J. Biosci., Vol. 6, Number 6, December 1984, pp. 795–809.

Molecular organization of great millet (*Sorghum vulgare*) DNA*

LAKSHMI SIVARAMAN, VIDYA S. GUPTA and
P. K.RANJEKAR**

Biochemistry Division, National Chemical Laboratory, Pune 411 008, India

MS received 30 April 1984; revised 13 October 1984.

Abstract. Approximately 52% of the nuclear genome of great millet (*Sorghum vulgare*) consists of repetitive DNA which can be grouped into very fast, fast and slow components. The reiteration frequencies of the fast and slow reassociating components are 7000 and 92 respectively. Approximately 90% of the genome consists of repeated sequences interspersed amongst themselves and with single copy sequences. The interspersed repeat sequences are of three sizes *viz.* > 1·5 kilobase pairs, 0·5–1·0 kilobase pairs and 0·15–0·30 kilobase pairs while the size of the single copy sequences is 3·0 kilobase pairs. Hence the genome organization of great millet is essentially of a mixed type.

Keywords. Great millet genome; nuclear DNA content; DNA sequence organization.

Introduction

Information about DNA sequence organization in eucaryotes is of fundamental importance as it provides a basis to understand several important parameters such as molecular events during evolution, origin of phenotypic variation and regulation of gene expression. In our laboratory, studies have been undertaken to analyse the genomes of a few Gramineae plant species with 1C nuclear DNA content less than 5 picograms (pg). The data on species with a DNA content of less than 5 pg like rice, finger millet and pearl millet (Gupta *et al*., 1981; Gupta and Ranjekar, 1981, 1982) have shown that their DNA sequence organization not only differs amongst themselves but also from that of wheat (Flavell and Smith, 1976), rye (Smith and Flavell, 1977) and maize (Hake and Walbot, 1980). While the latter species, having a DNA content of more than 5 pg, exhibit predominantly a short period interspersion pattern, the former species show a varying interspersion pattern.

Great millet (*Sorghum vulgare*) is a commonly cultivated Gramineae plant species with a haploid nuclear DNA content of 4·6 pg (Joshi, 1982). Since no molecular data are available on this species, studies were undertaken to assess the effect of nuclear DNA content on its DNA sequence organization.

---

* NCL communication No. 3527.

** To whom all correspondence should be addressed.

Abbreviations used: pg, Picograms; PIPES, piperazine-N-N′-bis (2-ethanol sulphonic acid); Tris, Tris (hydroxy methyl amino methane); kbp, Kilobase pairs; *Tm*, melting temperature; np, nucleotide pairs; HA, hydroxyapatite; RNase, ribonuclease.

Materials and methods

All the chemicals used throughout the work were of Analytical or Guaranteed Reagent grade obtained from British Drug House, Bombay, Sarabhai chemicals and/or E. Merck. RNase, piperazine-N-N′-bis(2-ethanol sulphonic acid) (PIPES) and Tris (hydroxy methyl amino methane) (Tris) were obtained from Sigma Chemical Company, St. Louis, Missouri, USA.

DN A preparation

Eight day old seedlings were used for DNA extraction. Prior to homogenization, the tissue was frozen in liquid nitrogen and powdered in a Remi blender. The DNA isolation procedure was essentially that of Ranjekar *et al.* (1976) with a minor modification. Protein was extracted twice with freshly distilled phenol saturated with Tris-HCl (pH 7·5) followed by one chloroform/isoamyl alcohol (24:1) extraction. The optical properties of the DNA were as follows:

$$\frac{A_{230}}{A_{260}} = 0{\cdot}45 \qquad \frac{A_{280}}{A_{260}} = 0{\cdot}55.$$

The absence of a 'foot' or shallow absorbance rise before the actual start of the DNA melting indicated the absence of single stranded DNA and RNA. The polysaccharide contamination in these preparations was negligible. Samples were routinely checked for their nativity on hydroxyapatite prior to reassociation and melting experiments.

