ⁻³		0.0552	-	-	-	-	-	-	-
Cot 2.65x10 ⁻³ to 1.00x10 ¹		0.1864	4.661x10 ¹	2.146x10 ⁻²	2.146x10 ⁻²	4.000x10 ⁻³	2.500x10 ²	3.396x10 ³	7.968x10 ⁴
Cot 1.00x10 ¹		0.7789	-	-	-	-	-	-	-

.....Contd.

TABLE 4.5 (Contd.)

1	2	3	4	5	6	7	8	9
<u>68°C (RMS error</u> <u>0.018)</u>								
Cot 2.29x10 ⁻³	0.0992	-	-	-	-	-	-	-
Cot 2.29x10 ⁻³ to 2.10x10 ⁻¹	0.1187	2.400x10 ⁰	4.167x10 ⁻¹	3.584x10 ⁻¹	4.254x10 ⁻²	2.351x10 ¹	3.612x10 ⁴	4.771x10 ³
Cot 1.0x10 ¹	0.7827	-	-	-	-	-	-	-
<u>75°C (RMS error</u> <u>0.015)</u>								
Cot 4.00x10 ⁻³	0.0808	-	-	-	-	-	-	-
Cot 4.00x10 ⁻³ to 1.00x10 ¹	0.0812	1.076x10 ⁰	9.294x10 ⁻¹	5.576x10 ⁻¹	4.528x10 ⁻²	2.209x10 ¹	3.844x10 ⁴	3.067x10 ³
Cot 1.00x10 ¹	0.8384	-	-	-	-	-	-	-
<u>Little millet</u>								
<u>55°C (RMS error</u> <u>0.050)</u>								
Cot 2.24x10 ⁻³	0.1350	-	-	-	-	-	-	-
Cot 2.24x10 ⁻³ to 1.50x10 ¹	0.3400	1.311x10 ⁰	7.513x10 ⁻¹	7.513x10 ⁻¹	2.554x10 ⁻¹	3.914x10 ⁰	2.168x10 ⁵	2.622x10 ²
Cot 1.50x10 ¹	0.5250	-	-	-	-	-	-	-
<u>62°C (RMS error</u> <u>0.030)</u>								
Cot 3.42x10 ⁻³	0.0954	-	-	-	-	-	-	-
Cot 3.42x10 ⁻³ to 1.50x10 ¹	0.3042	1.013x10 ⁰	9.872x10 ⁻¹	9.675x10 ⁻¹	2.943x10 ⁻¹	3.398x10 ⁰	2.499x10 ⁵	2.036x10 ²
Cot 1.50x10 ¹	0.6004	-	-	-	-	-	-	-

Contd..

TABLE 4.5 (Contd.)

1	2	3	4	5	6	7	8	9
<u>68°C (RMS error</u>								
<u>0.043)</u>								
Cot 2.86x10 ⁻³	0.0720	-	-	-	-	-	-	-
Cot 2.86x10 ⁻³	0.2884	7.792x10 ⁻¹	1.283x10 ⁰	1.129x10 ⁰	3.256x10 ⁻¹	3.071x10 ⁰	2.764x10 ⁵	1.745x10 ²
to 1.50x10 ¹								
Cot 1.50x10 ¹	0.6396	-	-	-	-	-	-	-
<u>75°C (RMS error</u>								
<u>0.011)</u>								
Cot 2.66x10 ⁻³	0.0685	-	-	-	-	-	-	-
Cot 2.66x10 ⁻³	0.2671	4.357x10 ⁻¹	2.295x10 ⁰	1.400x10 ⁰	3.739x10 ⁻¹	2.675x10 ⁰	3.174x10 ⁵	1.407x10 ²
to 1.50x10 ¹								
Cot 1.50x10 ¹	0.6644	-	-	-	-	-	-	-
<u>Great millet</u>								
<u>55°C (RMS=</u>								
<u>0.055)</u>								
Cot 2.32x10 ⁻³	0.1728	-	-	-	-	-	-	-
Cot 2.32x10 ⁻³	0.5119	3.610x10 ⁰	2.770x10 ⁻¹	2.687x10 ⁻²	1.376x10 ⁻¹	7.267x10 ⁰	1.168x10 ⁵	1.336x10 ³
to 2.00x10 ¹								
Cot 2.00x10 ¹	0.3170	-	-	-	-	-	-	-
<u>62°C (RMS=</u>								
<u>0.048)</u>								
Cot 3.19x10 ⁻³	0.1168	-	-	-	-	-	-	-
Cot 3.19x10 ⁻³	0.3835	2.394x10 ⁰	4.177x10 ⁻¹	4.094x10 ⁻¹	1.570x10 ⁻¹	6.369x10 ⁰	1.333x10 ⁵	8.769x10 ²
to 2.00x10 ¹								
Cot 2.00x10 ¹	0.5026	-	-	-	-	-	-	-

Contd....

TABLE 4.5 (Contd.)

	1	2	3	4	5	6	7	8	9
68°C (RMS error 0.029)									
Cot 2.21×10^{-3}	0.0910	-	-	-	-	-	-	-	-
Cot 2.21×10^{-3} to 2.00×10^{-1}	0.3560	1.786×10^0	5.599×10^{-1}	4.871×10^{-1}	1.734×10^{-1}	5.767×10^0	1.472×10^5	7.370×10^2	-
Cot 2.00×10^{-1}	0.5543	-	-	-	-	-	-	-	-
75°C (RMS error 0.026)									
Cot 2.56×10^{-3}	0.0698	-	-	-	-	-	-	-	-
Cot 2.56×10^{-3} to 2.00×10^{-1}	0.3466	8.140×10^{-1}	1.229×10^0	7.497×10^{-1}	2.598×10^{-1}	3.849×10^0	2.206×10^5	4.789×10^2	-
Cot 2.00×10^{-1}	0.5844	-	-	-	-	-	-	-	-

^{a,b}Values obtained from the optical reassociation curves in Fig. 4.9

^cCot 1/2 observed = (K observed)⁻¹

^dCot 1/2 adjusted = Cot 1/2 observed x rate relative to optimum rate (obtained from Table 4.4)

^eCot 1/2 pure = Cot 1/2 adjusted x fraction of genome

^fK pure = (Cot 1/2 pure)⁻¹

^gValues obtained from the standard relationship for E.coli DNA where Cot 1/2 is 5.3 (under our experimental conditions) and the number of nucleotide pairs is 4.5×10^6 .

^hCot 1/2 of unique DNA described in Table 3.1 of Chapter 3 divided by Cot 1/2 values of given fraction.

various repetitive fractions as a function of incubation temperature, differences are also observed with respect to their kinetic complexity, copy number and rate of reassociation (Table 4.5). The kinetic complexity, for example, increases with an increase in incubation temperature in all the cases. This increase is 18.75 fold in barn yard millet and is only 1.5 to 1.9 fold in little millet and great millet. The copy number, on the other hand, decreases 55 fold in barn yard millet, 1.9 fold in little millet and 2.8 fold in great millet. Finally, the reassociation rates are observed to decrease with increase in stringency of reassociation.

Methylation status of millet DNAs as determined using restriction endo-nucleases

The restriction enzymes Msp I and Hpa II are isoschizomers and recognize the sequence 5' CCGG 3'. However, Hpa II will not cleave the 5' CCGG 3' sequence if the internal cytosine residue is methylated, whereas Msp I will. The converse result is obtained when the external cytosine is methylated. Thus, these enzymes readily permit differentiation of methylated ^{from unmethylated} DNA, and give some information on the position of the methylated cytosine residues. The restriction enzyme Hae III recognizes the sequence 5' GGCC 3' cutting between the dinucleotide 5' Gp C 3', only when the internal cytosine residue is not

Figure 4.8

DIGESTIONS OF MILLET DNAs USING RESTRICTION ENZYMES.
4-5 μ gms OF MILLET DNAs WERE DIGESTED WITH ABOUT 10 UNITS
OF THE ENZYME FOR 1-2 hr AT 37°C. DNA FRAGMENTS WERE SEPARATED
BY ELECTROPHORESIS ON 1% AGAROSE GELS AT 50 V, FOR 1-2 hr.
LAMBDA DNA DIGESTED WITH HIND III AND PHAGE ϕ X174RF DNA DIGESTED
WITH Hae III SERVED AS MOLECULAR WEIGHT MARKERS.

Lane a: Lambda phage DNA digested with Hind III (23.3 kbp, 9.5 kbp,
6.4 kbp, 4.2 kbp, 2.2 kbp and 1.8 kbp).

Lane b,f,j & n: Barn yard millet DNA, fox tail millet DNA, little
millet DNA and great millet DNA controls.

Lane c,g,k,o: Barn yard millet, fox tail millet, little millet
and great millet DNAs digested with Msp I.

Lane d,h,l,p: Barn yard millet, fox tail millet, little millet
and great millet DNAs digested with Hpa II.

Lane e,i,m,q: Barn yard millet, fox tail millet, little millet and
great millet DNAs digested with Hae III.

Lane r: Phage ϕ X174 RF DNA digested with Hae III (1.32 kbp, 1.078 kbp,
0.872 kbp, 0.606 kbp, 0.310 kbp) and molecular weight marker of lane a.

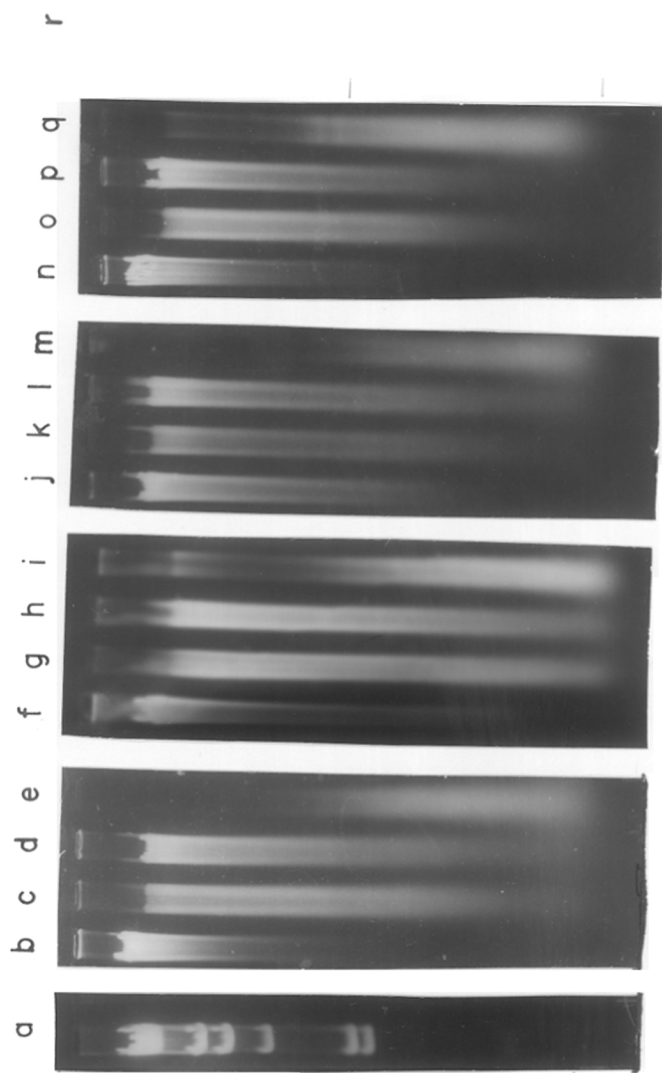


FIGURE 4·8

methylated. The above three enzymes were used to digest the millet DNAs to get an idea of the status of methylation in these DNAs. From Figure 4.8, it can be seen that the millet DNAs are extensively digested with Msp I to give a smear in entire region of 0.1 - 9.0 kbp indicating very few of the external cytosine residues in the sequence 5' CCGG 3' are methylated. There is almost no digestion of millet DNAs with Hpa II suggesting that several internal cytosine residues in the sequence 5' CCGG 3' are methylated. The extensive digestion with ~~Hae III~~ ^{Hae III} indicates that very few cytosine residues in the sequence 5' GGCC 3' are methylated. It thus appears that much of the methylation we see in millet DNAs is restricted to the internal cytosine residues of the sequence 5' CCGG 3'.

Occurrence of 5-methylcytosine as determined by HPLC

Since HPLC is a very sensitive technique, it was used to assess the presence of 5-methyl cytosine in the DNA of the four millet species. From Figure 4.9Ab, it can be seen that acid hydrolysed calf thymus DNA, used as a control, shows about 1% 5-methylcytosine. All millet DNAs also show the presence of 5-methylcytosine. Furthermore, it accounts for 3% of the total bases in barn yard millet, little millet and great millet and 8% in fox tail millet (Figure 4.9Ba, 4.9Ca,b,c). The identity of the 5-methylcytosine peak was

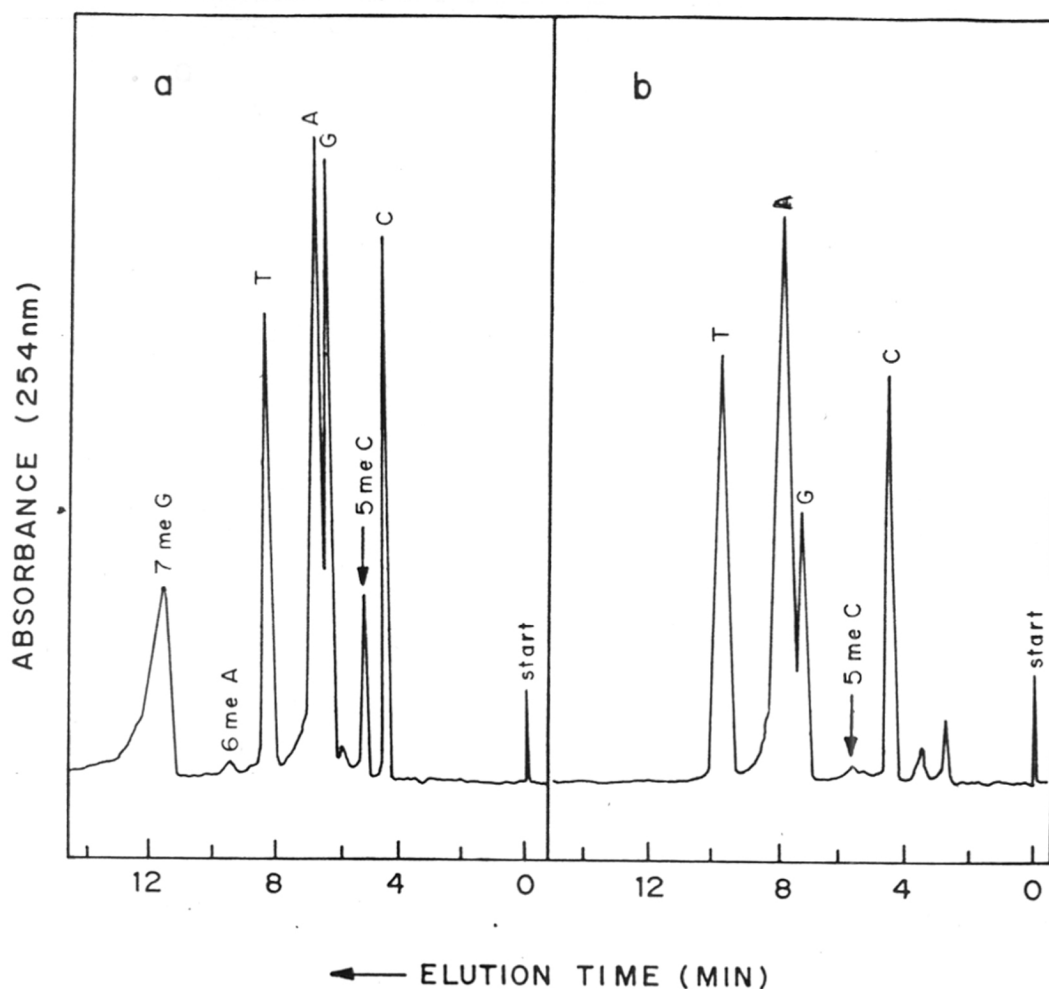


Figure 4.9A HPLC SEPARATION OF

- a A mixture of authentic bases adenine (A), guanine (G), cytosine (C), thymine T, 5-methyl cytosine (5 meC), 6-methyl adenine (6 meA) and 7-methyl guanine (7 meG).
- b Perchloric acid hydrolysed calf thymus DNA. The experimental details are outlined in the text.

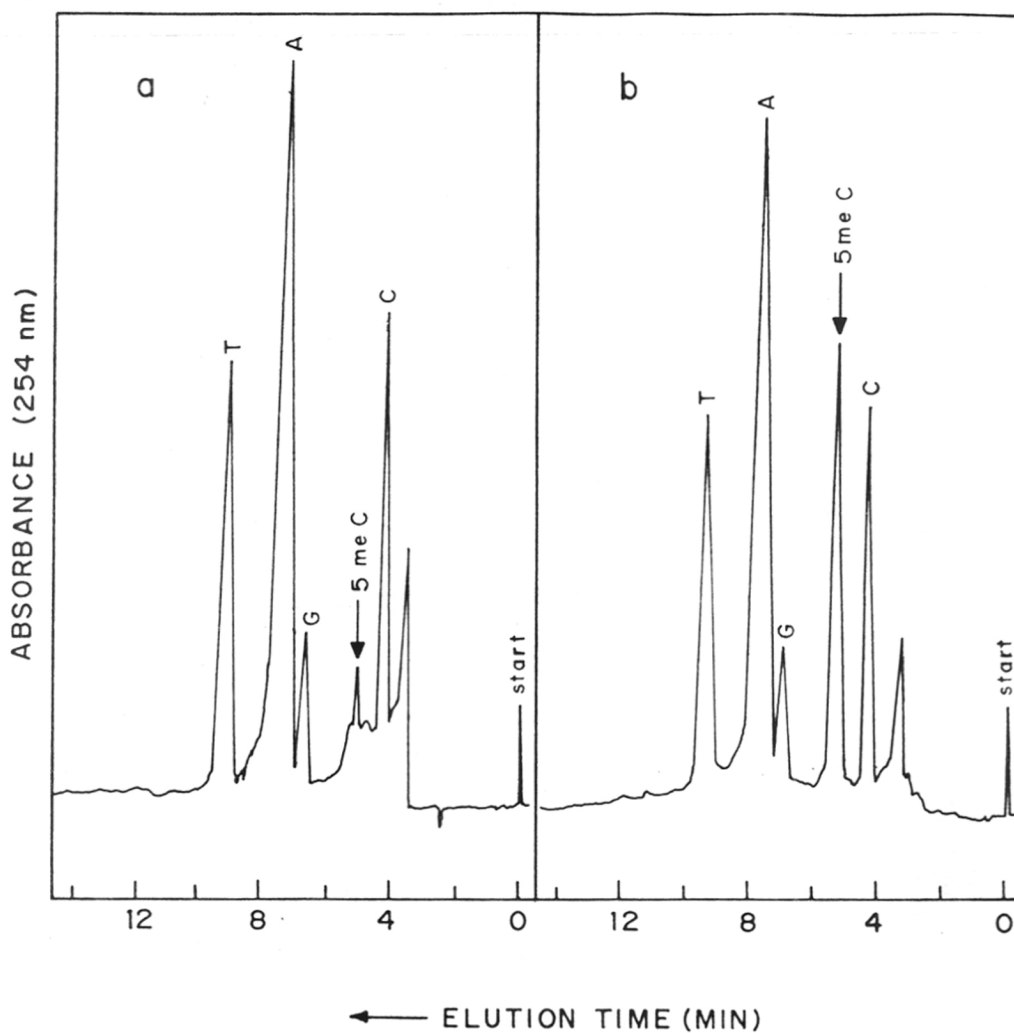


Figure 4.9B HPLC SEPARATION OF

- a Perchloric acid hydrolysed barn yard millet DNA.
- b Peak enhancement observed following the injection of an authentic sample of 5-methyl cytosine along with the hydrolysate of barn yard millet DNA.

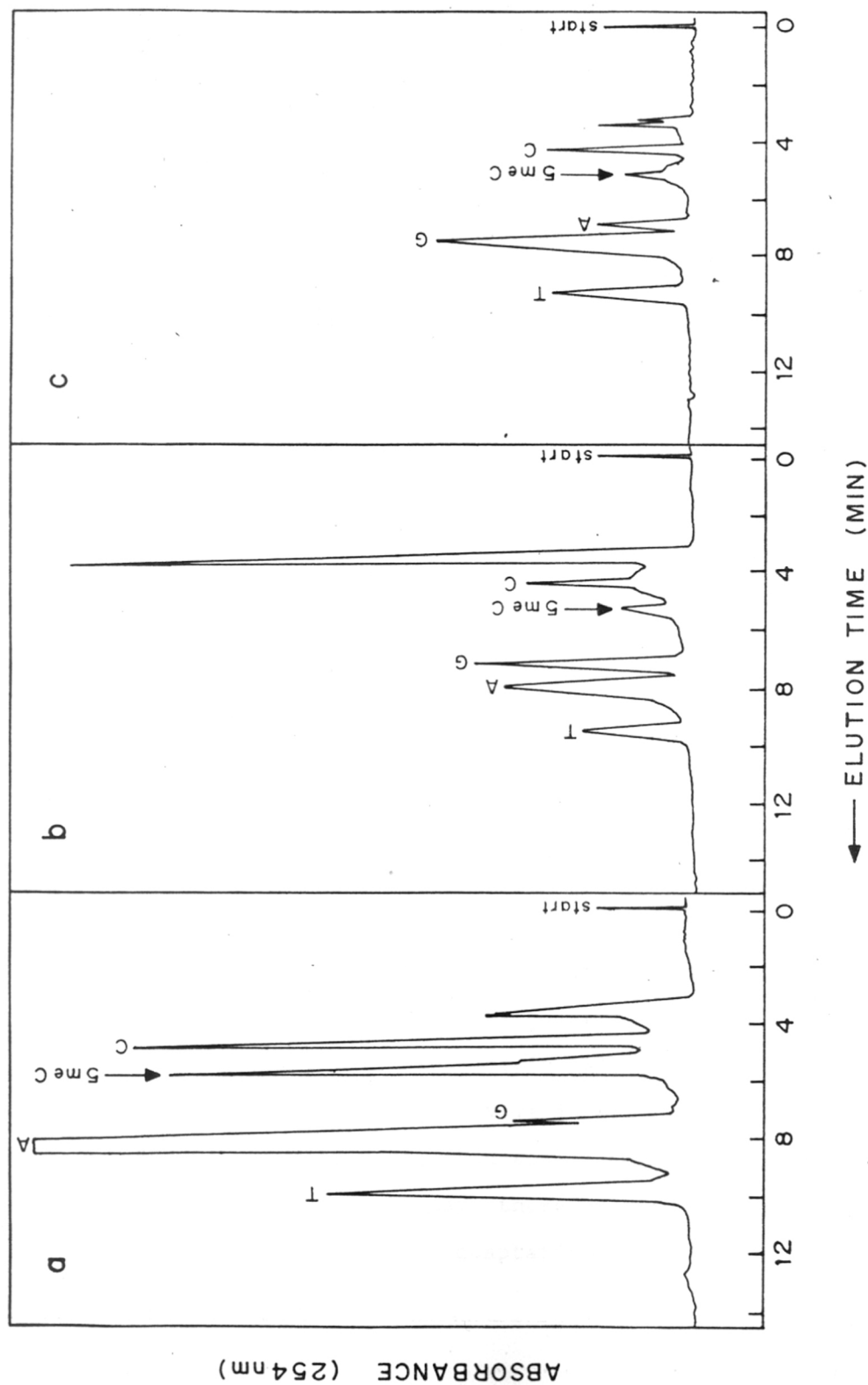


Figure 4.9C HPLC SEPARATION OF

confirmed by peak enhancement in case of barn yard millet (Figure 4.9Bb).

DISCUSSION

Possible function of GC rich sequences melting at high temperatures

Analysis of the individual Cot fractions which include DNA sequences reannealing in a narrow Cot range has enabled us to pinpoint specific differences which otherwise cannot be detected. Native 0.4 kbp long DNAs of the four millet species have melting profiles that are smooth and monophasic. However, the repetitive Cot fractions show a biphasic melting behaviour in all the millets except great millet. The minor high melting component (T_m : 96-100°C) constitutes 3-8% of the genomes of these millets and is present in the rapidly reassociating DNA (Cot 0.1 M.s.). Such a high melting component was earlier shown to be present in pearl millet (12). Its presence has not been shown in other grasses like wheat, oat, rye, barley, maize, rice or finger millet. The four millets, possessing the high melting component, are closely related (13); are adapted to a dry climate; can survive in sparse rainfall conditions and can grow in poor, thin soil (14). The high melting G + C rich component may, therefore, be playing an important role in the ecological adaptations of millets.

Since it is known that modification of cytosine by methylation results in a more radiation resistant derivative (15), we looked into the methylation of DNA sequences in these millets. Very little digestion of total millet DNAs by Hpa II suggests that the internal cytosine residues in several of the sequences 5' CCGG 3' are methylated. However, extensive digestion observed with Hae III indicates that the sequence 5' GGCC 3' occurred with high frequency with relatively less methylation in the 5' GpC 3' dinucleotide. Preliminary characterization of total millets DNAs by HPLC has also shown the presence of 3% 5-methyl cytosine in barn yard millet, little millet and great millet and about 8% in fox tail millet. However, no information as to whether the cytosine residues in the high melting DNA fraction are selectively more methylated than in the rest of the DNA is available. If this happens to be the case, it would certainly be a good mechanism to enable the millets to survive the extreme ecological conditions.

Sequence divergence of repeated DNA sequences in Gramineae

Thermal denaturation studies have indicated that the melting temperatures of total repetitive fractions are lower than that of the corresponding native DNAs in the four millets. The lowering in T_m is about 6.4 - 7.4°C. It is generally believed that the degree of base mismatch in a

given DNA fraction gives an indication of the extent of sequence divergence. Hence the sequence divergence of the repeat fractions in these millets is about 6-8% and is moderate.

In family Gramineae, data on thermal stability of repetitive duplexes have been reported for wheat, oat, rye, barley, maize, finger millet and pearl millet (12,16-19). Wheat, oat, rye, barley and maize show as much as 12-15% base mismatch in the repetitive duplexes indicating an extensive sequence divergence. The repetitive DNA fraction of finger millet shows a sequence divergence of 10-13% and thus resembles wheat, oat, rye, barley and maize. Sequence divergence in pearl millet is 9% and it resembles the millets in the present investigation. Sequence divergence accumulates with time and hence can give an idea about the age of a repeat. It, therefore, appears that the repeat sequences of millets (except finger millet) are of relatively recent origin than those in other grasses mentioned above.

The intermediately repetitive DNA of millets shows a greater sequence divergence than that in the highly repetitive class of DNA sequences. In maize, however, the highly repetitive DNA is twice as divergent as the intermediately repetitive DNA (19). Here, it is assumed that the highly repetitive DNA has originated from the intermediately repetitive DNA (showing moderate divergence)

and the latter from single copy DNA sequences (no divergence). Subsequent divergence at equal rates in both classes of repeated DNA results in the occurrence of highly divergent and highly repetitive sequences interspersed with precisely matched moderately repetitive DNA. In the millets studied here, it appears that the moderately repetitive DNA has originated from the highly repetitive DNA. A few sequences of the latter class of DNA sequences may have been translocated to new sites on the chromosome and may have diverged with time by amplification and base substitution. Highly repetitive DNA thus appears to maintain its sequence homogeneity by 'cross over fixation'.

Absence of distinct classes of repetitive DNA in millets

Analysis of mini Cot curves has enabled to detect the presence of an additional fast reassociating DNA component in the highly repetitive DNA class of only great millet. Since no subclass of repetitive DNA other than that observed in the reassociation kinetics of the total 0.4 kbp kinetic curves of millets is identified, it indicates that the repetitive classes in millets are not completely distinct. A continuum in Cot $1/2$ values suggests an interspersion of highly repetitive and intermediately repetitive DNA components. Maize is the only other Gramineae

plant species where mini Cot curves have been generated to obtain a direct estimate of rate constants of individual repetitive sequence classes (19). It has been shown here that each subclass of repetitive DNA is also contaminated to some extent with the other classes.

Reassociation kinetics studies at different incubation temperatures

Analysis of DNA reassociation kinetics at different incubation temperatures has been carried out only in one Gramineae plant species namely, barley. This species has been shown to have predominantly homogeneous families of repeat sequences (5,6). The basis of this approach is that if repeated DNA sequences are capable of reassociating with homologous DNA sequences representing a range of sequence divergence, changes in the temperature of reassociation (or criterion) would alter the number of available pairing sequences and therefore change the observed kinetics of reassociation. There are different views regarding composition of evolutionarily young and old families. According to Bendich and Anderson (5) all repeat families, young and old, should have the same copy number while according to Preisler and Thompson (20,21), older families should have lower copy number than the younger families due to sequence divergence and dispersion. Our data indicate

TABLE 4.6 - 'FOSSIL' REPEATS AND 'TRUE' SINGLE COPY SEQUENCES IN MILLET GENOMES

Plant species	Genome ^a size in nucleotide pairs	% single copy at permissive criterion (55°C)	% single copy ^c at standard criterion (62°C)	%fossil repeats	% true single copy	Kinetic ^f complexity of true single copy in nucle- otide pairs.
<u>Barn yard millet</u>	6.81 x 10 ⁸	0.7433	0.7789	4.57	95.43	6.499 x 10 ⁸
<u>Little millet</u>	7.99 x 10 ⁷	0.5250	0.6004	12.56	87.44	6.987 x 10 ⁷
<u>Great millet</u>	3.05 x 10 ⁸	0.3170	0.5026	36.93	63.07	1.924 x 10 ⁸

^aObtained from Table 3.1 of Chapter III^{b,c}From Table 4.5^d% Calculated as $\frac{c - b}{c} \times 100$ ^e% Calculated as $\frac{b}{c} \times 100$ ^f% Calculated as genome size x fraction of true single copy i.e. a x e/100.

that stable families reassociating at higher temperatures have a lower copy number and they reassociate slower than those reassociating at lower temperatures. These observations indicate that barn yard millet, little millet and great millet consist essentially of heterogeneous families of repeats. The increased complexity and decreased frequency of repetition when reassociation is carried out at rigorous criterion indicate that the family members of repetitive class in these millets are sufficiently related to form stable duplexes at the standard criterion of 62°C.

Applying definitions of Murray et al. (3) of ^(Table 4.6) fossil repeats and true single copy, we find that in barn yard millet almost all the single copy (95%) is true single copy. This value progressively decreases in little millet and great millet where the proportion of true single copy is 85% and 63% respectively. Thus, essentially only one-third (or less) of the single copy at permissive criteria behaves as fossil repeats leaving behind two-thirds or more of the single copy sequences to be true sequences. A smaller fraction of fossil repeats and a larger fraction of true single copy in these millet genomes indicates a relatively low rate of turn over in these genomes. Low turn over rates should give rise to homogeneous families of repeats (3). However, the situation in millets appears to be quite paradoxical in that they show heterogeneous families of

repeated DNA sequences. Thus, the notions of homogeneous versus heterogeneous repeat family organisations may be just two conceptual extremes and millets may be the instances of genomes falling along a continuum between them. It will be an interesting proposition to see how these repeat sequence families are organized with respect to single copy sequences since sequence organization patterns can also be used to interpret the rate of turn over of a genome.

REFERENCES

- 1 Murray, M.G., Cuellar, R.E. and Thompson, W.F. (1978). *Biochemistry*, 17: 5781-5790.
- 2 Murray, M.G., Palmer, J.D., Cuellar, R.E. and Thompson, W.F. (1979). *Biochemistry*, 18: 5259-5266.
- 3 Murray, M.G., Peters, D.L. and Thompson, W.F. (1981). *J. Mol. Evol.*, 17: 31-42.
- 4 Thompson, W.F., Murray, M.G. (1980). In: *Proceedings of the 4th John Innes Symposium*, Davies, D.R. and Hopwood, D.A. eds., pp.31-45, John Innes Charity.
- 5 Bendich, A.J. and Anderson, R.S. (1977). *Biochemistry*, 16: 4655-4663.
- 6 Bendich, A.J. (1979). In: *Molecular Biology of plants*, Rubenstein, I., Phillips, R., Green, C.E. and Gengenbach, B.G. eds. 1-30. Academic Press, New York.
- 7 Zimmerman, J.L., Goldberg, R.B. (1977). *Chromosoma (Berl.)*, 59: 227-252.
- 8 Wimpee, C.F., Rawson, J.R.Y. (1979). *Biochim. Biophys. Acta*, 562: 192-206.
- 9 Britten, R.J., Graham, D.E., Neufeld, B.R. (1974). In: *Methods in Enzymology*, Grossman, L. and Moldave, K. eds. Vol. 29E, pp.363-418, Academic Press, New York.
- 10 Mandel, M. and Marmur, J. (1968). In: *Methods in Enzymology*, Grossman, L. and Moldave, K. eds. Vol. 12B pp.195-206.
- 11 Bonner, T.I., Brenner, D.J., Neufeld, B.R. and Britten, R.J. (1973). *J. Mol. Biol.*, 81: 123-135.
- 12 Deshpande, V.G. and Ranjekar, P.K. (1980). *Hoppe-Seyler's Z. Physiol. Chem.*, 361: 1223-1233.
- 13 Hutchinson, J. (1973). In: *The families of flowering plants*, 3rd ed., Oxford University Press, Oxford.
- 14 Purseglove, J.W. (1972). In: *Tropical Crops Monocotyledons 1*, Longman Group Ltd., London.
- 15 Lindahl, T. (1981). *Nature (London)*, 290: 363-364.

- 16 Ranjekar, P.K., Lafontaine, J.G. and Pallotta, D (1974). *Chromosoma (Berl.)*, 48: 427-440.
- 17 Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976). *Biochim. Biophys. Acta*, 425: 30-40.
- 18 Smith, D.B. and Flavell, R.B. (1975). *Chromosoma (Berl.)*, 50: 223-242.
- 19 Hake, S. and Walbot, V. (1980). *Chromosoma (Berl.)*, 79: 251-270.
- 20 Preisler, R.S. and Thompson, W.F. (1981a). *J. Mol. Evol.*, 17: 78-84.
- 21 Preisler, R.S. and Thompson, W.F. (1981b). *J. Mol. Evol.*, 17: 85-93.

CHAPTER V

DNA SEQUENCE ORGANIZATION IN MILLET
GENOMES

SUMMARY

DNA sequence organization patterns have been studied in four millet plant species, namely, barn yard millet, fox tail millet, little millet and great millet. The reassociation kinetics of 1.5 kbp and 20 kbp long DNA fragments have shown that 85-95% of the millet genomes consist of alternating single copy and repetitive DNA sequences. The interspersed repetitive sequences are composed of long repeats (> 1.5 kbp) in all the four species. In addition, barn yard millet, little millet and great millet genomes consist of varying lengths of interspersed short repeats. Fox tail millet essentially shows the presence of long repeats. The length of interspersed single copy sequences varies in the narrow range of 1.7 - 2.8 kbp. The high thermal stability of the repetitive duplexes isolated from DNAs of different fragment lengths in all the four millets indicates that these duplexes are present in long arrays consisting of smaller elements varying slightly in their nucleotide sequences. The short repeats in barn yard millet and little millet are presumably of recent origin and have probably evolved from the long repeated DNA sequences. The rate of 'turn over' of the total genomes of these four species appears to be low compared to that in wheat, oat, rye, barley and maize.

INTRODUCTION

The genomes of most eukaryotic organisms contain DNA sequences in multiple copies (1). These repeat sequences are broadly grouped into two categories, namely satellite DNA and interspersed repetitive DNA (2). Generally satellite sequences are clustered in distinct regions of the chromosome and consist of tandemly linked simple sequences that have been attributed a functional role (3). On the other hand repetitive sequences are interspersed to varying degrees with single copy sequences and are assumed to play a regulatory role in gene expression (4-6). They have also been implicated in evolutionary processes (7). Studies on DNA sequence organization in plants have revealed a great deal of diversity with respect to organization patterns and length of interspersed repeats. A number of plant genomes have been characterized which show instances where their patterns of organization fall along a continuum between the two conceptual extremes of 'short' term (8) and 'long' term (9) models of interspersion. Interspersed repeats have been reported to have sizes ranging between 0.5 - 10 kbp. Current evidence suggests that different and complex modes of sequence organization can emerge even in closely allied phyletic lineages in plants. It will, therefore, be useful to compare the data from a wide variety of organisms including those with

varied genome size and repetitive DNA content, to understand the significance of the observed interspersal patterns and the limits within which different patterns vary. These studies, will lead to a better understanding of the amplification and deletion events responsible for the origin and evolution of repetitive DNA.

After extensive investigations about the heterogeneity in base composition (Chapter II), reassociation kinetics properties (Chapter III) and properties and rates of evolution of repeated DNA sequences (Chapter IV), further studies were directed towards the organization of repeat and single copy DNA sequences in the four millet species. These studies were undertaken to assess the changes that may have occurred in the genomes of these millets during evolution and thereby examine the functional constraints and/or flexibility imposed upon them by selective or random forces. For these studies, the reassociation kinetics of long DNA fragments were first determined. The reassociated duplexes were characterized by thermal denaturation. Sizing of S1 nuclease resistant duplexes was carried out by agarose column chromatography.

MATERIALS AND METHODS

Chemicals: PIPES buffer (Piperazine-N-N'-bis(2-ethane

sulfonic acid)) was obtained from Sigma Chemical Co., MO, USA. S1 nuclease was obtained either from Boehringer Mannheim, West Germany or Bethesda Research Laboratories, USA. Bio-Gel A-50 m (100-200 mesh) was obtained from Bio-Rad Laboratories, California, USA.

DNA isolation

DNA from the four millet species was isolated according to the procedure outlined in Chapter II.

DNA shearing and sizing

DNA samples were first dialyzed against 0.12 M sodium phosphate buffer, pH 6.8 and sheared to an average fragment length of 0.4 kbp as described in Chapter II. Sonication using the Biosonic III/^{fitted} with 1/4th inch probe for one minute at 30 setting yielded DNA fragments of an average fragment size of 1.5 kbp. DNA fragments in the size range of 3.0 to 9.0 kbp were obtained using a VirTis 60 K homogenizer by varying the conditions of speed and time between 30,000 to 58,000 rpm and 4 to 45 minutes respectively (Figure 5.1). The DNA concentrations during all sonications and homogenizations were kept constant at 400 - 450 µg/ml. The average fragment length of sheared DNAs was determined by agarose slab gel electrophoresis in 1% agarose using Hind III digest of bacteriophage lambda DNA and Hae III digest of phage ϕ x174 RF DNA.

Figure 5.1 AGAROSE SLAB GEL ELECTROPHORESIS OF DNA FRAGMENTS OBTAINED USING VIRTIS 60 K HOMOGENIZER AND BIOSONIC III. CALF THYMUS DNA WAS USED FOR THE STANDARDIZATION OF FRAGMENT SIZES OBTAINED BY VARIOUS SONICATIONS.

Lane a,i,k: Native unsheared calf thymus DNA

Lane h: Lambda phage DNA digested with Hind III (23.3 kbp, 9.5 kbp, 6.4 kbp, 4.2 kbp, 2.2 kbp, 1.8 kbp)

Lane m: Phage ϕ X 174 RF DNA digested with Hae III (1.342 kbp, 1.078 kbp, 0.872 kbp, 0.606 kbp and 0.310 kbp).

Lane	Instrument used	Setting	Time (min)	Average DNA fragment size (kbp)
b	Virtis 60 K homogenizer	57,000 rpm	45	8.2
c	"	50,000 rpm	15	7.3
d	"	48,000 rpm	7	5.6
e	"	30,000 rpm	30	4.9
f	"	40,000 rpm	2	4.5
g	"	32,000 rpm	4	3.0
j	Biosonik III fitted with $\frac{1}{4}$ th inch probe	40 setting	1	1.5
l	Biosonik III fitted with $\frac{1}{2}$ inch probe	Max. setting	3	0.4



FIGURE 5 · 1

as molecular weight markers. The molecular weight of native, unsonicated DNA from the four millet species was estimated to be approximately 20 kbp.

Reassociation kinetics of long DNA fragments

The reassociation kinetics of long DNA fragments (1.5 kbp and 20 kbp) millet DNAs were measured using hydroxapatite chromatography in the Cot range of 1×10^{-1} to 1×10^4 M.s. The experimental points were fitted by non-linear least squares regression analysis as described in Chapter III.

Thermal stability of repetitive duplexes

Thermal denaturation studies of repetitive DNA fractions isolated from DNAs of different fragment lengths after reassociation to their respective limit Cot values were carried out in a Gilford 250 Spectrophotometer as detailed in Chapter II. In all these studies, no correction was made for solvent expansion. Several parameters such as T_m , % base mismatch, % hyperchromicity, duplex content and average duplex length were calculated according to the standard procedures (10-12).

Determination of single strand collapse

Single strand collapse was calculated according to Graham *et al.* (13). Unreassociated unique DNA which included the DNA reassociating after Cot 75 M.s., Cot 5 M.s., Cot 15 M.s. and Cot 20 M.s. of barn yard millet,

fox tail millet, little millet and great millet respectively was isolated by hydroxyapatite column chromatography. This was then melted in 0.12 M ^{sodium}/phosphate buffer, pH 6.8 and the hyperchromicity was determined to be 1.3% for barn yard millet, little millet and great millet and 1.5% for fox tail millet.

Isolation of S1 nuclease resistant repetitive duplexes

These experiments were carried out according to the procedure of Murray et al. (14) with some modifications. Native DNAs of fragment size 20 kbp were dialyzed against 0.18 M NaCl containing 0.006 M PIPES buffer, pH 6.8. The samples were denatured in boiling water bath for 10 min. and allowed to reassociate at 62°C to the limit Cot values of each millet species. After reassociation, samples were adjusted to 25 mM sodium acetate (pH 4.5), 0.1 mM ZnSO₄, 25 mM 2-mercapto-ethanol (15,16). At least 1.5 to 2 mg of DNA was used for each experiment. 10 units of S1 nuclease per ug of DNA were added and the samples incubated for one hour at 37°C to digest single stranded DNA. The reaction was terminated by adjusting the sample to 0.12 M ^{sodium}/phosphate buffer (pH 6.8) and simultaneously cooling the samples to 4°C. S1 nuclease resistant duplexes were then separated on hydroxyapatite column by elution with 0.4 M sodium

phosphate buffer (pH 6.8) for further characterization of distribution of size classes of repeated DNA sequences.

S1 nuclease resistance studies of millet DNAs incubated to Cot 0.1 M.s. were also carried out as outlined above.

Sizing of S1 nuclease resistant reassociated duplexes

(i) Agarose column chromatography

The size distribution of S1 resistant duplexes was determined by gel filtration in agarose Biogel A-50 (100-200 mesh size) column previously calibrated with calf thymus DNA of known duplex lengths. Native calf thymus DNA (>1.5 kbp) was mixed with calf thymus DNA sonicated to 0.4 kbp and was used for standardization of the column.

(ii) Preparation of the agarose column

The agarose gel matrix (fully hydrated in 0.001 M Tris EDTA buffer) was poured around 4 mm glass beads in a 92 cm x 1.5 cm column. The column was equilibrated with 0.12 M sodium phosphate buffer (pH 6.8) and was reused a number of times after thorough washing before column operations.

(iii) Elution of DNA fragments

Approximately 70-100 μ g of S1 nuclease resistant duplexes, separated by hydroxyapatite column chromatography,

was loaded on the column and elution was carried out with 0.12 M sodium phosphate buffer (pH 6.8). The effluent was continuously monitored with the help of LKB 2138 UVICORD S at 253 nm (17,18). The exclusion limit of Agarose A-50 was 1.5 kbp.

(iv) Agarose slab gel electrophoresis

The DNA fractions in the excluded volume of Agarose A-50 were sized in 1% agarose slab gels (Chapter II).