Shearing and sizing of DNA

DNA fragments of an average length of 0·4 kilobase pairs (kbp) were obtained by sonication for 3 min using Biosonic III (Bronswell model, 250 W, 20 KHz) fitted with a half inch probe. Sonication with 1/4 inch probe for 1 min at 40 setting yielded DNA fragments of an average size of 1·5 kbp. DNA fragments in the range of 3·0–9·0 kbp were obtained using a Vir Tis 60 *K* homogenizer by varying the speed (30,000–58,000 rpm) and time (4–45 min). The DNA concentrations for sonication were kept constant at 400–450 μg/ml. The average size of the DNA fragments after shearing was determined by agarose slab gel electrophoresis using 1% agarose with Hind III digest of bacteriophage λ DNA and Hae III digest of ϕ × 174 RF DNA, as molecular weight markers. The molecular weight of native unsonicated DNA was estimated to be nearly 20 kbp.

Reassociation techniques

The reassociation kinetics of great millet DNA of different fragment sizes (0·4 and 20 kbp) were studied essentially by the method of Britten *et al.* (1974). Kinetic fractions enriched in highly repetitive (very fast and fast) and intermediately repetitive (slow) sequences were also prepared (Hake and Walbot, 1980). The highly repetitive DNA included sequences which were bound to hydroxyapatite at Cot 0·5 M.s. DNA sequences which bind to hydroxyapatite at Cot 20 M.s. but not at Cot 0·5 M.s. were considered to be intermediately repetitive. These DNA fractions were analyzed for their

thermal stability and kinetic heterogeneity. To fractionate populations of repeated sequences by their extent of relatedness, reassociation of 0·4 kbp long great millet DNA was followed optically at different stringency criteria in the range of 55°–75°C.

Melting studies

Thermal denaturation of total DNA and of repetitive DNA fractions was carried out as described earlier (Ranjekar *et al*., 1976). In these studies, the data were used without any correction for solvent expansion. Single strand collapse was determined according to Graham *et al.* (1974) by melting unreassociated unique great millet DNA in 0·12 M sodium phosphate buffer (pH 6·8). The value obtained was in the range of 1·2–1·6%.

Sizing of repetitive DNA sequences

Native DNA of an approximate fragment length of 20 kbp was used for S_1 nuclease digestion (Seshadri and Ranjekar, 1980). The size distribution of S_1 resistant duplexes was determined by gel filtration on an Agarose *A*-50 (100–200 mesh size, Bio Rad Laboratories) column (92 × 1·5 cms) previously calibrated with calf thymus DNA of known duplex length. The column was developed in 0·12 M sodium phosphate buffer (pH 6·8).

Results

Reassociation kinetics of short DNA fragments

Reassociation kinetics of 0·4 kbp long DNA fragments of great millet DNA are given in figure 1A. A moderately stringent standard annealing criterion equivalent to 62°C and 0·18 M Na^+ was employed for the construction of the reassociation curve. The data points were obtained using a combination of optical reassociation and hydroxyapatite chromatography. The curve was drawn through the data points by nonlinear least squares regression analysis, assuming second order kinetics and allowing all the parameters to free float. This analysis reveals that the curve can be modelled into three distinct second order components (root mean square = 0·027) differing in their reiteration frequency (table 1). About 6% of the genome reassociates very rapidly by Cot 1×10^{-3} M.s. and could not be resolved with the techniques used in the present investigation. A fast component comprising 17·4% of the genome reassociates with an observed second order rate constant of $1{\cdot}12 \times 10^2$ and a reiteration frequency of 7000. The slow repetitive component comprising 27·4% of the genome has a second order rate constant of $9{\cdot}26 \times 10^{-1}$ and a reiteration frequency of 91. While approximately 31% of the genome contains sequences which reassociate with a rate constant of $8{\cdot}97 \times 10^{-3}$ indicating the presence of mainly single copy sequences, the remaining 18% of the genome fails to reassociate at a Cot value of 1×10^4 M.s. This is attributed to DNA degradation resulting from long incubations and failure of some sequences to form stable duplexes at the above experimental conditions.

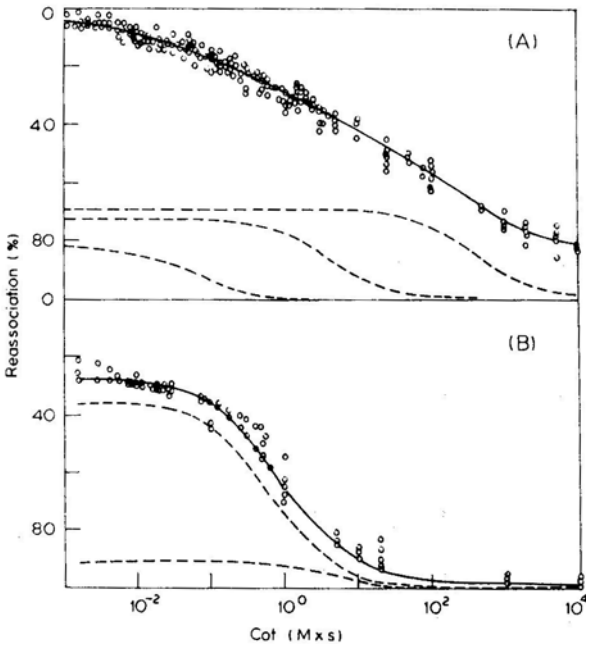

Figure 1. Reassociation kinetics of (A) 0·4 kbp long and (B) unsonicated 20 kbp long DNAs. The solid lines through the data points represent the least squares fit allowing all parameters to free float (root mean square, 0·027 for A and 0·061 for B). The lower dashed curves represent the reassociation kinetics of pure components.

Reassociation of fractions enriched with repetitive DNA sequences

For a direct demonstration of the existence of discrete families of nucleotide sequences with respect to repetition frequency, kinetics of reassociation of isolated fractions enriched in highly and intermediately repetitive DNA sequences were studied. Computer analysis of the mini-Cot curve of the highly repetitive DNA demonstrates the presence of two components (figure 2A and table 2). The faster of the two components, almost 40% of the total, has a very low Cot 1/2 pure value of 4·8 $\times 10^{-3}$ M.S. It contains sequences with a high copy number of 74,000 and a kinetic complexity of 1·5 $\times 10^{3}$ nucleotide pairs. The slower component of the fraction is typical of the intermediately repetitive class with a Cot 1/2 pure value of 2·31 $\times 10^{-1}$ M.S. and a kinetic complexity of 1·93 $\times 10^{5}$ nucleotide pairs. The intermediately repetitive class (figure 2B) could be modelled with only a single component with a Cot 1/2 pure value of 2·69 $\times 10^{-1}$ M.s. and a kinetic complexity of 2·24 $\times 10^{5}$ nucleotide pairs. The kinetics of total repetitive DNA (figure 2C) show the presence of two components with reiteration frequencies of 2620 and 256 and kinetic complexities of 2·93 $\times 10^{3}$ and 8·58 $\times 10^{5}$ nucleotide pairs respectively. From the analysis of the mini-Cot curves, it is evident that each fraction is contaminated to some extent with the other. This is partly due to the fact that highly repetitive and intermediately repetitive DNA sequences are interspersed.

Table 1. Kinetic analysis of 0·4 kbp and 20 kbp long great millet DNA.