RESULTS

Reassociation kinetics of long DNA fragments

The presence or absence of interspersions of repetitive and single copy DNA sequences can be determined by comparing the hydroxyapatite binding data of long and short DNA fragments over a range of Cot values. An increase in binding to hydroxyapatite as a function of DNA fragment length is considered to be an indication of the occurrence of interspersions of repetitive and non-repetitive DNA sequences (5,19). This increase in binding can be due to the increase in fragment length as well as due to the presence of single copy DNA sequences in an unreassociated form contiguous to repetitive duplexes.

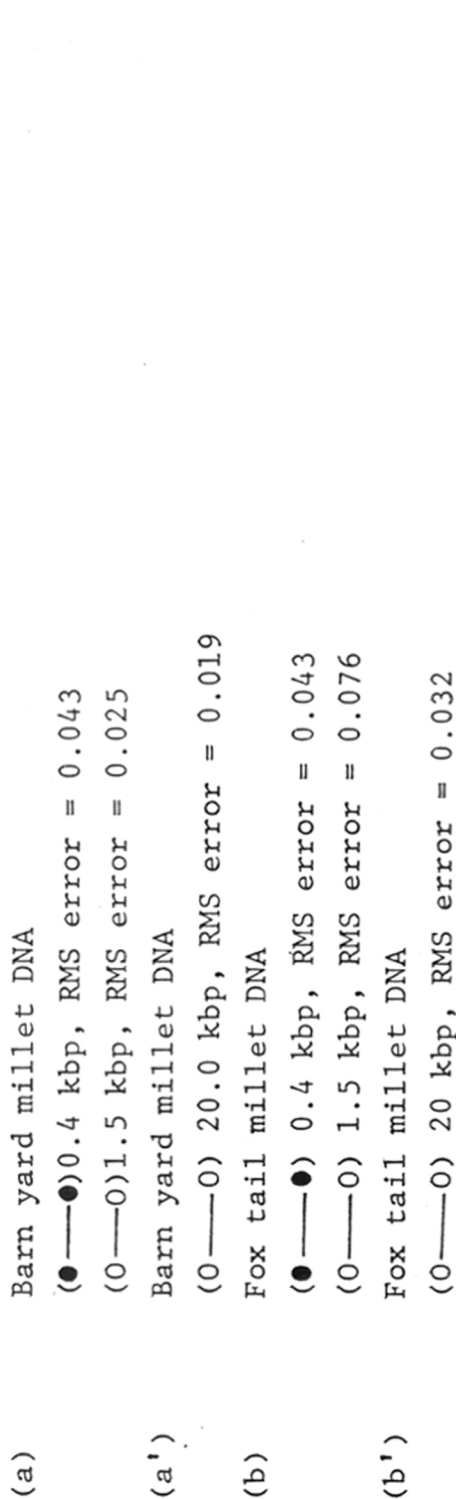
Fig 5.2 and 5.3 depict the reassociation kinetics of 1.5 kbp long and 20 kbp long DNA fragments of all the

four millets. The reassociations were studied in the Cot range of 1×10^{-1} to 1×10^4 M.s. The computer programme was used to find the best fit by allowing all parameters to free float. The results of this analysis are given in Table 5.1.

In barn yard millet, the Cot curves of 1.5 kbp and 20 kbp fragments indicated the presence of two components each reassociating with second order kinetics. We have earlier shown that the 0.4 kbp reassociation kinetics curve for barn yard millet could be modelled into three distinct kinetic components. At a Cot value of 5×10^{-2} M.s. the binding of duplexes increased from 15% at 0.4 kbp to 26% at 1.5 kbp and 63.6% at 20 kbp. At the limit repeat Cot value of 75 M.s., this increase is from 42% at 0.4 kbp to 65% at 1.5 kbp and 81% at 20 kbp. Such significant increase in binding as a function of increasing DNA fragment lengths at a low Cot value and at a limit repetitive Cot value indicates an extensive interspersion of repeated DNA sequences among themselves and with single copy DNA sequences in barn yard millet genome.

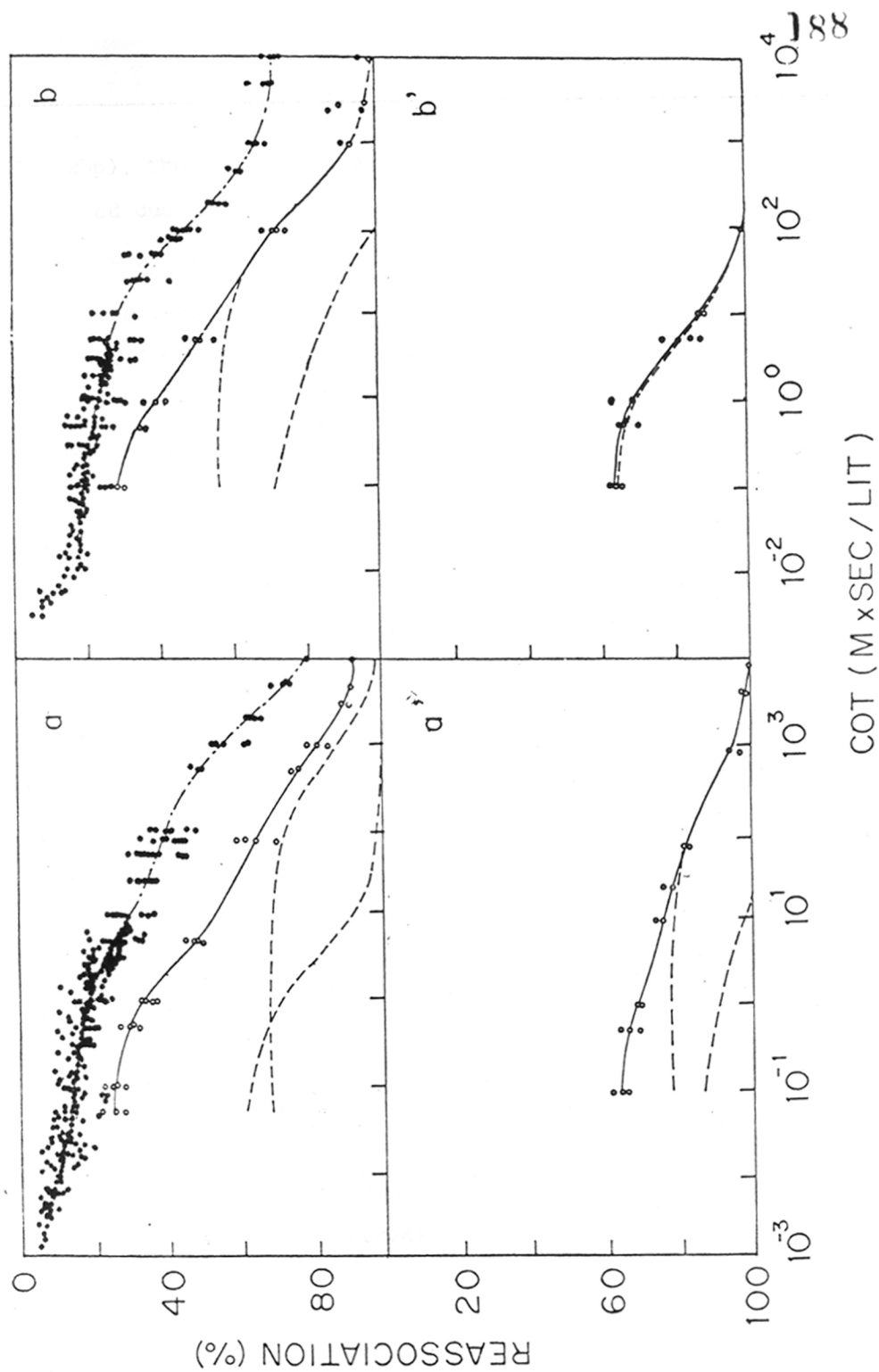
In fox tail millet, resolution of 0.4 kbp and 1.5 kbp reassociation kinetics curves into two components and of 20 kbp curve into a single component indicated the absence of a single copy component at the highest fragment size

Figure 5.2 REASSOCIATION KINETICS OF LONG DNA FRAGMENTS



The solid lines through the data points represent the best least squares fit, allowing all the parameters to free float. The lower dashed curves represent the reassociation kinetics of pure components.

FIGURE 5-2



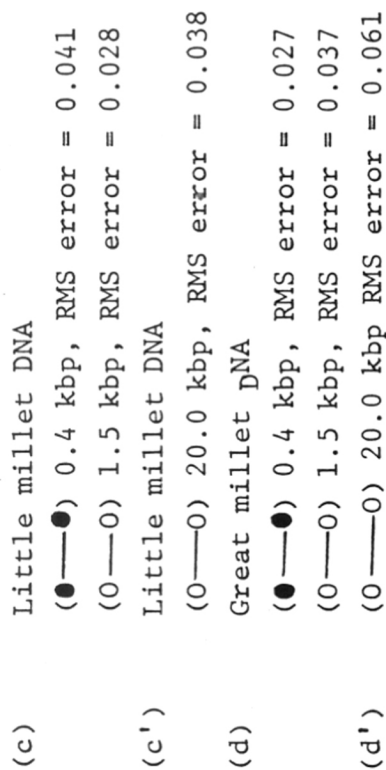
(20 kbp), the reassociation of the latter being completely obscured due to extensive interspersions of repeat and single copy components.

The reassociation kinetics of the little millet genome were resolved into two components at all the fragment lengths studied. At a Cot value of 5×10^{-2} M.s., the hydroxyapatite bound duplexes increased from 16% at 0.4 kbp to 30.5% and 62% at 1.5 kbp and 20 kbp respectively. At the limit repeat Cot value of 15 M.s., this increase was from 38% at 0.4 kbp to 66% at 1.5 kbp and 85% at 20 kbp.

The above data clearly point out that there is interspersions of repeated and single copy DNA sequences in fox tail millet and little millet genomes.

In case of great millet, the repetitive DNA could be resolved as only a single component at 1.5 kbp and 20 kbp as against the presence of two distinct second order components at a fragment size of 0.4 kbp indicating that the highly repetitive and intermediately repetitive DNA components are interspersed. Extensive interspersions of these components is indicated by the fact that at a Cot value of 5×10^{-1} , the binding of duplexes to hydroxyapatite increases from 25% at 0.4 kbp to 42% at 1.5 kbp to 55% at 20 kbp. At the ^{limit} ~~repetitive~~ Cot value of 20 M.s., the percent reassociation increases from 52% at 0.4 kbp to

Figure 5.3 REASSOCIATION KINETICS OF LONG DNA FRAGMENTS



The solid lines through the data points represent the best least squares fit, allowing all the parameters to free float. The lower dashed curves represent the reassociation kinetics of pure components.

FIGURE 5.3

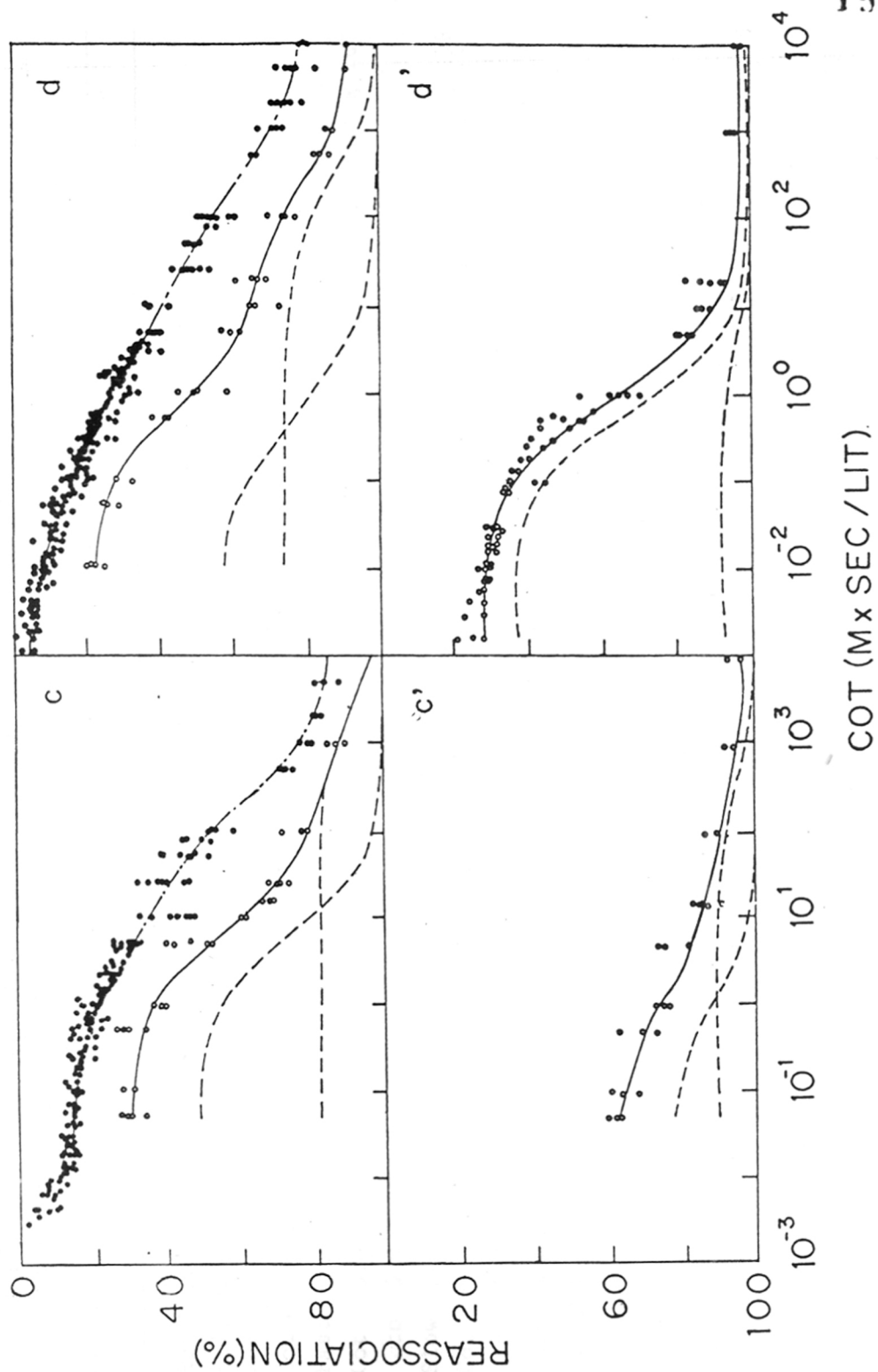


TABLE 5.1 - KINETIC COMPLEXITY OF LONG DNA FRAGMENTS

Species, DNA fragment size and DNA components	Fraction ^a of the genome	K observed ^b	K pure ^c	Frequency ^d of repeti- tion per 1 C genome	Kinetic ^e complexity in nucleotide pairs
	2	3	4	5	6
<u>Barn yard millet</u>					
<u>1.5 kbp</u>					
Very fast					
Cot $< 5.0 \times 10^{-2}$	0.2521	-	-	-	-
Fast					
Cot 5.0×10^{-2}	0.3844	3.36×10^1	8.74×10^{-1}	1.92×10^2	9.71×10^5
to 5.0×10^1					
Slow					
Cot 5.0×10^1	0.3161	1.75×10^{-3}	5.54×10^{-3}	1.00×10^0	1.53×10^8
to 1.0×10^4					
Unreassociated					
Cot $> 1.0 \times 10^4$	0.0537	-	-	-	-
<u>20 kbp</u>					
Very fast					
Cot $< 1.0 \times 10^{-1}$	0.6363	-	-	-	-
Fast					
Cot 1.0×10^{-1}	0.1404	4.38×10^{-1}	3.12×10^0	1.55×10^2	2.72×10^5
to 1.3×10^1					

.....Contd.

TABLE 5.1 (Contd.)

1	2	3	4	5	6
Slow					
Cot 1.3×10^1 to 1.0×10^4	0.2303	2.82×10^{-3}	1.23×10^{-2}	1.00×10^0	6.90×10^7
Unreassociated					
Cot $> 1.0 \times 10^4$	0.0000	-	-	-	-
<u>Fox tail millet</u>					
<u>1.5 kbp</u>					
Very fast					
Cot $< 1.0 \times 10^{-1}$	0.2787	-	-	-	-
Fast					
Cot 1.0×10^{-1} to 1.5×10^1	0.3056	7.24×10^{-1}	2.37×10^0	1.27×10^2	3.58×10^5
Slow					
Cot 1.5×10^1 to 1.0×10^4	0.4366	5.69×10^{-3}	1.30×10^{-2}	1.00×10^0	6.53×10^7
Unreassociated					
Cot $> 1.0 \times 10^4$	0.0000	-	-	-	-
<u>20 kbp</u>					
Very fast					
Cot $< 1.0 \times 10^{-1}$	0.6365	-	-	-	-
Fast					
Cot 1.0×10^{-1} to 1.0×10^4	0.3609	2.02×10^{-1}	5.60×10^{-1}	1.00×10^0	1.52×10^6

.....Contd.

TABLE 5.1 (Contd.)

1	2	3	4	5	6
Slow					
Unreassociated					
Cot $> 1.0 \times 10^4$	0.0000	-	-	-	-
<u>Little millet</u>					
<u>1.5 kbp</u>					
Very fast					
Cot $< 5.0 \times 10^{-2}$	0.3047	-	-	-	-
Fast					
Cot 5.0×10^{-2} to 1.7×10^2	0.5138	1.56×10^{-1}	3.04×10^{-1}	2.97×10^2	2.79×10^6
Slow					
Cot 1.7×10^2 to 1.0×10^4	0.1798	5.25×10^{-4}	2.92×10^{-3}	1.00×10^0	2.91×10^8
Unreassociated					
Cot $> 1.0 \times 10^4$	0.0057	-	-	-	-
<u>20 kbp</u>					
Very fast					
Cot $< 5.0 \times 10^{-2}$	0.6205	-	-	-	-
Fast					
Cot 5.0×10^{-2} to 1.0×10^2	0.2461	8.5×10^{-1}	3.48×10^0	3.09×10^2	2.44×10^5
Slow					
Cot 1.0×10^2 to 1.0×10^4	0.1156	2.7×10^{-3}	2.40×10^{-2}	1.00×10^0	3.54×10^7
Unreassociated					
Cot $> 1.0 \times 10^4$	0.0280	-	-	-	-

TABLE 5.1 (Contd.)

1	2	3	4	5	6
<u>Great millet</u>					
<u>1.5 kbp</u>					
Very fast					
Cot < 1.0×10^{-2}	0.2260	-	-	-	-
Fast					
Cot 1.0×10^{-2}	0.4408	1.68×10^0	3.80×10^0	3.70×10^2	2.23×10^5
to 2.0×10^1					
Slow					
Cot 2.0×10^1	0.2696	4.53×10^{-3}	1.68×10^{-2}	1.00×10^0	5.06×10^7
to 1.0×10^4					
Unreassociated					
Cot > 1.0×10^4	0.0708	-	-	-	-
<u>20 kbp</u>					
Very fast					
Cot < 1.6×10^{-3}	0.2781	-	-	-	-
Fast					
Cot 1.6×10^{-3}	0.6396	1.40×10^0	2.19×10^0	5.48×10^2	3.88×10^5
to 2.0×10^{-1}					
Slow					
Cot 2.0×10^{-1}	0.0837	2.56×10^{-3}	3.06×10^{-2}	1.00×10^0	2.78×10^7
to 1.0×10^4					
Unreassociated					
Cot > 1.0×10^4	0.0000	-	-	-	-

TABLE 5.1 (Contd.)

^{a,b} Values obtained from the respective reassociation curves (Fig. 5.2 and 5.3)

^c K observed/fraction of genome

^d K observed of given fraction/K observed of slow DNA component

^e Values obtained from the standard relationship for E. coli where $Cot\ 1/2$ is 5.3×10^6 (determined under our experimental conditions) and genome size of 4.5×10^6 nucleotide pairs.

67% at 1.5 kbp and 92% at 20 kbp.

Hyperchromicity of repetitive duplexes of varying lengths

Another approach of assessing the occurrence of DNA sequence interspersion is to study the optical melting behaviour of reassociated duplexes isolated from DNAs of varying lengths. Figure 5.4 displays the melting profiles of the DNAs of different fragment lengths reassociated to their respective limit Cot values, and the results are summarized in Table 5.2. It can be seen from the above analysis that though duplexes of high thermal stability are formed, hyperchromicity decreases with increasing fragment length. The decrease in the hyperchromicity is from 17.7% to 13.3% in barn yard millet, 21.63% to 11.97% in fox tail millet, 18.4% to 10.0% in little millet and 20.1% to 14.4% in great millet. Such a decrease in the hyperchromicity of a repetitive fraction with an increase in DNA fragment size is attributed to the increasing presence of single strand regions in the reassociated fragments. These single strand regions are most likely to be unique sequences which remain unreassociated at Cot 75, Cot 5, Cot 15 and Cot 20 M.s. in case of barn yard millet, fox tail millet, little millet and great millet respectively. These results thus provide an additional evidence for the existence of repetitive and non-repetitive DNA sequences on the same strand in all the four millet genomes.

Figure 5.4 THERMAL DENATURATION PROFILES OF REPETITIVE
DNA DUPLEXES ISOLATED AT THE LIMIT REPETITIVE COT

- (a) Barn yard millet DNA, Cot 75 M.s
- (b) Fox tail millet DNA, Cot 5 M.s
- (c) Little millet DNA, Cot 15 M.s
- (d) Great millet DNA, Cot 20 M.s

where

(O—O)	represents duplexes isolated from 0.4 kbp long DNA
(Δ—Δ)	" duplexes isolated from 1.5 kbp long DNA
(X—X)	" duplexes isolated from 5.6 kbp long DNA
(▲—▲)	" duplexes isolated from 20.0 kbp long DNA

The melting profile of native 0.4 kbp long DNA (●—●) is included for comparison. Each point in these graphs represents an average of at least three experiments.