	Fraction[a] of the total DNA	Cot 1/2[a] values observed (M.s.)	Cot 1/2[b] values pure M.s.	*K* pure[c]	Kinetic complexity in base pairs[d]	Frequency of repetition per 1C genome[e]
0·4 kbp long DNA						
Very fast						
Cot $< 1 \times 10^{-3}$ M.s.	0·0635	—	—	—	—	—
Fast						
Cot 1×10^{-3} to 5×10^{-1} M.s.	0·1736	$5{\cdot}17 \times 10^{-2}$	$8{\cdot}97 \times 10^{-3}$	$1{\cdot}12 \times 10^{2}$	$6{\cdot}94 \times 10^{3}$	$7{\cdot}62 \times 10^{3}$
Intermediate						
Cot 5×10^{-1} to $2{\cdot}0 \times 10^{1}$ M.s.	0·2741	$3{\cdot}92 \times 10^{0}$	$1{\cdot}08 \times 10^{0}$	$9{\cdot}26 \times 10^{-1}$	$3{\cdot}40 \times 10^{6}$	$9{\cdot}17 \times 10^{1}$
Unique						
Cot $2{\cdot}0 \times 10^{1}$ to 1×10^{4} M.s.	0·3103	$3{\cdot}59 \times 10^{2}$	$1{\cdot}11 \times 10^{2}$	$8{\cdot}97 \times 10^{-3}$	$3{\cdot}05 \times 10^{8}$	$1{\cdot}00 \times 10^{0}$
Unreassociated	0·1823	—	—	—	—	—
Cot $> 1 \times 10^{4}$ M.s.						
20 kbp long DNA						
Very fast Cot $< 1{\cdot}6 \times 10^{-3}$ M.s.	0·2781	—	—	—	—	—
Fast Cot $1{\cdot}6 \times 10^{-3}$ to $2{\cdot}0 \times 10^{1}$ M.s.	0·6396	$7{\cdot}14 \times 10^{-1}$	$4{\cdot}57 \times 10^{-1}$	$2{\cdot}19 \times 10^{0}$	$3{\cdot}88 \times 10^{5}$	$5{\cdot}48 \times 10^{2}$
Intermediate						
Cot $2{\cdot}0 \times 10^{1}$ to $1{\cdot}0 \times 10^{4}$ M.s.	0·0837	$3{\cdot}91 \times 10^{2}$	$3{\cdot}27 \times 10^{1}$	$3{\cdot}06 \times 10^{-2}$	$2{\cdot}78 \times 10^{7}$	$1{\cdot}00 \times 10^{0}$
Unreassociated Cot $> 1{\cdot}0 \times 10^{4}$ M.s.	0·0014	—	—	—	—	—

[a] Values obtained from figure 1.
[b] Cot 1/2 values observed × fraction of genome.
[c] K pure = (Cot 1/2 pure)$^{-1}$
[d] Values obtained by comparing the Cot 1/2 of 5·3 M.s. under our experimental conditions and genomes size $4{\cdot}5 \times 10^{6}$ base pairs of *Escherichia coli*.
[e] Cot 1/2 of unique DNA divided by Cot 1/2 values of the given fraction.

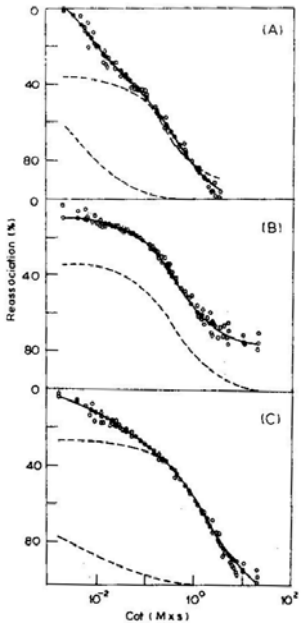

Figure 2. Reassociation kinetics of (A) highly repetitive (Cot < 0·5 M.s.) (B) intermediately repetitive (Cot 0·5–20 M.s.) and (C) total repetitive (Cot 20 M.s.) fractions isolated from 0·4 kbp long great millet DNA as described earlier in Materials and methods. The solid lines through the data points represent the best least squares solutions for the three fractions.

Melting properties of total DNA and repetitive DNA fragments

Native great millet DNA (20 kbp long) melts in 0·12 M sodium phosphate buffer, pH 6·8 with a melting temperature (*Tm*) of 87·7°C and an average hyperchromicity of 25%. The optical dispersion, defined by Mahler and Dutton (1964) is estimated to be 7·5°C and the $G + C$ content as estimated from its *Tm* is 44·9%. Sonicated DNA (0·4kbp) has a *Tm* of 86·9°C. The Cot 5×10^{-1} fraction isolated on hydroxyapatite, which mainly includes foldback sequences and highly repetitive DNA, has a thermal stability (*Tm* 80·6°C) lower than that of native DNA and shows a sequence divergence of 6·84%. The intermediately repetitive DNA (Cot 0·5–2·0) has a *Tm* of 77·25°C and a sequence divergence of 10 % (table 3).

Kinetic analysis of 0·4 *kbp long DNA at different stringency criteria of temperature*

The optical reassociation of 0·4 kbp long great millet DNA was studied till a Cot value of $3{\cdot}0 \times 10°$ M.s. at 55°, 68° and 75°C apart from the standard criterion of 62°C. By Cot $3{\cdot}0 \times 10°$ M.s., all the highly repetitive and a part of the intermediately repetitive DNA sequences reassociate. The reassociation behaviour of these DNA sequences under different stringency conditions was compared with that at the standard criterion of 62°C. It can be seen from figure 3 that the reassociation curves obtained at 55°C and 62°C could be resolved into two components while those at 68°C and 75°C showed only a single component. The rate constant obtained for the fast reassociating component at

Table 2. Kinetic analysis of mini Cot curves.

	Fraction[a]	Cot 1/2[a] values observed (M.s.)	Cot 1/2[b] values pure (M.s.)	*K* pure[c]	Frequency[d] of repetition per 1C genome	Kinetic[e] complexity in base pairs
Cot 0·5 *fraction*						
Very fast Cot $2{\cdot}3 \times 10^{-3}$ to $4{\cdot}0 \times 10^{-2}$ M.s.	0·37	$4{\cdot}84 \times 10^{-3}$	$1{\cdot}77 \times 10^{-3}$	$5{\cdot}65 \times 10^{2}$	74174	$1{\cdot}48 \times 10^{3}$
Fast Cot $4{\cdot}0 \times 10^{-2}$ to $3{\cdot}0 \times 10^{0}$ M.s.	0·64	$3{\cdot}61 \times 10^{-1}$	$2{\cdot}31 \times 10^{-1}$	$4{\cdot}33 \times 10^{0}$	995	$1{\cdot}93 \times 10^{5}$
Cot 0·5–*Cot* 20 *fraction*						
Fast Cot $2{\cdot}0 \times 10^{-3}$ to $2{\cdot}0 \times 10$ M.s.	0·66	$4{\cdot}08 \times 10^{-1}$	$2{\cdot}69 \times 10^{-1}$	$3{\cdot}72 \times 10^{0}$	880	$2{\cdot}24 \times 10^{5}$
Cot 20 *fraction*						
Very fast Cot $1{\cdot}6 \times 10^{-3}$ to $1{\cdot}13 \times 10^{-1}$ M.s.	0·26	$1{\cdot}37 \times 10^{-2}$	$3{\cdot}51 \times 10^{-3}$	$2{\cdot}85 \times 10^{2}$	2620	$2{\cdot}93 \times 10^{3}$
Fast Cot $1{\cdot}13 \times 10^{-1}$ to $2{\cdot}00 \times 10^{1}$ M.s.	0·73	$1{\cdot}40 \times 10^{0}$	$1{\cdot}03 \times 10^{0}$	$9{\cdot}71 \times 10^{-1}$	256	$8{\cdot}58 \times 10^{5}$

[a] Values obtained from figure 3. [b] Cot 1/2 value observed × fraction of genome.
[c] *K* Pure = (Cot 1/2 pure)$^{-1}$. [d] Cot 1/2 of unique DNA obtained from figure 2A divided by Cot 1/2 value of the given fraction.
[e] Values obtained by comparing to Cot 1/2 of 5·3 M.s. under our experimental conditions and genome size $4{\cdot}5 \times 10^{6}$ base pairs of *E. coli*.

Table 3. Melting analysis of total DNA and repetitive DNA fractions.

DNA	*Tm* (°C)	Hyperchromicity[a] (%)	Base mismatch[b] (%)
Unsonicated 20 kbp long DNA	87·80 ± 0·21	24·88 ± 2·5	—
Native 0·4 kbp long DNA	86·90 ± 0·90	23·80 ± 1·0	—
Cot 0·5 DNA	80·06 ± 0·84	17·30 ± 0·8	6·84
Cot 0·5–20 DNA	77·25 ± 0·45	20·57 ± 1·2	9·65
Cot 20 DNA	79·55 ± 0·25	18·25 + 0·7	7·35

[a] Hyperchromicity (H) was calculated using the formula:

$$H = \frac{A_{260}(98°C) - A_{260}(60°C)}{A_{260}(98°C)}$$

[b] Per cent base mismatching is calculated as: lowering in *Tm* by 1°C corresponds to 1% base mismatch (Ullman and McCarthy, 1973).