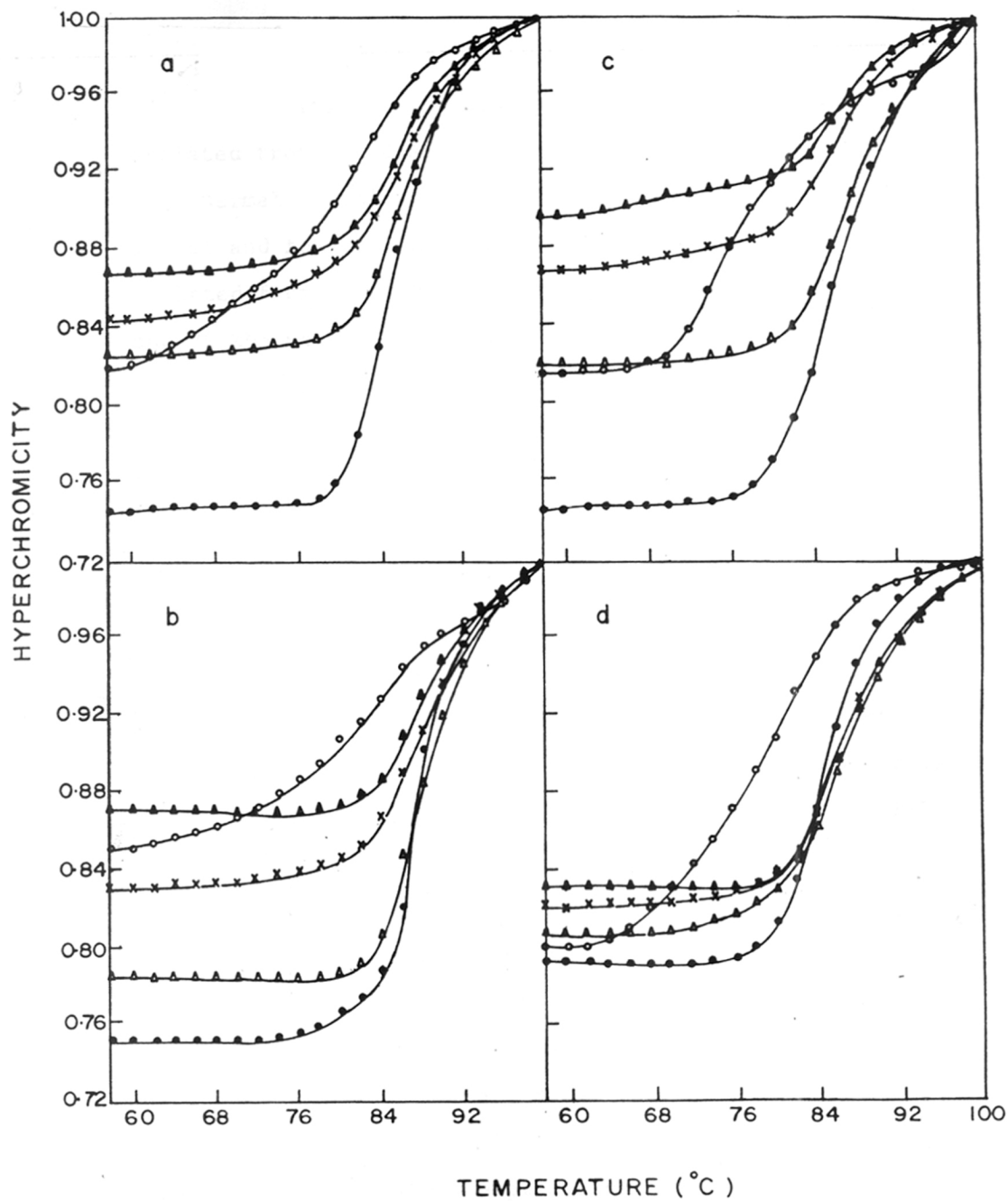


FIGURE 5.4

The melting temperatures of the repetitive duplexes isolated from the DNAs of different fragment lengths show high thermal stability (Table 5.2). In case of barn yard millet and fox tail millet, the T_m of the repetitive duplexes isolated from DNAs of fragment lengths ranging between 1.5 kbp and 20 kbp varies in the narrow range of 86°C - 87°C and 87°C - 88°C respectively and is higher than that of the native DNAs. In case of little millet and great millet these values are 86°C - 86.8°C and 86.2°C - 87°C respectively and are close to that of the native DNAs. Thus the repeat sequences of the four millets show little or no sequence divergence.

The high melting temperatures of the repetitive duplexes cannot be solely attributed to fragment lengths (20,21). The melting temperatures of the repetitive duplexes can be predicted using the formula $T_n - T_m = B/L$ where T_n is the melting temperature of native DNA, L is the fragment length and B is a constant equal to 650 (11). Using this formula, the melting temperatures should be in the range of 84.4 - 86.7°C in the four millets. This range is much narrower than the observed range of 78.9 - 87.2°C in barn yard millet, 80.4 - 88.2°C in fox tail millet, 79.3 - 86.8°C in little millet and 79.5 - 87°C in great millet. A likely explanation for this observation

TABLE 5.2 - MELTING ANALYSIS OF REPETITIVE DUPLICATIONS ISOLATED FROM DNAs OF DIFFERENT FRAGMENT LENGTHS

Species and DNA fragment size (kbp)	Fraction ^a bound to HA at limit Cot	Hyperchro-micity (%)	Tm ^c (°C)	Duplex ^d content	Average duplex length ^e
1	2	3	4	5	6
<u>Barn yard millet Cot 75 M.s</u>					
Native sonicated DNA (0.4 kbp)	-	24.91±0.37	85.97±0.37	-	-
0.4	0.450	14.40±2.00	LMF 78.9±0.80 HMF 96.1±1.90	0.5549	221.96
1.5	0.600	17.72±0.80	87.1±0.50	0.6955	1043.25
5.6	0.805	15.59±0.90	86.3±1.50	0.6053	3389.68
8.2	0.865	14.40±0.80	86.3±1.50	0.5549	4550.18
20.0	0.900	13.33±0.33	87.2±0.33	0.5095	10190.00
<u>Fox tail millet Cot 5 M.s</u>					
Native sonicated DNA (0.4 kbp)	-	25.01±1.90	86.8±0.35	-	-
0.4	0.400	14.30±1.40	LMF 80.4±1.00 HMF 96.0±2.40	0.5485	219.40
1.5	0.500	21.63±0.05	88.2±0.25	0.8578	1286.70
5.6	0.715	17.43±0.31	87.9±0.10	0.6806	3811.36
8.2	0.800	16.68±2.39	87.9±0.20	0.6490	5321.80
20.0	0.850	11.97±2.88	87.4±0.45	0.4444	8888.00

Contd..

TABLE 5.2 (Contd.)

1	2	3	4	5	6
<u>Little millet Cot 15 M.s</u>					
Native sonicated DNA (0.4 kbp)	-	23.00±1.20	86.4±0.20	-	-
0.4	0.450	15.50±2.70	IMF 79.3±1.0 HMF 96.4±1.7	0.6544	261.76
1.5	0.670	18.44±1.31	86.0±0.50	0.7899	1184.85
5.6	0.760	14.39±1.29	86.8±0.43	0.6032	3377.92
8.2	0.795	13.81±1.38	86.7±0.42	0.5765	4727.30
20.0	0.850	10.07±0.50	86.4±0.08	0.4042	8084.00
<u>Great millet Cot 20 M.s</u>					
Native sonicated DNA (0.4 kbp)	-	23.0±0.80	86.9±0.90	-	-
0.4	0.500	19.5±2.63	79.5±0.19	0.8390	335.60
1.5	0.600	20.1±0.15	86.2±0.15	0.8670	1300.50
5.6	0.815	18.6±1.12	86.4±1.12	0.7980	4468.80
8.2	0.865	17.8±1.90	86.9±0.20	0.7610	6240.20
20.0	0.935	14.0±4.60	87.0±0.30	0.5860	11720.00

^aObtained from Figure 5.8^bObtained using the formula as described in the caption to Table 2.1^cObtained from Figure 5.4 or their respective denaturation profiles.

TABLE 5.2 (Contd.)

^dThe average duplex content (D) of bound fragment was estimated using the formula:

$$D = \frac{H - \text{Single strand collapse}}{H_{(\text{native sonicated DNA})} - \text{Single strand collapse}}$$

Single strand collapse was calculated as described by Graham et al., (13)¹⁴ and was 1.3% for barn yard millet, little millet and great millet and 1.5% for fox tail millet.

^eThe average length of duplex regions is the product of the duplex content (D) and the fragment length of DNA.

LMF and HMF stand for low melting fraction and high melting fraction respectively (Chapter IV).

is that much of the repeated duplexes of these millets are present in long arrays consisting of smaller elements varying slightly in their nucleotide sequences. When the DNAs are sheared, there is a greater probability of mismatch in the reassociated duplexes than if the entire stretch were left intact. When the long fragments are reassociated without shearing, they are just zipped together to form stable duplexes with minimum mismatch between the shorter elements.

Hyperchromicity measurements of reassociated long fragments of DNA can be used to estimate their duplex content(8,13,15). Using hyperchromicity values, we have determined that about 51% of barn yard millet, 44.5% of fox tail millet, 40% of little millet and 58.6% of great millet genome is base paired at the highest fragment length (20 kbp). These values are in good agreement with the repetitive DNA content values obtained from the re-association curves of 0.4 kbp long DNAs allowing $\pm 5\%$ experimental scatter in the percentage reassociation at each Cot value, except in case of fox tail millet where this value is higher.

Characterization of S1 nuclease resistant repetitive duplexes

An increase in hydroxyapatite binding and a decrease

in hyperchromicity with an increase in DNA fragment length have revealed the occurrence of interspersion of repeated and single copy DNA sequences in all the four millet genomes. Use of S1 nuclease is another approach to detect the presence of single stranded DNA stretches in repetitive duplexes isolated from the DNAs of different fragment lengths. Moreover, S1 nuclease can be used to isolate repetitive duplexes which can be further characterized with respect to duplex content, T_m , hyperchromicity and size distribution.

(i) Duplex content

Native high molecular weight DNAs of the four millets were reassociated to their respective limit Cot values (Chapter III). S1 nuclease resistant duplexes were isolated by hydroxyapatite column chromatography. The enzyme has been used under such conditions that it can digest only single stranded regions but not individual mismatched base pairs⁽²²⁾. In all the four millets, the fraction bound to hydroxyapatite at the repetitive Cot reduces considerably after S1 nuclease digestion (Table 5.3). Approximately 52% of barn yard millet, 41.3% of fox tail millet, 48% of little millet and 55% of great millet DNAs are S1 resistant. These values of duplex content compare well with those estimated using hyperchromicity measurements (Table 5.4).

TABLE 5.3 - MELTING ANALYSIS OF S1 NUCLEASE RESISTANT DUPLICES

Species	No. of S1 nuclease expts. carried out.	Native unsonicated DNA ^a		S1 resistant T _m ^b (°C)	DNA duplexes		S1 resistant ^d fraction of the genome (%)
		T _m (°C)	Hyperchro- micity (%)		Hyperchro- micity (%)		
<u>Barn yard millet</u>	5	87.48±0.25	23.76±0.97	85.14±0.24	23.88±0.77	0.52±0.002	
<u>Fox tail millet</u>	5	87.82±0.40	23.74±1.38	86.85±0.39	21.96±0.71	0.41±0.005	
<u>Little millet</u>	5	87.28±0.35	25.50±1.30	85.40±0.50	23.22±0.50	0.48±0.040	
<u>Great millet</u>	5	87.80±0.21	24.88±2.54	83.57±1.90	23.60±1.50	0.55±0.008	

^aFrom Chapter II^bFrom Figure 5.5^cCalculated as described in the caption to Table 2.1^dObtained as $\frac{\mu\text{g of S1 resistant repetitive duplexes bound to hydroxyapatite column}}{\mu\text{g of S1 nuclease treated DNA loaded to hydroxyapatite column}}$

(ii) Melting temperature

The melting temperatures of S1 resistant duplexes are in the range of 85.4°C - 86.9°C (Figure 5.5) and are slightly lower than the respective T_m values of the native unsheared DNAs (Table 5.3). The mismatch in repetitive duplexes as calculated from the lowering of T_m varies in the narrow range of 0.97 - 1.88% in case of all the millets, except great millet which shows a mismatch of 4.23% (Table 5.3).

(iii) Hyperchromicity

The hyperchromicity of the duplexes before S1 nuclease digestion is about 10-14%, which increases to almost 23% after S1 nuclease digestion. This substantial increase in hyperchromicity of S1 nuclease resistant duplex is a further evidence for the occurrence of interspersion of single copy and repetitive duplexes in these millet genomes.

Taking into account the nucleotide base mismatching between the T_m s of native DNAs and S1 resistant duplexes, we can calculate the expected hyperchromicities (11) of the S1 nuclease resistant DNA fractions (Table 5.4). For example, native sonicated barn yard millet DNA has a T_m of 85.95°C and hyperchromicity of 24.91%. The S1 nuclease resistant duplexes have a T_m of 85.74% and thus a base mismatch of 0.23%. Due to this small base mismatch, the hyperchromicity of the S1 resistant fraction is expected

Figure 5.5 THERMAL DENATURATION PROFILES OF DNA DUPLEXES BEFORE (▲—▲) AND AFTER (■—■) S1 NUCLEASE TREATMENT.

The DNA duplexes were isolated at the respective limit Cot of each DNA.

- (a) Barn yard millet DNA, Cot 75 M.s
- (b) Fox tail millet DNA, Cot 5 M.s
- (c) Little millet DNA, Cot 15 M.s
- (d) Great millet DNA, Cot 20 M.s.

The melting profile of native 0.4 kbp long DNA (●—●) is included for comparison. The experiments were done as detailed in Materials and Methods. Each point represents an average of minimum three different experiments using three different DNA preparations.

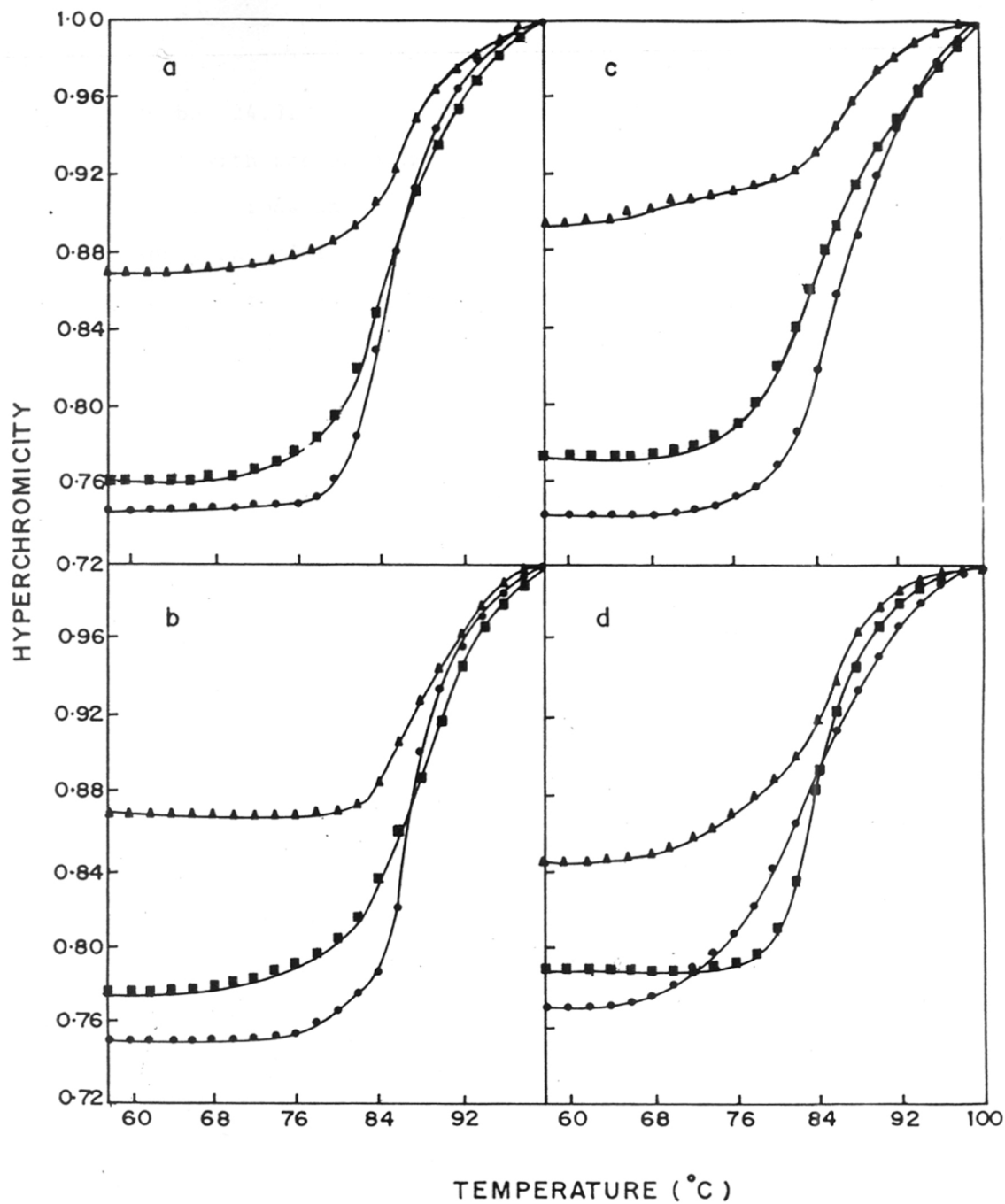


FIGURE 5.5

to be $24.91 \times (1.0 - 0.0021) = 24.86\%$ which compares well with the experimental value of 23.88%. Similar calculations in case of fox tail millet, little millet and great millet show that there is a good agreement between the expected and experimental hyperchromicity values (Table 5.4).

(iv) Size distribution

The duplex length values as obtained from hyperchromicity measurements represent only the average size estimates of the interspersed repeated segments and do not distinguish between distinct size classes. A direct estimate of the size distribution of repetitive DNA sequences is obtained by analyzing the S1 nuclease resistant repetitive duplexes on Agarose A-50 column which has an exclusion limit of 1.5 kbp. Calf thymus DNA, of known fragment lengths below and above the exclusion limit of Agarose A-50, was used as standard. From Figure 5.6, it is clear that there are different sizes of repetitive duplexes in the four millet species. In barn yard millet, for example, approximately 38% of the duplexes have a size > 1.5 kbp and 62% have a heterogenous size distribution in the range of 0.3 - 1.0 kbp. In fox tail millet, majority of the duplexes have a size > 1.5 kbp. Little millet exhibits two distinct size classes: 36% with a size > 1.5 kbp and 64% with a size range of 0.15 - 0.25 kbp.

TABLE 5.4 - EXPERIMENTAL AND EXPECTED HYPERCHROMICITIES FOR S1 NUCLEASE RESISTANT DUPLEXES

Species	Native 0.4 kbp long DNA ^a		S1 resistant DNA duplexes		
	T _m (°C)	Hyperchromicity (%)	T _m ^b (°C)	Hyperchromicity experimental (%) ^c	Hyperchromicity expected (%) ^d
Barn yard millet	85.95±0.37	24.91±0.37	85.74±0.24	23.88±0.77	24.86
Fox tail millet	86.80±0.35	25.01±1.90	86.85±0.39	21.96±0.71	25.02
Little millet	86.40±0.20	23.00±1.20	85.40±0.50	23.32±0.50	23.00
Great millet	86.90±0.90	23.80±1.00	83.57±1.90	23.60±1.50	23.01

^aFrom Chapter IV

^{b,c}Calculated as explained in caption to Table 5.3

^dCalculated using the formula (Graham et al., 13) : Fractions base paired x sonicated, native DNA hyperchromicity. For example native, sonicated barn yard millet DNA (0.4 kbp) has a T_m of 85.95°C and a hyperchromicity of 23.38%. The S1 nuclease Cot 75.0 M.s. fraction has a T_m of 85.74°C and thus a base mismatch of (85.95 - 85.74) = 0.21%. Due to this base mismatch, the hyperchromicity of this DNA fraction is expected to be (24.91 x (1.0 - 0.0021) = 24.86%).

Figure 5.6

SIZING OF S1 NUCLEASE RESISTANT REPETITIVE DUPLICES ISOLATED
AT LIMIT REPETITIVE COT BY AGAROSE A-50 COLUMN CHROMATOGRAPHY

- (a) Calf thymus DNA of known duplex lengths used for the calibration of the column
- (b) Barn yard millet
- (c) Fox tail millet
- (d) Little millet
- (e) Great millet

The column is developed with 0.12M sodium phosphate buffer, pH 6.8, and the elution profiles were obtained by monitoring the effluent with the help of an LKB 2138 UVICORD S at 253 nm.