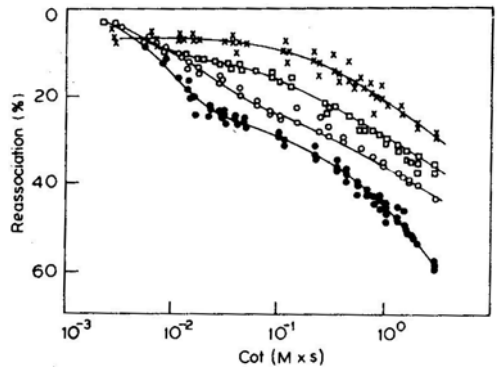

Figure 3. Optical reassociation kinetics of 0·4 kbp long great millet DNA followed upto Cot 3·0 M.S. at different conditions of stringency. (X), 75°C; (□), 68°C; (○), 62°C; (●) 55°C. The solid lines through the data points represent the best least squares solutions.

62°C is compared with the rates of the single component obtained at 68° and 75°C and the rate of the fast component obtained at 55°C, as the second component did not reassociate completely at a Cot value of 3·0 × 10° M. s. It can be seen from figure 3 that when DNA reassociation is studied at 75°C, the extent of reassociation at Cot 3 reduces from 60% to 30%, indicating that at this temperature some of the repeated sequences no longer behave as repetitive. After the reassociation rate is corrected for the effect of temperature and the decreased extent of reaction (Bonner *et al.*, 1973; Braun *et al.*, 1978), it can be seen from table 4 that the observed rate of reaction shows a 35 fold decrease (from 60·38 to 1·73) as the reassociation temperature increases from 62° to 75°C. This is because at a higher temperature, only well matched duplexes reanneal (*Tm* 83·5°C) indicating an increase in their specificity. At a lower reassociation temperature, the rate decreases (9·44 M.s. at 55°C) 6 fold but duplexes formed show

mismatch (*Tm* 76·5°C) indicating decreased specificity. Melting analysis of the duplexes formed under different stringency criteria of temperature (table 4) shows that the duplexes formed at 75°C are more stable than those at 68°, 62° or 55°C

Reassociation kinetics of long DNA fragments

The reassociation kinetics of 20 kbp great millet DNA is shown in figure 1B and the results are summarized in table 1. The repetitive DNA could be resolved only as a single component at this fragment length indicating that the highly repetitive and intermediately repetitive DNA components are linked together. Extensive interspersion of these components indicates the fact that at a Cot value of 5×10^{-1} M.s. the binding of duplexes to hydroxyapatite increases from 25 to 55%. At the repetitive Cot value of 20 M.s., the per cent reassociation increases from 52 to 92%. This suggests that the repetitive components are also extensively interspersed with single copy DNA sequences in the great millet genome. There is also an increase in the DNA binding to hydroxyapatite from 6 to 30% at a low Cot value of 1×10^{-3} M.s. This is attributed to the interspersion of very highly repetitive and/or foldback sequences along with other repetitive and single copy sequences.

Hyperchromicity of reassociated duplexes

The reassociated DNA duplexes from 0·4 kbp, 1·5 kbp and 20 kbp long DNA fragments were isolated at Cot 20 M.s. and melted to determine their thermal stability and hyperchromicity. From table 5, it can be seen that the *Tm* of the reassociated duplexes (85°–87°C) from 1·5 and 20 kbp DNA is very close to that of the native unsheared DNA, whereas the *Tm* of those isolated from 0·4 kbp DNA is much lower (79·5°C). Long repetitive duplexes thus have a greater thermal stability than the shorter ones. The hyperchromicity of the 0·4 kbp duplexes is 19·5% in contrast to 14% for 20 kbp fragments. This is mainly due to the increasing presence of single stranded regions in the reassociated fragments. Using the hyperchromicity values, it is estimated that 84 % of the 0·4 kbp long DNA and 58% of the 20 kbp long fragment is base paired. This is expected if there is extensive interspersion of single copy and repetitive DNA sequences.