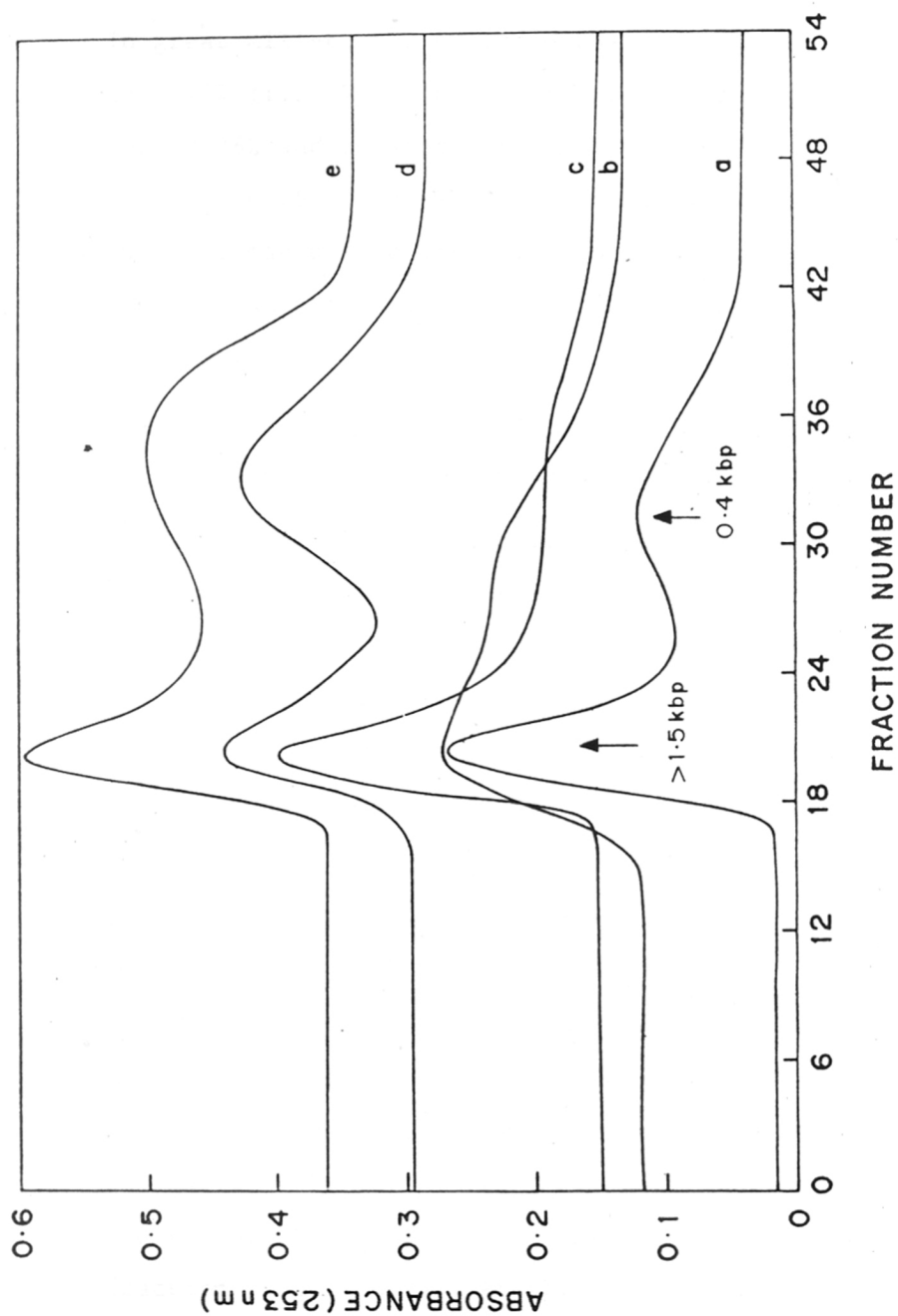


FIGURE 5.6

In great millet, the excluded DNA fraction accounts for 43 - 48% (i.e. > 1.5 kbp). The second fraction accounts for 32-38% and consists of DNA sequences in the size range of 0.15 - 0.30 kbp. The third class of repetitive duplexes has a size distribution in the range of 0.5 - 1.0 kbp and accounts for about 15% of the total S1 resistant duplexes.

(v) Exact sizing of duplexes in the excluded volume of Agarose A-50 column chromatography

The peak fractions obtained in the exclusion volume were analyzed by Agarose slab gel electrophoresis as detailed in Chapter II, for the determination of precise size using appropriate molecular weight markers. In case of all the four millets, the DNA fragments in the exclusion volume showed a heterogeneous size distribution in the range of 2.0 - 9.0 kbp (Figure not included).

(vi) Repetitive sequence classes reassociating at lower Cot values

Experiments were also carried out where the DNAs of high molecular weight (> 20 kbp) were reassociated to Cot 0.1 and then treated with S1 nuclease and fractionated on Agarose A-50 to determine the distribution of long and short repeats in the highly repetitive (fast reassociating) fraction (Figure 5.7). At this Cot value, the proportion of

Figure 5.7 SIZING OF S1 NUCLEASE RESISTANT REPETITIVE DUPLICES ISOLATED
 AT COT 0.1 M.s. BY AGAROSE A-50 COLUMN CHROMATOGRAPHY

(a)	Calf thymus DNA of known duplex lengths used for the calibration of the column
(b)	Barn yard millet
(c)	Fox tail millet
(d)	Little millet
(e)	Great millet

The column is developed with 0.12 M sodium phosphate buffer, pH 6.8 and the elution profiles were obtained by monitoring the effluent with the help of an LKB 2138 UVICORD§ at 253 nm.

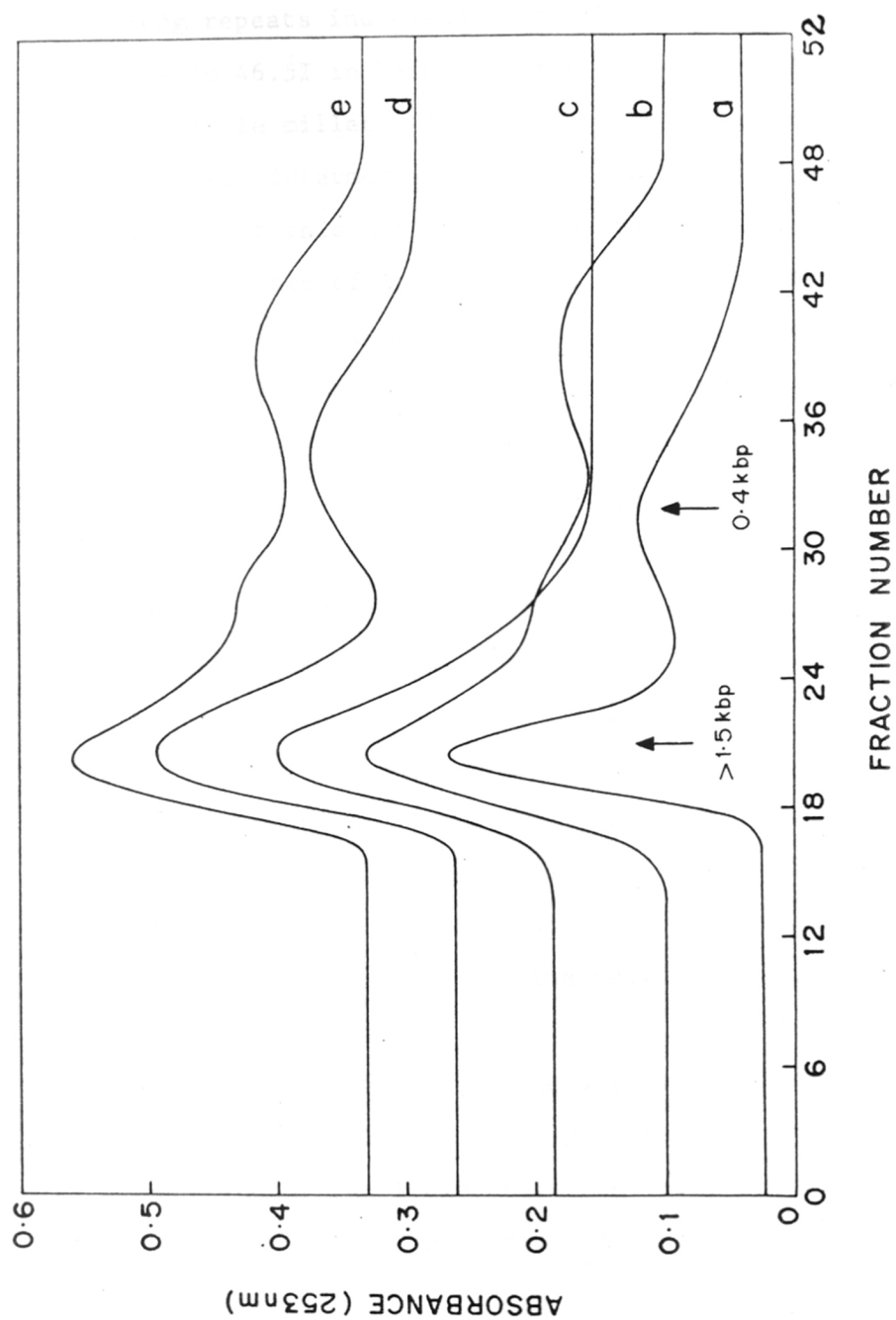


FIGURE 5.7

long repeats increased from 38% of the total repetitive DNA to 46.5% in barn yard millet and from 36% to 57.6% in little millet. In great millet the proportion of the long and intermediate repeats together did not significantly change at this Cot value. Fox tail millet showed only the presence of long repeats. This observation suggested that the fast reassociating DNA fraction in case of barn yard millet and little millet is composed largely of long repeats.

Spacing of repetitive sequence elements.

Interspersion of repetitive and single copy sequences causes a certain amount of single copy DNA to get bound to hydroxyapatite at a Cot value at which only repetitive sequences are fully reassociated. Therefore, at the repetitive Cot, the amount of unreassociated single copy DNA existing as single-stranded tails on reassociated duplexes will increase as a function of DNA fragment length. This will hold true only upto a fragment length which is equal to the average spacing between repetitive sequences. Thus, the estimation of the average length of single copy DNA can be made by plotting a curve relating to the fraction of DNA fragments bound to hydroxyapatite as a function of DNA fragment length. The results of such experiments are depicted in Figure 5.8. Two slopes are

Figure 5.8

THE PERCENTAGE BINDING OF REPETITIVE DNA DUPLICES (AT LIMIT
REPETITIVE COT) TO HYDROXYAPATITE AS A FUNCTION OF DNA FRAGMENT
LENGTH

- (a) Barn yard millet DNA at Cot 75 M.s.
- (b) Fox tail millet DNA at Cot 5 M.s
- (c) Little millet DNA at Cot 15 M.s.
- (d) Great millet DNA at Cot 20 M.s

The DNAs were sheared to different fragment lengths as detailed in the caption to Figure 5.1. The percentage binding at the limit Cot values was determined by hydroxyapatite column chromatography.

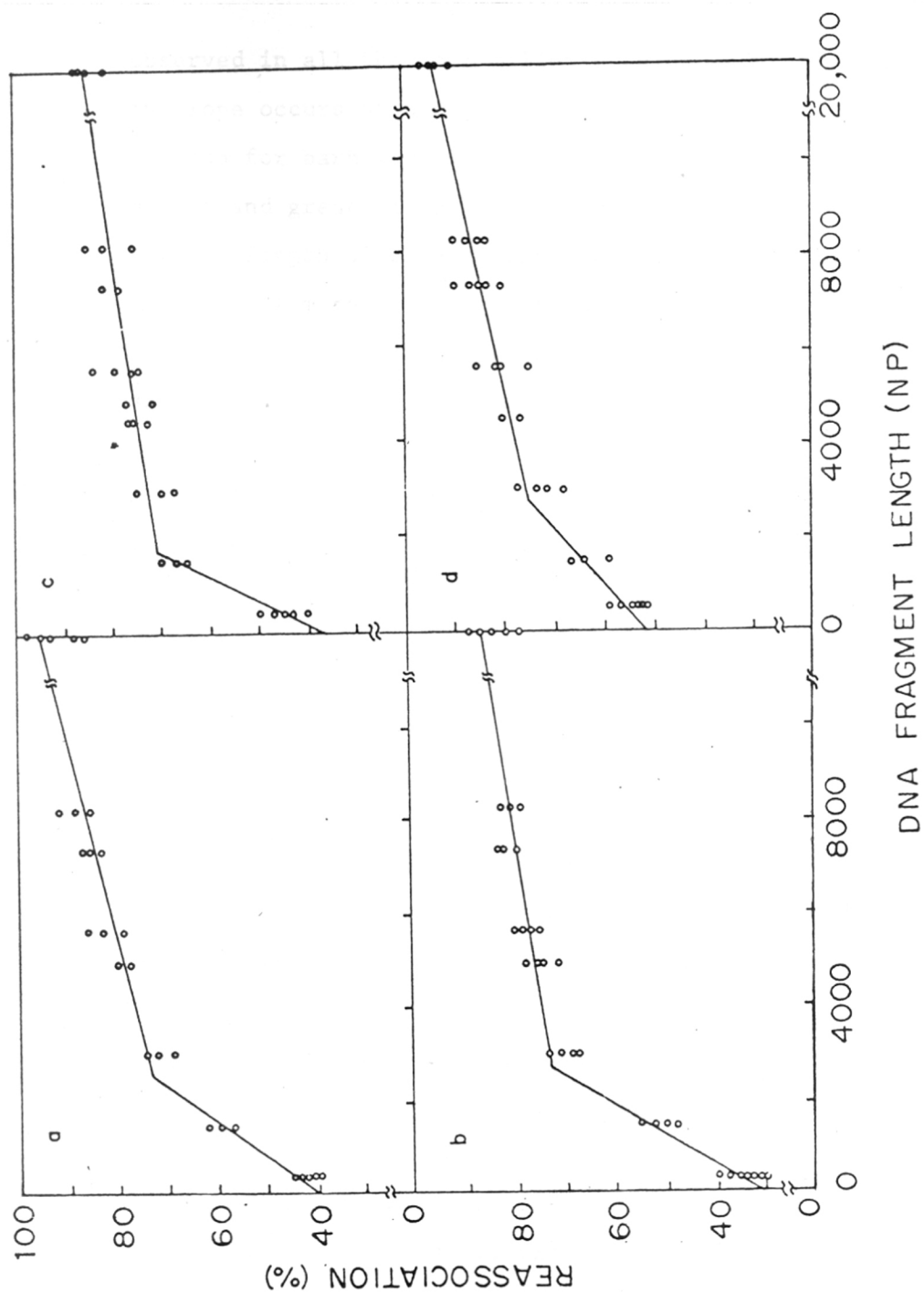


FIGURE 5·8

observed in all the four millet species and the change in slope occurs at a fragment length of 2.6, 2.7, 1.7 and 2.8 kbp for barn yard millet, fox tail millet, little millet and great millet respectively. This is the estimated average length of single copy DNA sequences in these millets. A more gradual increase in slope after these values indicates the presence of a wide spectrum of single copy lengths greater than 2.6, 2.7, 1.7 and 2.8 kbp in barn yard millet, fox tail millet, little millet and great millet respectively. An extrapolation of the line with the greater slope to the ordinate indicates the proportion of total repetitive DNA in each species (Table 5.5).

DISCUSSION

Modes of interspersion patterns

This chapter describes experiments which give a comprehensive information about the DNA sequence organization in the genomes of four related millet species. The reassociation kinetics of long DNA fragments reveal that 85-95% of the millet genomes consist of alternating single copy and repetitive DNA sequences. Barn yard millet, little millet and great millet resemble each other in having the presence of both long and short interspersed repeat DNA sequences in their nuclear genomes. The size of the short repeats, however, is different in the three species

TABLE 5.5 - DUPLEX CONTENT (D) OF REPETITIVE DNA FRACTIONS OF FOUR MILLET SPECIES FROM VARIOUS INDEPENDENT METHODS

Method	<u>Barn yard</u> <u>millet</u>	<u>Fox tail</u> <u>millet</u>	<u>Little</u> <u>millet</u>	<u>Great</u> <u>millet</u>
1. From reassociation kinetics of 0.4 kbp long millet DNAs (Table 3.1)	42.00	30.00*	38.00	52.00
2. Hyperchromicity measurements (Table 5.2)	50.95	44.44	40.42	58.60
3. S1 nuclease resistance measurement (Table 5.3)	52.00	41.33	48.00	55.00
4. From graph of % reassociation vs fragment length (Fig.5.8)	40.00	30.00	38.00	52.00

It can be seen from the above analysis that there is slight variation in the duplex content estimated by different methods. The value obtained from the graph of % reassociation vs fragment length (Fig.5.8) is surely influenced by the reassociation at repeat Cot value of 0.4 kbp long DNAs. So these values match well. However, the value of duplex content obtained from reassociation kinetics of 0.4 kbp long DNAs is lower than that estimated by methods 2 & 3. This is possibly because shearing DNA to an average size of 0.4 kbp results in average fragments which are entirely repetitive or single copy giving rise to estimated duplexes with minimum of single strand tails than has normally been encountered in other genomes. Hence 0.4 kbp duplexes have little or no over estimation.

(0.35 - 1.0 kbp in barn yard millet, 0.15 - 0.25 kbp in little millet and 0.15 - 1.0 kbp in great millet). Fox tail millet genome essentially shows the presence of long repeats. The long repeats in all the four millets have a heterogenous size distribution ranging from 2.0 - 9.0 kbp. The size range of single copy DNA sequences is 1.7 - 2.8 kbp. Thus barn yard millet, little millet and great millet have a mixed sequence interspersation pattern while fox tail millet essentially has a long period interspersation pattern (Table 5.6).

Thermal stability of repetitive duplexes

The melting temperatures of repeated sequences isolated from the DNAs of different fragment lengths (> 1.5 kbp) are close to or even higher than that of native DNAs indicating little or no sequence divergence. This high T_m of repetitive duplexes is not solely attributed to fragment lengths alone because the repeated duplexes of these millets may be present in long arrays consisting of smaller elements varying slightly in their nucleotide sequences. These long repeated sequences maintain internal homogeneity ^{probably} by cross over fixation (23) or ~~may be~~ of recent origin and hence have had little chance to diverge by base substitution.

Dynamics of sequence interspersation patterns

It has been earlier argued by Hinegardner (24) that

TABLE 5.6 - INTERSPERSION PATTERNS IN THE FOUR MILLET GENOMES

Species	% of the genome inter-spersed	Length of interspersed repeats (kbp)		% of interspersed repeats	Length of interspersed single copy (kbp)		% of uninter-spersed single copy	Type of inter-spersion
		long	short		long	short		
<u>Barn yard millet</u>	95	1.5	0.35-1.00	38	62	2.6	5	Mixed
<u>Fox tail millet</u>	85	1.5	-	100	-	2.7	15	Long
<u>Little millet</u>	85	1.5	0.15-0.25	36	64	1.7	15	Mixed
<u>Great millet</u>	94	1.5	0.15-1.00	60	40	2.8	6	Mixed

much of the DNA in eukaryotes is secondary and serves no strongly sequence dependent role in gene regulation. Such sequences are free to evolve by amplification, deletion and base substitution. Much of the repetitive DNA in eukaryotes is considered to originate from periodic amplification events with members of the resulting sequence families being subjected to translocation, base substitution and deletion. When both amplification and deletion occur during evolution, the portion of the genome subjected to these processes can be viewed as 'turning over' over geological time scales. The genome size remains constant under 'steady state' conditions and the rates of amplification and deletion are equal for any one species although different states may exist for different species.

In the family Gramineae, cereals with high DNA content ($1C > 5$ pg) such as wheat, rye and maize have been extensively characterized with respect to their genome organization patterns. In these species, repeated DNA sequences constitute a major portion of the total DNA. These sequences are predominantly short (0.35 - 1.0 kbp) and exhibit extensive mismatching (25-28) and are finely dispersed with single copy sequences. According to Flavell et al. (29), the genomes of wheat, oat, rye and barley are 'turning over' rapidly. Short DNA segments are translocated to new chromosomal sites to

produce new sequence combinations, some of which are further propagated by secondary amplifications. These processes lead to a dispersion of closely related sequences among unrelated sequences giving rise to the observed short period interspersed pattern. It also appears that the repeat sequences in these species have evolved much earlier and have accumulated sufficient sequence divergence. All the millets in the present study show the presence of long repeats in their genomes. Short repeats are also present in case of barn yard millet, little millet and great millet. Furthermore, the thermal stability of long and short repeats together is very similar or even higher than that of native DNA indicating very little sequence divergence. It, therefore, appears that these genomes are slowly turning over and the repeat sequences are of relatively recent origin and hence not diverged from each other by base substitutions.

It can be further assumed that the short repeats have evolved recently from the long repeats which are conserved in these millets. The low content of repetitive DNA (30 - 52%) can be explained by assuming that only a smaller fraction of the millet genome is evolving by rapid amplification and deletion.

REFERENCES

- 1 Britten, R.J. and Kohne, D.E. (1968). Science, 161: 529-540.
- 2 Thompson, W.F. and Murray, M.G. (1981). In: The Biochemistry of Plants. A comprehensive Treatise Proteins and Nucleic acids. Marcus, A. ed. Vol. 6, pp.1-81, Academic Press, New York.
- 3 John, B. and Miklos, G.L.G. (1979). In: International Review of Cytology. Bourne, G.M. and Danielli, J.F. eds., Vol. 58, pp. 1-114, Academic Press, New York.
- 4 Britten, R.J. and Davidson, E.H. (1969). Science, 165: 349-357.
- 5 Davidson, E.H. and Britten, R.J. (1973). Quart. Rev. Biol., 48: 565-613.
- 6 Davidson, E.H. and Britten, R.J. (1979). Science, 204: 1052-1059.
- 7 Bouchard, R.A. (1982). In: International Review of Cytology, Bourne, G.H. and Danielli, J.F. eds., Vol. 76, pp.113-193, Academic Press, New York.
- 8 Davidson, E.H., Hough, B.R., Amenson, C.S. and Britten, R.J. (1973). J. Mol. Biol., 77: 1-23.
- 9 Manning, J.E., Schmid, C.W. and Davidson, N. (1975). Cell, 4: 141-155.
- 10 Mandel, M. and Marmur, J. (1968). In: Methods in Enzymology, Grossman, G.L. and Moldave, K. eds. Vol. 12B, pp.195-206, Academic Press, New York.
- 11 Britten, R.J., Graham, D.E., Neufeld, B.R. (1974). In: Methods in Enzymology, Grossman, G.L. and Moldave, K. eds., Vol. 29E, pp.363-418, Academic Press, New York.
- 12 Zimmerman, J.L., Goldberg, R.B. (1977). Chromosoma (Berl.), 59: 227-252.
- 13 Graham, D.E., Neufeld, B.R., Davidson, E.H. and Britten, R.J. (1974). Cell, 1: 127-137.
- 14 Murray, M.G., Cuellar, R.E. and Thompson, W.F. (1978). Biochemistry, 17: 5781-5790.

- 15 Goldberg, R.B., Crain, W.R., Ruderman, J.V., Moore, G.P., Barnett, T.R., Higgins, R.C., Gelfand, R.A., Galau, G.A., Britten, R.J. and Davidson, E.H. (1975). *Chromosoma* (Berl.), 51, 225-251.
- 16 Britten, R.J., Graham, D.E., Eden, F.C., Painchaud, D.M. and Davidson, E.H. (1976). *J. Mol. Evol.*, 9: 1-23.
- 17 Sachs, D.H. and Painter, E. (1972). *Science*, 165: 781-782.
- 18 Wimpee, C.F. and Rawson, J.R.Y. (1979). *Biochim. Biophys. Acta*, 562:192-206.
- 19 Wetmur, J.G. and Davidson, N. (1968). *J. Mol. Biol.*, 31: 349-370.
- 20 Crothers, D.M., Kallenbach, N.R. and Zimm, B.N. (1965). *J. Mol. Biol.*, 11: 802-820.
- 21 Hayes, F.N., Lilly, E.H., Ratliff, R.L., Smith, D.A. and Williams, D.L. (1970). *Biopolymers*, 9: 1105-1117.
- 22 Vogt, V.M. (1973). *Eur. J. Biochem.*, 33: 192-200.
- 23 Smith, G.P. (1976). *Science*, 191: 528-535.
- 24 Hinegardner, R. (1976). In: *Molecular Evolution*, Ayala, F.J. ed., pp. 179-199. Sinauer, Sunderland, MA.
- 25 Smith, D.B. and Flavell, R.B. (1975). *Chromosoma* (Berl.), 50: 223-242.
- 26 Flavell, R.B. and Smith, D.B. (1976). *Heredity*, 37: 231-252.
- 27 Ranjekar, P.K., Lafontaine, J.G. and Pallotta, D. (1974). *Chromosoma* (Berl.), 48: 427-440.
- 28 Ranjekar, P.K., Pallotta, D. and Lafontaine, J.G. (1976). *Biochim. Biophys. Acta*, 425: 30-40.
- 29 Flavell, R.B., O'Dell, M. and Hutchinson, J. (1981). *Cold Spring Harbor Symposium Quant. Biol.*, Vol. XLV, pp.501-508.

CHAPTER VI

SEQUENCE HOMOLOGIES AMONG THE REPEATED
DNA SEQUENCES OF MILLET GENOMES

SUMMARY

The technique of in vitro DNA-DNA hybridization was employed to determine the sequence homologies among repeated DNA sequences in five millet plant species. Using total pearl millet DNA as tracer, the homology between the repeated DNA sequences of pearl millet and those of other millets is 8-14%. When repetitive DNA of barn yard millet was used as tracer, its homology with those of other repeated DNA sequences increased to 34-37% except in fox tail millet where it is only 12%. The melting profiles of majority of heterologous duplexes are distinctly biphasic. A significant proportion of radioactive counts is eluted at high temperatures indicating that some GC rich sequences are conserved in majority of the millets. Earlier reports of low homology among other Gramineae plant species and the present report of low homology among repeated DNA sequences in millets appear to support the polyphyletic origin of grasses.

INTRODUCTION

Several approaches have been made to understand the origin, accumulation and function of repeated DNA elements in related organisms. One type of analysis focuses on determination of size, complexity and arrangement of repeated DNA components, while more direct comparisons employ reassociation of radiolabelled DNA probes from one species with its own and with DNAs of related species to assess the proportion of shared DNA sequences and their degree of similarity. The thermal dissociation studies of DNA heteroduplexes can reveal the extent to which the base sequences of each species have diverged from one another. Such studies can yield quantitative data on DNA relatedness and DNA sequence divergence among organisms and are, therefore, useful in systematics and speciation events.

Since eukaryotic genomes contain both repeated and single copy DNA sequences, we can expect to find differences in the nucleotide sequences of coding regions and in the sequence organization and frequency of reiteration of moderately repetitive DNA and/or highly repeated DNA sequences during the course of evolution (1). Since single copy sequences are known to have evolved essentially by base substitution, interspecies comparison between single

copy sequences provides a measure of evolutionary sequence divergence (1). Such changes normally occur less frequently than other genomic changes and affect only a single gene. Moreover, they can be detected by amino acid replacement, isozyme analysis, immunological tests on proteins as well as direct DNA sequencing. Repetitive DNA sequences, on the other hand, have evolved as a result of several molecular events like amplification, deletion, translocation and base substitution (2,3). These changes can have pleiotropic effects on genomes (1), and can be detected at intra as well as inter-species level. Interspecies homology studies of repeated sequences in closely related phyletic lineages can thus help to study not only their origin but also a possible way in which they have accumulated in the genome and the probable functional role they play in the genome (4).

DNA-DNA hybridization studies in plants have been confined to species belonging to Gramineae (5-12), Leguminosae (5,13,14), genus Cucurbita (15), Atriplex (16-18) and Osmunda (19,20). These data have yielded information with respect to phylogenetic relatedness, identification of genome donors during evolution and mechanism of species evolution.

In this Chapter, we detail DNA-DNA hybridization experiments to determine the homology of repeated DNA sequences in five millet plant species, namely, pearl millet, barn yard millet, fox tail millet, little millet and great millet. Two approaches have been used to carry out these studies. In the first approach, total pearl millet DNA was used as the radiolabelled probe. Pearl millet is closely related to the three millets, namely, barn yard millet, fox tail millet and little millet. All these millets belong to the tribe Paniceae and show a high melting component in their highly repetitive DNA sequences. Moreover, pearl millet DNA has been used extensively in our laboratory for cloning studies and hence the labelled DNA was continuously available.

In the second approach, isolated repeated DNA fraction of barn yard millet was used as the radiolabelled probe. Barn yard millet was selected because the homologous DNA hybridization experiments involved incubations to a Cot value of 75 M.s. which is the highest repeat Cot value as compared to the range of Cot 5-20 M.s. for the remaining millet species in the present study.

MATERIALS AND METHODS

Seed material

Pearl millet seeds (Pennisetum americanum, variety

Awsari) were purchased locally. Barley seeds (Hordeum vulgare) were obtained from Mahatma Phule Agricultural University, Rahuri, Ahmednagar district, Maharashtra State. Other seeds were procured as indicated in Chapter II.

Extraction, shearing and sizing of total unlabelled DNAs

DNAs from different millet species and from wheat and barley were isolated according to the procedure outlined in Chapter II. The DNA samples were dialyzed against 0.12 M sodium phosphate buffer (pH 6.8) and sheared to an average size of 0.4 kbp as described in Chapter III. A minimum of three DNA preparations were used for the hybridization experiments for each species.

Isolation of repetitive DNA of barn yard millet

4-5 mg of barn yard millet DNA, 0.4 kbp long, in 0.12 M sodium phosphate buffer (pH 6.8) was heat-denatured and incubated to its limit repetitive Cot value of 75 M.s. The reassociated duplexes were separated from the single stranded DNA by hydroxyapatite column chromatography (Chapter IV). These duplexes were used for nick-translation.

In vitro radio labelling of DNA

2-3 μ g of native unsheared pearl millet DNA (molecular weight > 20 kbp) in 10 mM Tris-HCl (pH 7.5) and repetitive DNA of barn yard millet DNA in 400 mM sodium phosphate buffer (pH 6.8) was labelled by nick-translation according

to the procedure of Rigby et al. (21). In case of pearl millet, the DNA was first nicked with pancreatic DNase I (1×10^{-4} units/ μ g) prior to labelling. In case of repetitive DNA fraction of barn yard millet the reaction mixture was diluted in such a way that the effective concentration of phosphate did not exceed 0.05 M in the final volume. The reaction mixture consisted of 50 mM Tris-HCl (pH 8.0), 5 mM MgCl_2 , 20 μ M each of cold deoxynucleotide triphosphates i.e. dATP, dCTP, dGTP and $\alpha^{32}\text{P}$ labelled dTTP (specific activity 3000 Ci/mM). E. coli DNA polymerase I was added to a final concentration of 2-3 units/ μ g of DNA and the incubation was done at 15°C for 90 min. which was found to be adequate to obtain the required specific activity. After the completion of the reaction, sodium dodecyl sulfate was added to give an effective concentration of 0.1% in the final mixture. The mixture was deproteinized twice with freshly distilled phenol saturated with Tris-HCl (10 mM, pH 7.5) and finally with chloroform/isoamylalcohol (24:1 v/v). The aqueous layer was adjusted to 0.12 M sodium phosphate buffer and loaded on a hydroxyapatite column previously equilibrated with the same buffer. Unincorporated dNTPs and single stranded material were removed by washing the column thoroughly with 0.12 M sodium phosphate buffer (pH 6.8). Labelled DNA duplexes were eluted with 0.4 M

sodium phosphate buffer (pH 6.8).

Radioactivity of the DNA solution was measured by Cerenkov counting method (22) using a Beckman Liquid Scintillation counter, Model LS-100C and taking a constant volume of solution (3 ml).

Purified nick-translated pearl millet DNA was then sonically sheared to an average fragment length of 0.4 kbp using Biosonic III (Chapter III).

Sizing of labelled DNA

Nick-translated pearl millet DNA and repetitive fraction of barn yard millet DNA were sized by agarose (1% w/v) slab gel electrophoresis. Unsheared and sheared (0.4 kbp) cold DNAs were also run simultaneously along with labelled DNAs with appropriate DNA molecular weight markers. Cold DNAs and DNA molecular weight marker bands were visualized using the ultraviolet transilluminator. The labelled DNA was eluted from the gel cut into 1 cm pieces using the procedure of Maniatis et al. (23) and the radioactivity of the eluted samples was measured. Molecular weight of the labelled DNAs was then calculated with respect to the mobility of the molecular weight markers.

Reassociation studies

When a mixture containing separated DNA strands of two organisms is allowed to reanneal under proper conditions,

the DNA strands will reassociate to the extent of their homology and give rise to heteroduplexes or hybrid DNA molecules. The quantity of hybrid DNA is easily measurable when one of the two strands is radiolabelled. The melting temperature difference between the homologous and heterologous hybrids is a measure of base sequence divergence between the two DNAs under study.

In experiments involving duplex formation between labelled and unlabelled DNA fragments, the latter was taken in approx. 10,000 fold excess to the former. Approximately 50,000 counts of tracer DNA (1×10^{-2} μ g) and 75-100 μ g of driver DNA were used for each hybridization experiment. Incubations were carried out in 0.36 M sodium phosphate buffer, pH 6.8 and the corresponding corrections for reassociation rate were applied. Samples were denatured in sealed ampoules in a boiling water bath for 10 min. and incubated to the repeat Cot values of the driver DNA. They were then diluted to 0.12 M sodium phosphate buffer with double distilled water prior to loading on hydroxyapatite column equilibrated with 0.12 M sodium phosphate buffer pH 6.8, at 62°C. DNA samples eluting with 0.12 M sodium phosphate buffer pH 6.8 at 62°C were considered to be single stranded and those eluting with 0.4 M sodium phosphate buffer (pH 6.8)

at and above 62°C as double stranded. Recovery of the total counts from the column was 100-105%. The fraction reassociated was computed using the formula:

$$\% R = \frac{\text{Counts of fraction eluted with 0.4 M sodium phosphate buffer (pH 6.8) at 62°C and 100°C}}{\text{Total counts eluted from column with 0.12 M and 0.4 M sodium phosphate buffer (pH 6.8) at 62°C and 100°C}}$$

All heterologous hybridization experiments were done with a corresponding homologous experiment as a control to determine the self-reassociation of tracer DNA.

Hydroxyapatite thermal denaturation analysis

The duplex DNA bound to hydroxyapatite (after elution of single strand DNA with 0.12 M sodium phosphate buffer, pH 6.8 at 62°C) was recovered stepwise by raising the temperature of the columns in 5°C increments upto 100°C. At each temperature, the denatured DNA duplexes were eluted with one column volume of 0.12 M sodium phosphate buffer (pH 6.8). Duplexes eluting at successive temperatures after 62°C in 0.12 M sodium phosphate buffer and finally at 100°C in 0.4 M sodium phosphate buffer (pH 6.8) were considered as 100% of amount bound and the temperature at which 50% of the duplexes eluted from the column was taken as the melting temperature of the duplexes.

To get an idea about the fraction of the total counts eluted at each successive temperature rise, a graph of counts eluted between successive temperature rise versus temperature was plotted. Using these methodologies of reassociation and thermal denaturation, the following hybridization experiments were carried out.

1. Radiolabelled total wheat DNA was hybridized to barley DNA (0.4 kbp long).
2. Radiolabelled total pearl millet DNA was hybridized to cold DNAs of barn yard millet, fox tail millet, little millet and great millet (0.4 kbp long).
3. Radiolabelled repetitive DNA of barn yard millet was hybridized to cold DNAs of fox tail millet, little millet, great millet and pearl millet (0.4 kbp long).

All incubations were carried out to the limit Cot values of the driver DNA.

RESULTS

Self-reassociation of tracer DNA

In all the hybridization experiments, unlabelled DNA (driver DNA) was used in 10,000 fold excess to the labelled DNA (tracer DNA) to ensure that all the labelled DNA was involved in hybridization with the unlabelled DNA.

However, the tracer DNA may still undergo self-reassociation, which was the case in the present experiments and the extent of self-reassociation of tracer DNA was found to be in the range of 12-22% depending on the incubation time. Smith and Flavell (24) have reported a self-reassociation value of 18.7% in wheat DNA. All the homologous and heterologous hybridization values have, therefore, been corrected for self-reassociation of tracer DNA incubated under comparable conditions.

Control experiments

The homology of repeated DNA sequences between wheat and barley was determined under our experimental conditions according to Smith and Flavell (24) and Flavell *et al.* (25). Total labelled wheat DNA of an average fragment length of 0.4 kbp was hybridized with excess cold barley DNA of the same size to a Cot value of 50 M.s. (24) and 90 M.s. (25) to allow all the repeated sequences to form duplexes. The homologous reassociation value (wheat x wheat) was 76.67% ($\pm 3.45\%$) whereas heterologous reassociation values (wheat x barley) was 63.62% ($\pm 3.58\%$). After correction for self-reassociation of the tracer DNA ($24.97 \pm 2.95\%$) and normalizing the homologous hybridization value to 100%, the percentage homology between the repeated sequences of wheat and barley

Figure 6.1

HYDROXYAPATITE THERMAL DENATURATION PROFILES OF HOMOLOGOUS (O—O) AND HETEROLOGOUS (●—●) DNA DUPLEXES.

Homologous DNA duplexes were obtained by reannealing wheat tracer DNA (0.4 kbp long) with excess of wheat driver DNA (0.4 kbp long). Heterologous DNA duplexes were obtained by reannealing wheat tracer DNA (0.4 kbp long) with excess of barley driver DNA (0.4 kbp long).

(a) DNA duplexes obtained at Cot 50 M.s.

(b) DNA duplexes obtained at Cot 90 M.s.

Graphs are plotted as cumulative percentage counts eluted at each temperature interval of 5°C.

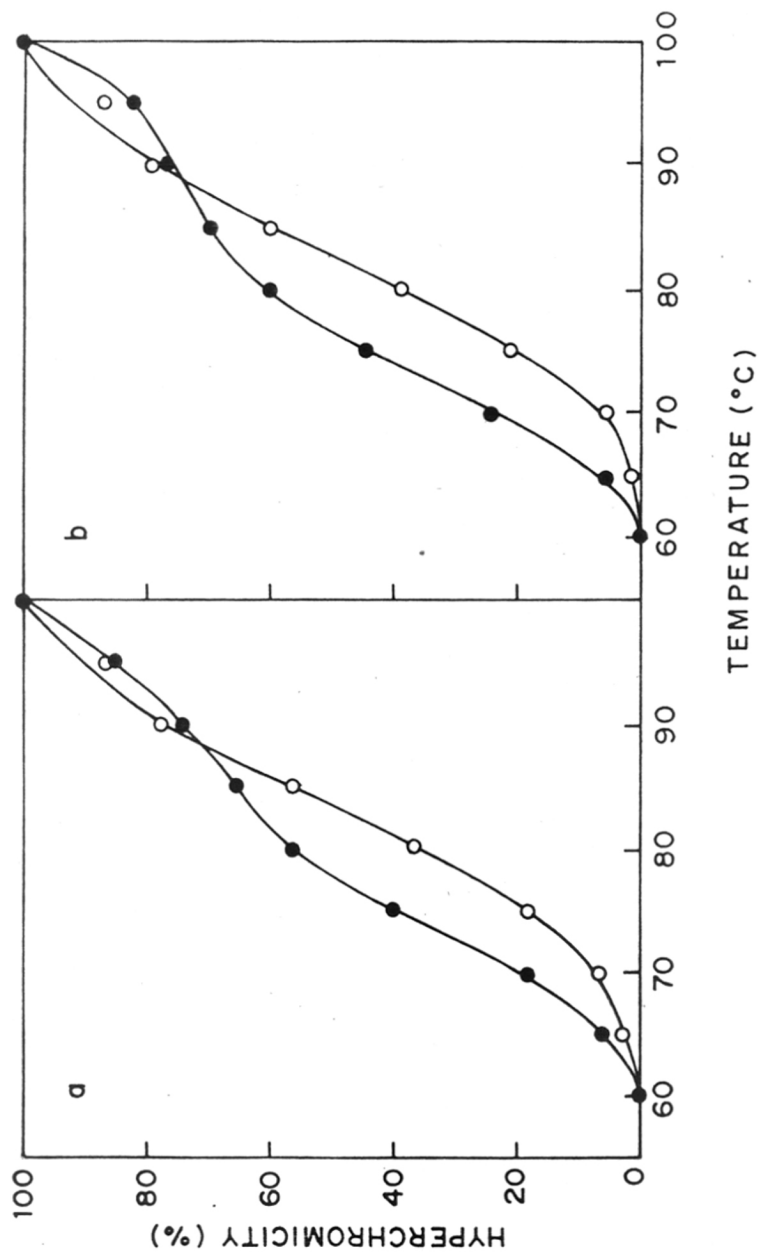


FIGURE 6.1

Figure 6.2

DERIVATIVE THERMAL ELUTION PROFILES OF HOMOLOGOUS (O—O) AND HETEROLOGOUS (X—X) DNA DUPLEXES.

Homologous DNA duplexes were obtained by reannealing wheat tracer DNA (0.4 kbp long) with excess of wheat driver DNA (0.4 kbp long). Heterologous DNA duplexes were obtained by reannealing wheat tracer DNA (0.4 kbp long) with excess of barley driver DNA (0.4 kbp long).

(a) DNA duplexes obtained at Cot 50 M.s.

(b) DNA duplexes obtained at Cot 90 M.s.

Denaturation results are graphed as the percentage of the counts eluted at each temperature rise of 5°C.

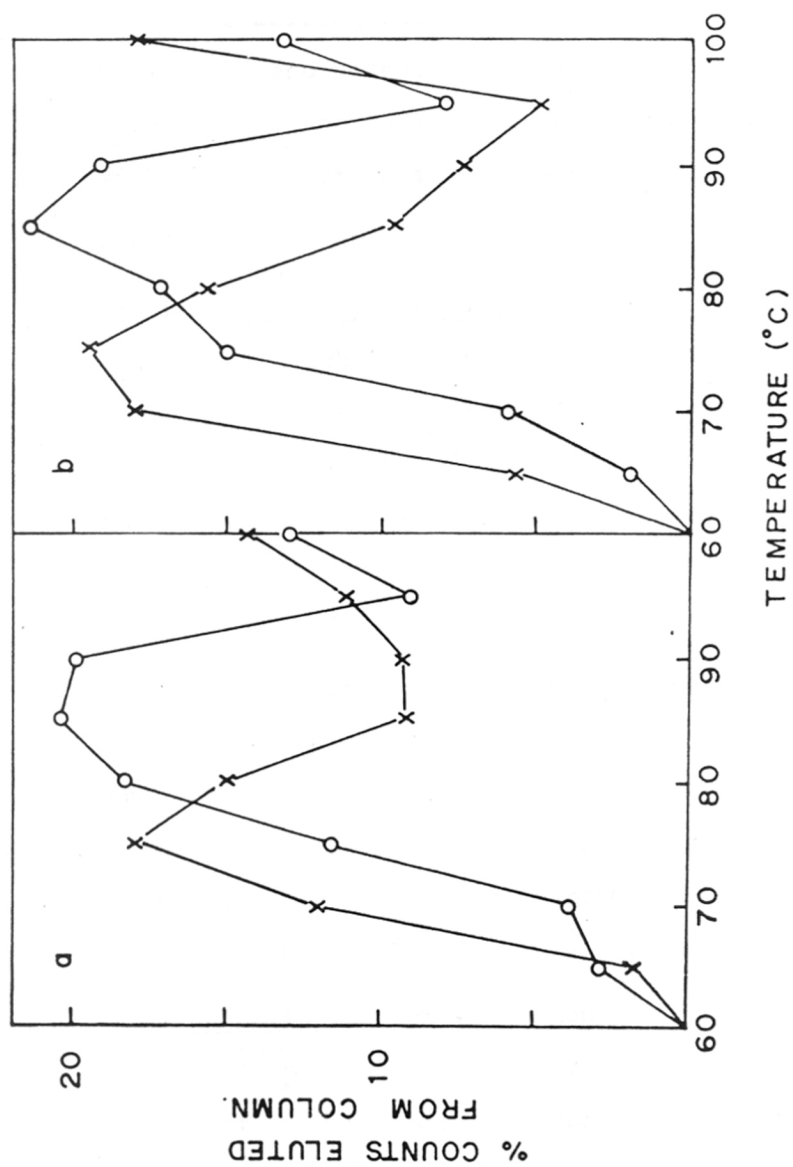


FIGURE 6.2

was estimated to be 74.75%.

The melting and derivative profiles of the homologous and heterologous duplexes were also determined (Figures 6.1 and 6.2). All the results were comparable with those reported in literature (24,25).

Use of total nick-translated pearl millet DNA as tracer

(i) Homologous hybridization

$\alpha^{32}\text{P}$ labelled pearl millet DNA was hybridized in excess of its own cold DNA to Cot values of 5 M.s., 15 M.s., 20 M.s. and 75 M.s. (Table 6.1). The reassociation values, after correction for self-reassociation, are 22.88% at Cot 5 M.s., 23.79% at Cot 15 M.s., 23.79% at Cot 20 M.s. and 32.13% at Cot 75 M.s.