Size distribution of repetitive and single copy DNA duplexes

An independent approach to determine the presence of single strand regions in the reassociated duplexes is to subject the duplexes from 20 kbp DNA to S_1 nuclease digestion and study the melting properties of the S_1 resistant fraction. Approximately 55% of the 20 kbp DNA which binds to hydroxyapatite at Cot 20 M.s. is S_1 resistant and hence in the duplex form. The hyperchromicity of the S_1 resistant DNA is 23·6% as against 14% of the duplexes not treated with S_1 nuclease.

To estimate the size distribution of repeated DNA sequences, the S_1 resistant duplexes were fractionated on Agarose A-50 Biogel column. From figure 4B, it is clear that the repetitive duplexes have a heterogeneous size distribution and can be grouped into three classes. The proportion of the excluded DNA fraction (> l·5 kbp) ranges from 43 to 48%. The second fraction which accounts for 32–38% consists of DNA

Table 4. Analysis of 0·4 kbp great millet DNA by optical reassociation and melting at different criteria of temperature stringency.

Incubation temperature and Cot range	Fraction[a]	Cot 1/2[b] observed (M.s.)	K[c] observed	Tm[d] (°C)	Tp[e] (°C)	$T_i - T_p$[f]	Rate[g] Relative to optimum rate	Adjusted[h] K	Kinetic[i] complexity in base pairs	Frequency of[j] repetition per 1C genome
55°C ($RMS = 0{\cdot}014$)										
$1{\cdot}6 \times 10^{-2} - 4{\cdot}27 \times 10^{-1}$	0·15	$9{\cdot}63 \times 10^{-2}$	10·38	76·5	81·7	−26·70	1·10	9·44	$1{\cdot}36 \times 10^{4}$	3095
$4{\cdot}27 \times 10^{-1} - 3{\cdot}00 \times 10^{0}$	0·64	$5{\cdot}08 \times 10^{0}$	0·20					0·18	$3{\cdot}02 \times 10^{6}$	20
62°C ($RMS = 0{\cdot}026$)										
$3{\cdot}2 \times 10^{-2} - 5{\cdot}00 \times 10^{-1}$	0·28	$1{\cdot}69 \times 10^{-2}$	59·17	79·4	83·15	−21·15	0·98	60·38	$4{\cdot}04 \times 10^{3}$	6731
$5{\cdot}0 \times 10^{-1} - 3{\cdot}00 \times 10^{0}$	0·31	$2{\cdot}69 \times 10^{0}$	0·37					0·38	$6{\cdot}93 \times 10^{5}$	42
68°C ($RMS = 0{\cdot}020$)										
$2{\cdot}2 \times 10^{-3} - 3{\cdot}00 \times 10^{0}$	0·29	$2{\cdot}54 \times 10^{-1}$	3·94	82·0	84·45	−16·45	0·75	5·25	4.67×10^{4}	585
75°C ($RMS = 0{\cdot}026$)										
$2{\cdot}6 \times 10^{-3} - 3{\cdot}00 \times 10^{0}$	0·30	$9{\cdot}04 \times 10^{-1}$	1·11	83·5	85·20	−10·20	0·64	1·73	$1{\cdot}47 \times 10^{5}$	193

[a,b] Values obtained from the optical reassociation curves in figure 3.

[c] $K = 1/\text{Cot } 1/2$ observed.

[d] Tm of duplexes formed after reassociation at different criteria of temperature stringency.

[e] $T_p = (T_n + T_m)/2$. T_n is the melting temperature of well matched DNA, 86·9°C for 0·4 kbp long great millet DNA.

[f] T_i = incubation temperature.

[g] Taken from figure 8 of Bonner *et al.* (1973).

[h] Calculated from (observed rate)/(rate relative to optimum rate).

[i] Values 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 $4{\cdot}5 \times 10^{6}$.

[j] Cot 1/2 value of unique DNA as described in table 1 (0·4 kbp long) divided by Cot 1/2 value of given fraction.