The melting profiles of the homologous duplexes at all the Cot values were biphasic with a major low melting fraction (T_m : 77.5°C - 79.0°C) and a minor high melting fraction (T_m : 94.6°C - 96.6°C). The native 0.4 kbp long pearl millet DNA has a T_m of 87°C (26). The homologous duplexes, therefore, exhibit a base mismatch of 8 - 9.5%.

(ii) Heterologous hybridization

$\alpha^{32}\text{P}$ labelled pearl millet DNA was hybridized with excess of cold 0.4 kbp long DNAs of barn yard millet (Cot 75 M.s.), fox tail millet (Cot 5 M.s.), little millet

TABLE 6.1 - HOMOLOGY AMONG REPEATED DNA SEQUENCES USING RADIOLABELLED PEARL MILLET DNA

Driver DNA	No. of expt.	Cot value (M.s.)	Reassociation ^a (%)	Self reassocia- tion of tracer (%)	Reassocia- ^b after correction ^c (%)	Homology ^d
<u>Pearl millet</u>	3		46.11±2.81		32.13	100.00
<u>Barn yard millet</u>	3	75	17.71±0.17	13.98±2.28	3.73	10.20
<u>Pearl millet</u>	3	5	35.07±3.23	12.19±1.96	22.88	100.00
<u>Fox tail millet</u>	4		14.11±3.51		1.92	8.40
<u>Pearl millet</u>	3	15	37.04±4.88	13.25±2.20	23.79	100.00
<u>Little millet</u>	2		16.55±2.00		3.30	13.87
<u>Pearl millet</u>	3	20	37.04±4.88	13.25±2.20	23.79	100.00
<u>Great millet</u>	3		15.67±1.98		2.42	10.17

^aReassociation between homologous and heterologous hybrids at the limit Cot value of driver DNA.

^bObtained by allowing tracer DNA to reassociate alone for the corresponding Cot value used in homologous and heterologous reassociations.

^cObtained as $c = a - b$

^dCalculated by normalizing the corresponding homologous experiment to 100% homology.

Figure 6.3 HYDROXYAPATITE THERMAL DENATURATION PROFILES
OF HOMOLOGOUS AND HETEROLOGOUS DNA DUPLEXES.

- (a) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 75 M.s.
(●—●) Pearl millet tracer DNA x barn yard millet driver DNA at Cot 75 M.s.
- (b) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 5 M.s.
(●—●) Pearl millet tracer DNA x fox tail millet driver DNA at Cot 5 M.s.
- (c) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 15 M.s.
(●—●) Pearl millet tracer DNA x little millet driver DNA at 15 M.s.
- (d) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 20 M.s.
(●—●) Pearl millet tracer DNA x great millet driver DNA at Cot 20 M.s.

Renaturation was carried to the respective limit Cot values of the driver DNA. Graphs were plotted as cumulative percentage counts eluted at each temperature interval.

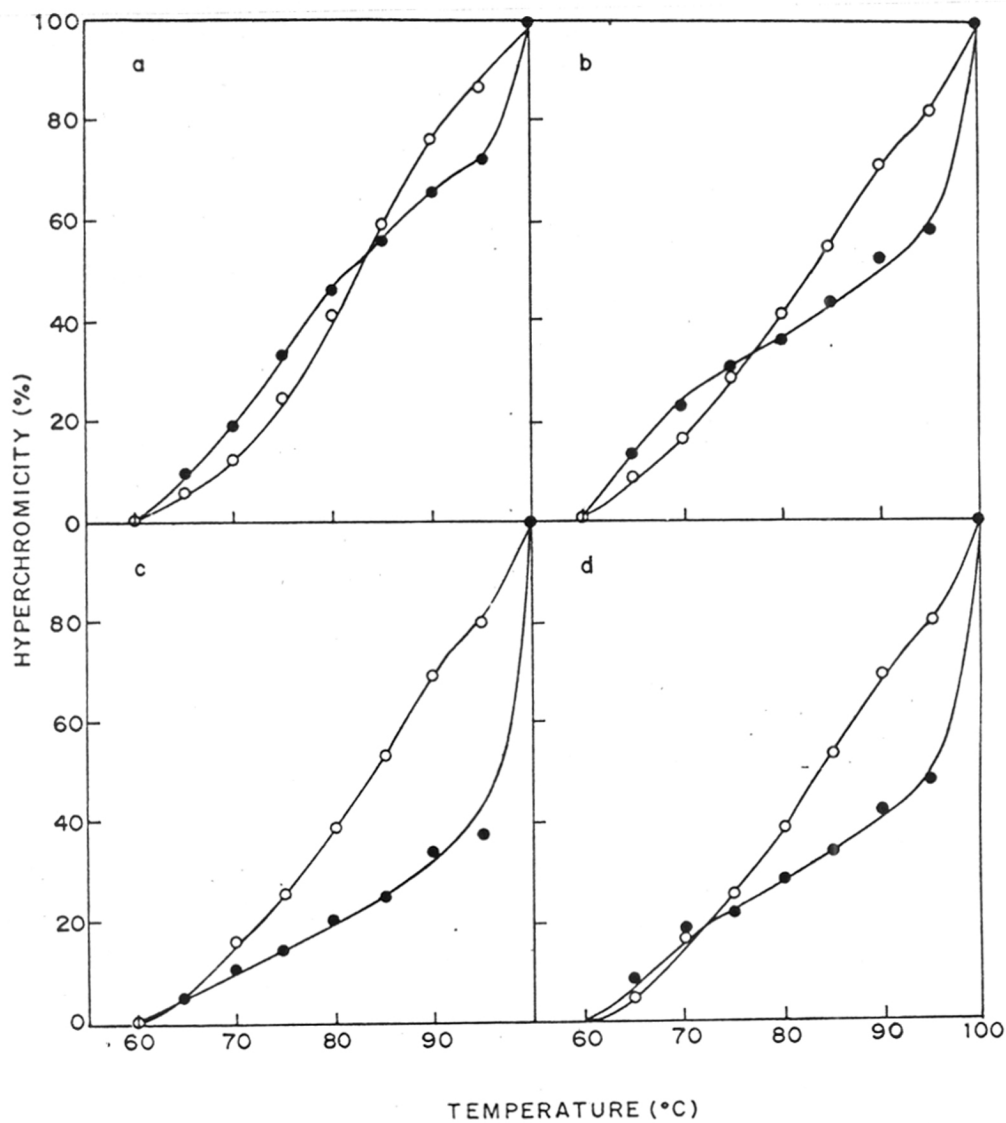


FIGURE 6.3

Figure 6.4 DERIVATIVE THERMAL DENATURATION PROFILES
OF HOMOLOGOUS AND HETEROLOGOUS DNA DUPLEXES.

- (a) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 75 M.s.
(X—X) Pearl millet tracer DNA x barn yard millet driver DNA at Cot 75 M.s.
- (b) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 5 M.s.
(X—X) Pearl millet tracer DNA x fox tail millet driver DNA at Cot 5 M.s.
- (c) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 15 M.s.
(X—X) Pearl millet tracer DNA x little millet driver DNA at Cot 15 M.s.
- (d) (O—O) Pearl millet tracer DNA x pearl millet driver DNA at Cot 20 M.s.
(X—X) Pearl millet tracer DNA x great millet driver DNA at Cot 20 M.s.

Renaturation was carried out to the respective limit Cot values of the driver DNA. Denaturation results are graphed as the percentage of the counts eluted at each temperature rise of 5°C.

•

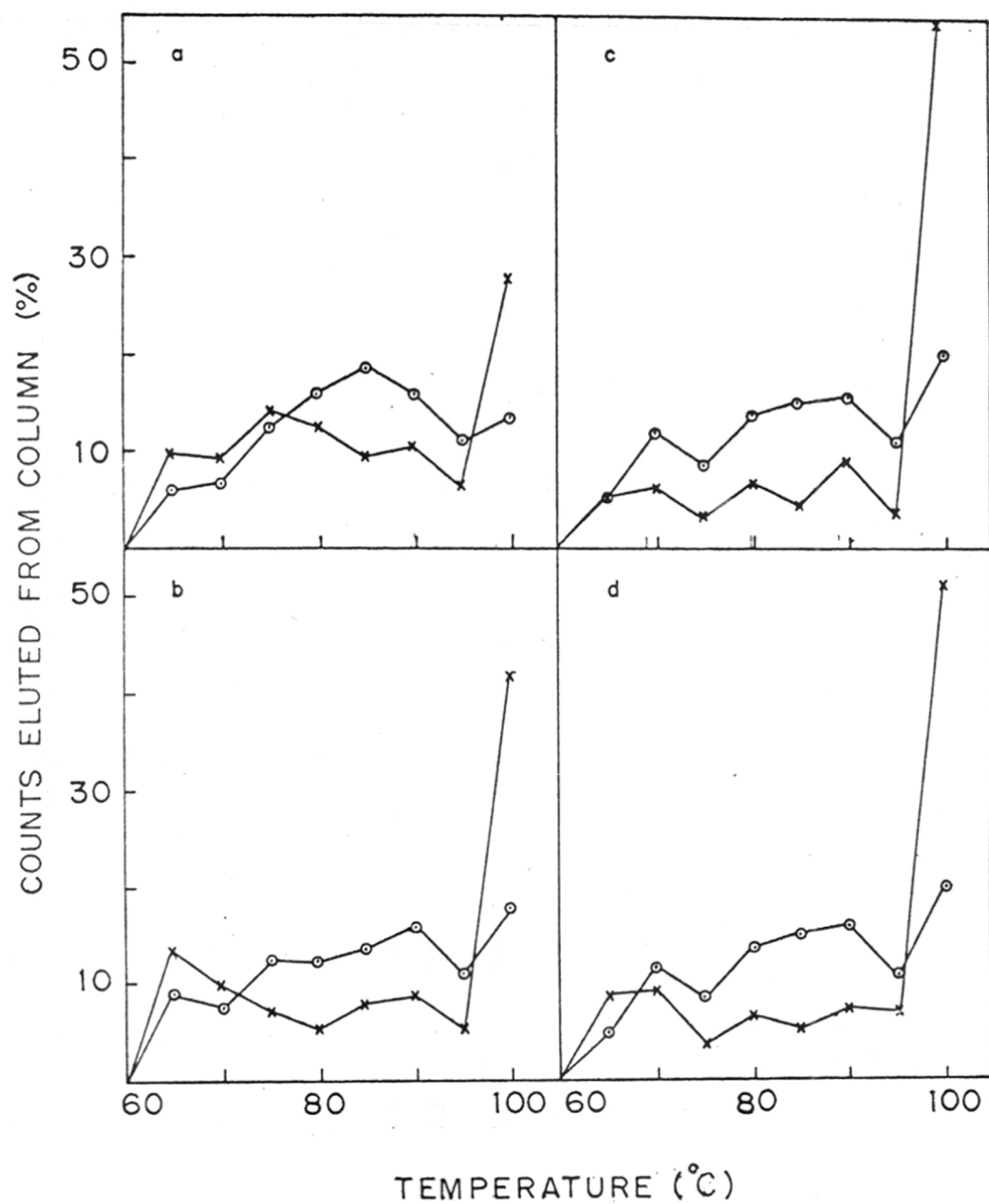


FIGURE 6-4

TABLE 6.2 - MELTING ANALYSIS OF DUPLICES USING RADIOLABELLED TOTAL PEARL MILLET DNA AS PROBE

Driver DNA	Cot value M.S	Proportion ^a (%)		T _m ^b (°C)		Δ T _m ^c (°C)		Base mismatch ^d (%)	
		Low melting	High melting	Low melting	High melting	Low melting	High melting	Low melting	High melting
<u>Pearl millet</u>		70	30	78.2	94.6	-	-	-	-
<u>Barn yard millet</u>	75	61	39	74.0	97.2	4.2	-2.6	4.2	-2.6
<u>Pearl millet</u>		70	30	77.5	96.3	-	-	-	-
<u>Fox tail millet</u>	5	50	50	72.7	98.2	4.8	-2.1	4.8	-2.1
<u>Pearl millet</u>		70	30	79.0	96.6	-	-	-	-
<u>Little millet</u>	15	25	75	73.5	97.2	5.5	-0.6	5.5	-0.6
<u>Pearl millet</u>		70	30	79.0	96.6	-	-	-	-
<u>Great millet</u>	20	40	60	74.0	98.0	5.0	-1.4	5.0	-1.4

^{a,b}Obtained from Fig.6.3 and 6.4

^cObtained using the formula: melting temperature of homologous duplexes - melting temperature of heterologous duplexes

^dObtained assuming that 1°C lowering in T_m corresponds to 1% mismatch.

(Cot 15 M.s.) and great millet (Cot 20 M.s.). The values of reassociation after correction for self-reassociation of tracer DNA are as low as 3.56% in fox tail millet DNA, 3.33% in little millet DNA and 2.42% in great millet DNA. After normalizing the corresponding homologous reassociation values to 100%, it is calculated that 10.2%, 8.4%, 13.87% and 10.17% of the repeated sequences of barn yard millet, fox tail millet, little millet and great millet respectively, are homologous to the repeated sequences of pearl millet.

The melting profiles of all the heteroduplexes are also biphasic as in homoduplexes (Figure 6.3) and the melting temperatures of the low melting fraction and high melting fractions lie in the range of 72.7 - 74.0°C and 97.2 - 98.2°C respectively (Table 6.2). After comparing the T_m of the low melting fractions of the heterologous duplexes with that of homologous duplexes, the base sequence divergence in heterologous duplexes is estimated to be 4.2 - 5.5%. The T_m values of the high melting fractions in the heterologous duplexes are higher than that of the corresponding high melting fraction of the homologous duplexes. This negative value of base mismatch (Table 6.2) indicates that the heterologous duplexes formed at high temperatures are thermally more stable than the corresponding homologous duplexes and show

no sequence divergence. As seen from Figure 6.4, as much as 40-75% of the counts of the heteroduplexes between pearl millet and the other millets are eluted at very high temperatures indicating an extensive homology between the G + C rich sequences.

Use of nick-translated repetitive DNA of barn yard millet as tracer

(i) Homologous hybridization

When ^{32}P labelled repetitive DNA of barn yard millet was incubated to a Cot value of 75 M.s. with its own cold DNA, the extent of reassociation was 70.42%. Self-reassociation of tracer DNA incubated to an equivalent of Cot 75 M.s. showed a reassociation of 22.04%. After correction for the self-reassociation of tracer DNA, the homologous reassociation is estimated to be 48.38%.

The melting temperature of the homologous duplexes is 83.5°C (Figure 6.5) and that of 0.4 kbp long DNA is 85.95°C (Chapter IV). The homologous duplexes, therefore, show a base mismatch of 2.45%.

(ii) Heterologous hybridization

Radiolabelled repetitive DNA of barn yard millet was hybridized with cold DNAs (0.4 kbp long) of fox tail

TABLE 6.3 - HOMOLOGY AMONG REPEATED DNA SEQUENCES OF MILLETS USING RADIOLABELLED REPETITIVE BARN YARD MILLET DNA

Driver DNA	Cot value M.s	Reassociation ^a (%)	Self reasso- b ciation of tracer (%)	Reassociation ^c after correction (%)	Homology ^d (%)	No. of experi- ments
<u>Barn yard millet</u>	75	70.42 ± 2.39	22.04 ± 1.05	48.38	100.00	3
<u>Fox tail millet</u>	5	18.22 ± 0.75	12.39 ± 0.90	5.85	12.05	3
<u>Little millet</u>	15	31.85 ± 3.21	13.64 ± 1.77	18.21	37.64	2
<u>Great millet</u>	20	32.18 ± 4.36	14.96 ± 2.75	17.22	35.59	3
<u>Pearl millet</u>	10	30.15 ± 0.11	13.64 ± 1.77	16.51	34.13	2

^aReassociation between homologous or heterologous hybrids at the limit Cot value of driver DNA

^bObtained by allowing tracer DNA to reassociate alone for the corresponding time as homologous and heterologous reassociation.

^cObtained as $c = b - a$

^dCalculated by normalizing the corresponding homologous experiment to 100% homology.

Figure 6.5 HYDROXYAPATITE THERMAL DENATURATION PROFILES
OF HOMOLOGOUS AND HETEROLOGOUS DNA DUPLEXES.

- (O—O) Repetitive DNA of barn yard millet
as tracer DNA x barn yard millet driver DNA
at Cot 75 M.s.
- (a) (●—●) Repetitive DNA of barn yard millet
as tracer DNA x fox tail millet driver DNA
at Cot 5 M.s.
- (b) (●—●) Repetitive DNA of barn yard millet
as tracer DNA x little millet driver DNA at
Cot 15 M.s.
- (c) (●—●) Repetitive DNA of barn yard millet as
tracer DNA x great millet driver DNA at Cot
20 M.s.
- (d) (●—●) Repetitive DNA of barn yard millet
as tracer DNA x pearl millet driver DNA at Cot 10 M.s.

Renaturation was carried out to the respective
limit Cot values of the driver DNA. Graphs
were plotted as cumulative percentage counts
eluted at each temperature interval of 5°C.

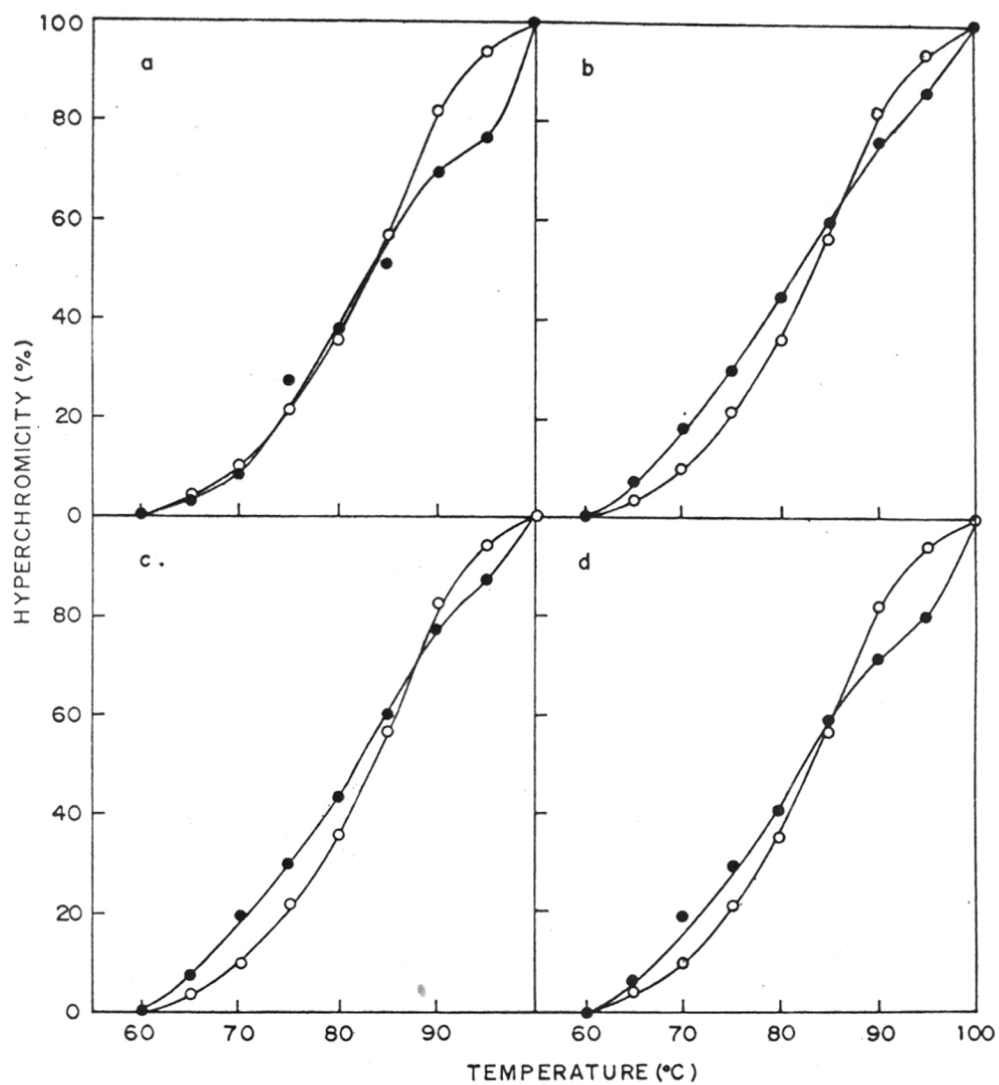


FIGURE 6.5

Figure 6.6

DERIVATIVE THERMAL DENATURATION PROFILES OF HOMOLOGOUS AND
HETEROLOGOUS DUPLEXES

- (a) (O—O) Repetitive DNA of barn yard millet as tracer DNA
x barn yard millet driver DNA at Cot 75 M.s
- (b) (X—X) Repetitive DNA of barn yard millet as tracer DNA
x fox tail millet driver DNA at Cot 5 M.s.
- (c) (Δ — Δ) Repetitive DNA of barn yard millet as tracer DNA
x little millet driver DNA at Cot 15 M.s.
- (d) ($\bullet$ — $\bullet$) Repetitive DNA of barn yard millet as tracer DNA
x great millet driver DNA at Cot 20 M.s.
- (e) ($\blacktriangle$ — $\blacktriangle$) Repetitive DNA of barn yard millet as tracer DNA
x pearl millet driver DNA.

Renaturation was carried out to the respective limit Cot values of the driver DNA. Denaturation results are graphed as the percentage of the counts eluted at each temperature rise of 5°C.

FIGURE 6.6

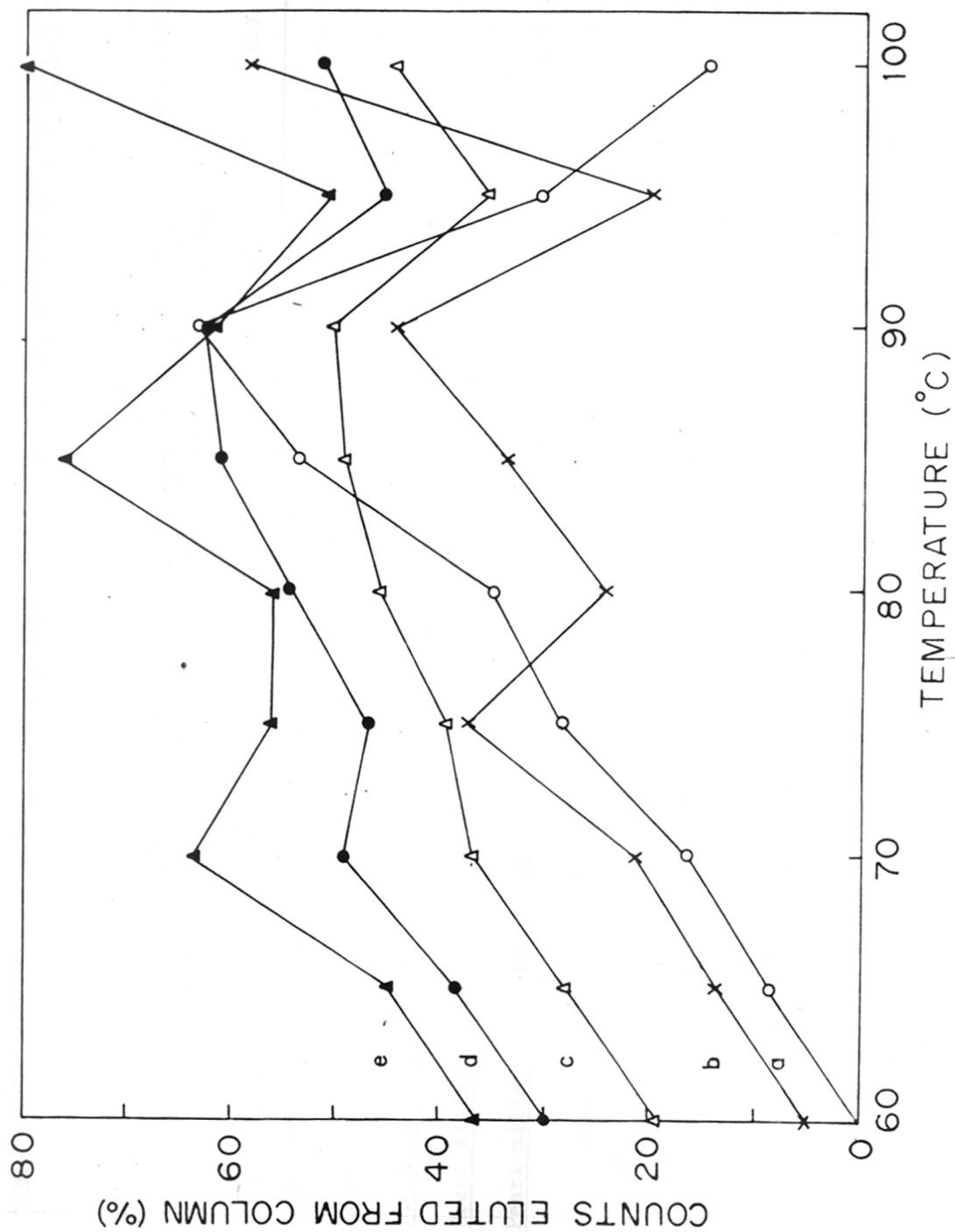


TABLE 6.4 - MELTING ANALYSIS OF DUPLEXES USING RADIOLABELLED REPETITIVE BARN YARD MILLET DNA

Driver DNA	Cot value M.s	Proportion ^a		T _m ^b		T _m ^c using low melting frac- tion	Base mismatch ^d (%)
		Low melting	High melting	Low melting	High melting		
Barn yard millet	75	100		83.5		-	-
Fox tail millet	5	65	35	78.6	96.8	4.9	4.9
Little millet	15	100		82.0		1.5	1.5
Great millet	20	100		82.0		3.5	3.5
Pearl millet	10	70	30	77.0	96.6	6.5	6.5

^{a,b}Obtained from Fig.6.5 and 6.6

^cObtained using the formula: melting temperature of homologous duplexes — melting temperature of heterologous duplexes.

^dObtained assuming that 1°C lowering in T_m corresponds to 1% mismatch.

millet, little millet, great millet and pearl millet to Cot 5 M.s., Cot 15 M.s., Cot 20 M.s. and Cot 10 M.s. respectively. The reassociation values for these heterologous reassociations are 5.83%, 18.2%, 17.22% and 16.51% respectively after correction for self-reassociation of tracer DNA (Table 6.3). Normalizing the corresponding homologous reassociation values to 100%, the sequence homologies of the repeat sequences of fox tail millet, little millet, great millet and pearl millet to the repeated sequences of barn yard millet are estimated to be 12.05%, 37.04%, 35.59% and 34.13% respectively.

The melting profiles of barn yard millet x little millet and barn yard millet x great millet heteroduplexes are monophasic, with melting temperatures in the range of 80 - 82°C (Figure 6.5). The melting profiles of barn yard millet x fox tail millet and barn yard millet x pearl millet heteroduplexes are, however, biphasic (Figure 6.5). The low melting and the high melting fractions have T_m s of 77.0 - 78.6°C and 96.6 - 96.8°C respectively. The base mismatch in the low melting fraction is in the range of 4.9 - 6.5% (Table 6.4). The high melting fractions show a negative base mismatch indicating the absence of sequence divergence. In case of little millet and great millet, the heteroduplexes show a base mismatch of 1.5 - 3.5% (Table 6.4).

DISCUSSION

Low homology among repeated DNA sequences of millets and its implication in the origin of grasses

Based on their very similar morphological features, all the millets in the present study, except great millet, have been classified together under the tribe Paniceae. Great millet belongs to the tribe Andropogoneae. We, therefore, can expect a substantial homology between the DNA sequences of the millets belonging to Paniceae. Our studies of DNA-DNA hybridization among the repeated sequences have however, revealed a strikingly low homology of 8-14%, when total pearl millet DNA is used as tracer. When isolated repetitive DNA of barn yard millet is used as tracer, the homology among the repetitive DNA sequences increases to 34 - 37% except in case of fox tail millet where still a very low homology of 12% is observed. The increase in homology in the latter case is because of the use of only repetitive DNA as tracer. Using isolated repetitive DNA of barn yard millet as tracer, we have estimated that 30.15% of the repeated sequences of pearl millet are homologous to the repeated sequences of barn yard millet. Repeated sequences constitute 42% of the total genome in barn yard millet (Chapter III). Hence, it is estimated that 12.66% ^{ie.} $\wedge \frac{(30.15 \times 42)}{100}$ of the genomes

of barn yard millet and pearl millet are homologous. This value compares well with that of 10.20% when total pearl millet DNA was used as the radiolabelled probe.

In family Gramineae, homologies among repeated DNA sequences have been extensively studied in the genomes of wheat, oat, rye and barley (9,10,24,25). Cross hybridizations between these cereals have revealed varying amounts of homology and thermal stabilities of heterologous duplexes. For example, there is extensive heteroduplex formation between wheat and rye repeated sequences. Oat DNA, on the other hand, forms considerably less heteroduplexes with all the DNAs, indicating extensive divergence of oat DNA from barley, rye and wheat DNAs. About 22.7 - 29.7% of the sequences are common to all the above four cereals which suggests that they have evolved from a common ancestor. Most of the distinct subclasses of DNA sequences which can be identified in each of these genomes have been added on after the bifurcation of lineages.

The millets in the present study show a very low homology among their repeat sequences. So it appears that they have diverged at a point where each species has acquired a new set of repeated sequences. It is also possible that they have evolved to such an extent that their

sequences are no longer homologous. The low homology which is observed in millets is comparable to the homology between wheat and maize reported by Hake and Walbot which is only 10% (27). Also it has been reported that the homology between rice and great millet is only 16% (28).

Our studies between radiolabelled wheat and great millet DNA have indicated almost a complete absence of homology between them (data not shown). These observations of low homology between grasses tend to support the idea that the origin of grasses is polyphyletic. From an ancestral stock, distinct evolutionary lines have evolved, partly through adaptations to different climatic zones and partly through the establishment of distinct differences in anatomy and reproductuary parts.

Specific probes of individual repetitive sequences are needed to be precise in describing the function and evolution of a repetitive sequence. A cloned repeat sequences will help us to determine how similar different families of repeats are, at what stage amplification has taken place and how the organization of repeats has evolved. The present investigation has formed the basis on which further detailed studies can be undertaken in future.

REFERENCES

- 1 Peacock, W.J., Dennis, E.S. and Gerlach, W.L. (1982).
In: Mechanisms of speciation, pp.123-142, Alan, R. Liss,
Inc., 150 5th Avenue, New York, N.Y. 10011.
- 2 Flavell, R.B., O'Dell, M. and Hutchinson, J. (1981).
Cold spring Harbor Symposium Qyant. Biol.Vol. XLV,
pp.501-508.
- 3 Thompson, W.F. and Murray, M.G. (1981). In: The
Biochemistry of Plants. A comprehensive Treatise
Proteins and Nucleic acids, Marcus, A. ed., Vol.6,
pp.1-81,Academic Press, New York.
- 4 Bouchard, R.A. (1982). In: International Review of Cytology,
Bourne, G.H. and Danielli, J.F. eds. Vol.76, pp.113-193,
Academic Press, New York.
- 5 Bendich, A.J. and Bolton, E.T. (1967). Plant Physiol.,
42: 959-967.
- 6 Bendich, A.J. and McCarthy, B.J. (1970a) Genetics, 65:
545-565.
- 7 Bendich, A.J. and McCarthy, B.J. (1970b) Genetics, 65:
567-573.
- 8 Smith, D.B. and Flavell, R.B. (1974). Biochem. Genet.,
12: 243-246.
- 9 Rimpau, J., Smith, D.B. and Flavell, R.B. (1978).
J. Mol. Biol., 123: 327-359.
- 10 Rimpau, J., Smith, D.B. and Flavell, R.B. (1980).
Heredity, 44: 131-149.
- 11 Nath, J., Hanzel, J.J., Thompson, J.P., McNay, J.W.(1984).
Biochem. Genet. 22:37-50.
- 12 Nath, J., McNay, J.W., Paroda, C.M., Gulati, S.C. (1983).
Biochem. Genet., 21: 745-760.
- 13 Chooi, W.J. (1971). Genetics, 68: 213-230.
- 14 Narayan, R.K.J. and Rees, H. (1977). Chromosoma (Berl.),
63: 101-107.

- 15 Goldberg, R.B., Bemis, W.P. and Siegel, A. (1972). Genetics, 72: 253-266.
- 16 Strong, H.G. and Thompson, W.F. (1975). Carnegie Inst. Wash. Yearbook, 74: 789-792.
- 17 Belford, H.S. and Thompson, W.F. (1981). Heredity, 46: 109-122.
- 18 Belford, H.S. and Thompson, W.F. (1981). Heredity, 46: 91-108.
- 19 Stein, D.B. and Thompson, W.F. (1975). Science, 189: 888-890.
- 20 Stein, D.B., Thompson, W.F. and Belford, H.S. (1979). J. Mol. Evol., 13: 215-232.
- 21 Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). J. Mol. Biol., 113: 237-251.
- 22 Berger, S.L. (1984). Anal. Biochem., 136: 515-519.
- 23 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, pp.164-169.
- 24 Smith, D.B. and Flavell, R.B. (1974). Biochem. Genet., 12: 243-256.
- 25 Flavell, R.B., Rimpau, J. and Smith, D.B. (1977). Chromosoma (Berl.), 63: 205-222.
- 26 Deshpande, V.G. and Ranjekar, P.K. (1980). Hoppe-Seylers Z. Physiol. Chem., 361: 1223-1233.
- 27 Hake, S. and Walbot, V. (1980). Chromosoma (Berl.), 79: 251-270.
- 28 Guangyu, Z., Yishen, Z., and Wanxia, Y. (1981). Scientia Sinica, Vol. XXIV, pp.701-709.

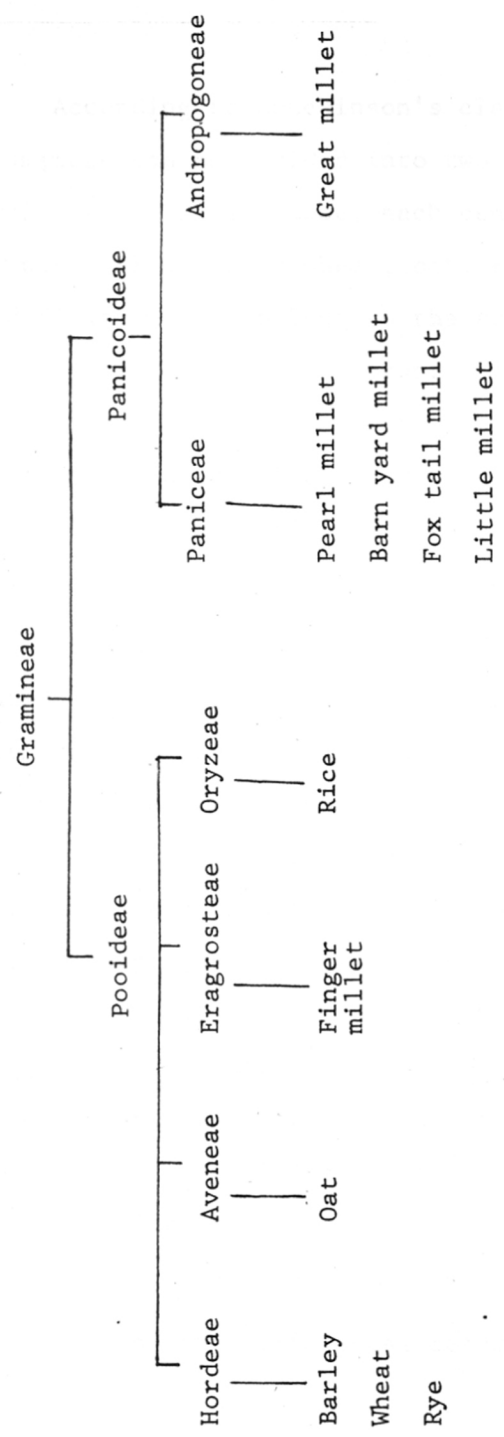
CHAPTER VII

GENERAL DISCUSSION

GENERAL DISCUSSION

Plants are rather unique in having large variations in their nuclear DNA content, high proportion of repetitive DNA and diverse nuclear organization. Information about genome organization using several molecular probes is extremely useful as it can provide answers to basic questions such as the mechanism of gene regulation, basis of chromosome evolution and phenotypic expression and functional role of different types of DNA sequences. Although many animal genes have been studied in detail using recombinant DNA technology, plant genes have seldom been the target for such studies. Moreover, before one can attempt to clone any gene from plants, a basic understanding of the organization of its genome will be very useful. Grasses, comprising the large and distinctive family Gramineae, are particularly interesting for such studies. They are very widely distributed and highly evolved flowering plants, inhabiting a wide range of soil and climatic conditions. To botanists as well as the molecular biologists, the grasses have always been an interesting but at the same time a difficult group of plants. Efforts are continuously being made to trace the descent of the present day grasses from the hypothetical ancestors, but they remain pure speculations.

FIGURE 7.1 CLASSIFICATION OF GRAMINEAE PLANT SPECIES ACCORDING TO HUTCHINSON (1973)*



* Hutchinson, J (1973). In the families of flowering plants, 3rd ed., Oxford University Press, Oxford.

According to Hutchinson's classification, family Gramineae can be divided into two subfamilies namely Pooideae and Panicoideae, each consisting of several tribes. Plants like wheat, oat, rye, barley, maize, rice and finger millet belong to the subfamily Pooideae and are classified under different tribes (Fig. 7.1). Pearl millet and the millets in the present study also known as minor millets are classified under the subfamily Panicoideae and belong to the tribe Paniceae except great millet which belongs to the tribe Andropogoneae.

In view of the above classification, it was an interesting proposition to examine whether the millets differed from the other grasses at the molecular level and if so, did these changes offer a selective advantage to them. Selective advantages may be in the form of reassortment of genes and other sequences, high genetic flexibility and/or extensive specialization. In fact, the simultaneous evolution of complexity and flexibility has been the principal feature of evolution of the genetic architecture of plant chromosomes. Our studies have shown that the minor millets have several distinguishing features that are different from those of the other cereals. These differences have been summarized in Table 7.1.

Most of the differences observed between the millets

TABLE 7.1 DISTINGUISHING FEATURES OF TROPICAL AND TEMPERATE GRASSES (FAMILY: GRAMINEAE)

1. Subfamily	Pooideae	Panicoideae
2. Plant species	Wheat, oat, rye, barley and maize	Pearl millet, barn yard millet, fox tail millet, little millet and great millet.
3. Climatic adaptations	Temperate	Tropical
4. Nuclear DNA content (1C)	>5 pg	<5 pg
5. Total repetitive DNA	4.3 - 13.8 pg	0.24 - 2.4 pg
6. Unique DNA	1.2 - 3.5 pg	0.56 - 2.2 pg
7. High melting fraction in the highly repetitive DNA	Absent	Present 3-8% of the total genome except in great millet
8. Sequence divergence among repeated DNA sequences	13-15%	8-10%
9. Nature of repeated DNA sequences	Homogeneous in barley. Data in other species not reported.	Heterogeneous in all millets and not known in pearl millet.
10. Interspersion pattern	Predominantly short	Predominantly long period in fox tail millet and pearl millet Mixed in barn yard millet, little millet and great millet.
11. Length of interspersed repeat sequences	Predominantly 0.3-1.0 kbp	Heterogeneous, 0.15-9.0 kbp
12. Homology among repeated DNA sequences.	14-58% using total DNAs as radiolabelled probes.	8-12% using total pearl millet DNA as the radiolabelled probe.

and other grasses have been discussed at length at appropriate places in various Chapters. All the data put together make it possible to draw conclusions about the rate of turn over and its implications on the wide ecological plasticity and amplitude of millets.

Based on reassociation kinetics of 0.4 kbp long DNAs at different criteria of temperature stringency, the millet genomes have been fractionated into 'fossil' repeats and 'true' single copy sequences. Essentially one third (or less) of the single copy DNA at permissive criteria behaves as 'fossil' repeats and a larger fraction is 'true' single copy. A significant fraction of 'true' single copy indicates that millet genomes are 'turning over' relatively slowly. The presence of long tandem arrays of repeats with little or no sequence divergence further substantiates the proposed low turn over rates in millets. The long repeats are assumed to maintain their internal homogeneity by cross over fixation. Further, the short repeats appear to have recently evolved from the long repeats and hence have not undergone sufficient sequence changes leading to the high thermal stability of the total repeated DNA sequences. Also the presence of a smaller fraction of repeated DNA sequences in millets indicates that only a small portion of the millet genomes is evolving by amplification and other evolutionary processes.

The short and long repeats in millets probably perform different functions. The short repeated DNA sequences are probably a product of long repeats produced by amplification and concomitant dispersion to various locations in the genome. They are possibly associated with single copy sequences and contain the regulatory domains. The frequency of repetition of the repeat DNA sequences possibly determines whether or not the adjacent single copy DNA is to be transcribed. If the frequency of repetition of the basic repeat is very high then the adjacent single copy DNA may not be transcribed at all. The long repeats in the millet genomes probably function as chromosomal spacers and bring about a folding of the chromosome. If some gene interactions are dependent on the spatial orientation, alterations in chromosomal positioning could have major effects on gene expression.

It thus appears that specific functional constraints are imposed on millet genomes to maintain a low rate of turn over. We postulate that environmental factors are responsible for the maintenance of low turn over rates. Turn over rates in turn reflect on the genome size. Presence of long tandem arrays of clustered repeats is assumed to bring about a folding of chromosomes and thus a reduction in genome size. Genome size in turn is very useful in predicting plant responses to climate, crop breeding and in the analysis

of niche differentiations in plant communities. A selective force which determines the relationship between genome size and growth under different ecological conditions arises from the differential effect of temperature on cell division and cell expansion. Species growing at low temperatures (temperate climates) have long mitotic and meiotic cycles resulting in enlargement of cells. This enlargement is not observed in plants growing in warm and arid environments and they usually show smaller genomes. Thus, temperature appears to control 'turn over' rates in millets which in turn control the genome size. The resulting small genomes are helpful in the ecological adaptations of millets.

LIST OF PUBLICATIONS

- 1 Novel molecular features of millet genomes
 Indian J. of Biochem. Biophys. (1984) (In press).
- 2 Molecular analysis of great millet (Sorghum vulgare)
 DNA
 Journal of Biosciences (1984) (In press).
- 3 DNA sequence organization in three millet
 plant species. (1984).
 (Communicated)
- 4 Fine melting analysis of DNAs of ten Gramineae
 plant species. (1984)
 (Communicated).
- 5 Nature of repeated DNA sequences in four millet
 genomes. (1984)
 (Manuscript in preparation).

PAPERS PRESENTED AT SCIENTIFIC MEETINGS

- 1 Nuclear DNA comparisons among three millet species.
Vidya Deshpande, S. Lakshmi & P.K. Ranjekar
At Second FAOB Congress & Golden Jubilee Meeting
of the SBC (India) at Bangalore, 14-18 Nov. 1980.
- 2 Characterization of repetitive DNA in three millet
species.
S. Lakshmi & P.K. Ranjekar
At the 50th Annual Meeting of the SBC (India) at
Baroda, 18-20 November 1981.
- 3 Fine melting analysis of the DNAs of three Gramineae
plant species.
S. Lakshmi
At the 52nd Annual Meeting of the SBC (India) at
Pune, 26-29 November 1983.
- 4 Molecular analysis of great millet DNA
V.S. Gupta and S. Lakshmi
At the 52nd Annual Meeting of the SBC (India) at
Pune, 26-29 November 1983.
- 5 Occurrence of highly reiterated GC rich DNA
sequences in millets.
S. Lakshmi & P.K. Ranjekar
At the XV International Congress of Genetics,
New Delhi (India), 12-21 December 1983.
- 6 Organization of plant nuclear genomes
Vidya Gupta, S. Lakshmi, M. Bhawe, S. Patankar,
S.A. Ranade and P.K. Ranjekar
At the Third International Congress on Cell Biology,
Tokyo, Japan, 26-31 August 1984.