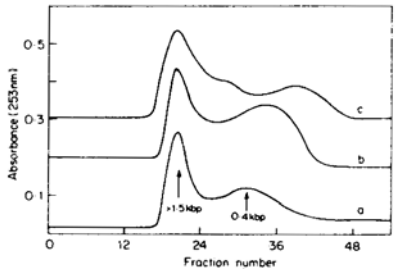

Figure 4. Agarose A-50 column chromatographic profiles of (a) calf thymus DNA of known duplex lengths (b) S_1 nuclease resistant Cot 20 great millet DNA and (c) S_1 nuclease resistant Cot 0·1 great millet DNA.

sequences of 0·15–0·30 kbp. The third class of repetitive duplexes has a size distribution in the range of 0·5 to 1·0 kbp and accounts for about 14–25% of the total duplexes.

Experiments were also carried out where the DNAs of high molecular weight (> 20 kbp) were reassociated to Cot 0·1 M.s. and then treated with S_1 nuclease and the S_1 resistant duplexes fractionated on Agarose A-50 to determine the distribution of long and short repeats in the highly repetitive fast reassociating fraction (figure 4, profile C). Three distinct sizes of repeat classes are observed in the elution profile. The proportion of the short repeats is about 28–34%, that of the intermediate repeats is about 28–34% and that of the long repeats in 35–41% (table 6).

The length of interspersed single copy sequences can be estimated from a curve relating the fraction of DNA fragments binding to hydroxyapatite as a function of DNA fragment length. In figure 5, two slopes are observed and the change in slope occurs at a fragment length of 2·5–3·0 kbp which is the estimated length of the single copy DNA. A more gradual increase in slope after 3·0 kbp indicates the presence of a spectrum of single copy sequences with lengths greater than 3·0 kbp. Extrapolation of the line with the greater slope to the ordinate indicates the proportion of repetitive DNA in the great millet genome to be 54%.

Discussion

The present studies have yielded information about the gross organization of the nuclear DNA of great millet. The repetitive DNA content as determined by the reassociation kinetics of 0·4 kbp long DNA is 52%. This value compares well with those obtained using other experimental approaches cited in the present investigation (table 7). Mini Cot curves have enabled us to detect the presence of a very fast reassociating DNA fraction containing DNA sequences with high reiteration frequency ($7{\cdot}4 \times 10^4$) and an average complexity of $2{\cdot}3 \times 10^3$ nucleotide pairs. Melting analysis of enriched fractions indicates that the highly repetitive sequences are more stable than the intermediately repetitive DNA. Studies on reassociation at different stringency conditions have demonstrated that stable families reassociating at a higher temperature have lower copy number and reassociation rate than those reassociating at a lower temperature. It, therefore, appears that repeat families in the great millet

Table 5. Melting analysis of Cot 20 DNA isolated from DNAs of different fragment lengths.

DNA fragment size (nucleotide pairs)	400	1500	5600	8200	20000	Native DNA (0·4 kbp)
Fraction bound to HA at Cot 20[a]	0·5	0·60	0·815	0·865	0·935	—
Hyperchromicity[b]	19·5 ± 2·63	20·1 ± 0·15	18·6 ± 1·12	17·8 ± 1·90	14·0 ± 4·6	23·0 ± 0·8
Tm°C	79·5 ± 0·19	86·2 ± 0·15	86·4 ± 1·12	86·9 ± 0·2	87·0 ± 0·3	—
Duplex content (D)[c]	0·839	0·867	0·798	0·761	0·586	—
Average duplex length[d]	335·6	1300·5	4468·8	6240·2	11720·0	—
Duplex content from S_1 nuclease[e]	—	—	—	—	0·550	—

[a] Obtained from figure 5.

[b] Hyperchromicity (H) was calculated using the formula:

$$H = \frac{A_{260}(98°C) - A_{260}(60°C)}{A_{260}(98°C)}$$ (Zimmerman and Goldberg, 1977; Wimpee and Rawson, 1977).

[c] The average duplex content (D) of bound fragment was estimated using the formula

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

[d] The average length of duplex region is the product of the duplex content (D) and the fragment length of DNA.

[e] Duplex content of S_1 nuclease treated repetitive DNA fragment was obtained as:

$$\frac{\mu g \text{ of } S_1 \text{ nuclease resistant repetitive duplexes bound to hydroxyapatite column}}{\mu g \text{ of } S_1 \text{ nuclease treated total DNA loaded on hydroxyapatite column}}$$

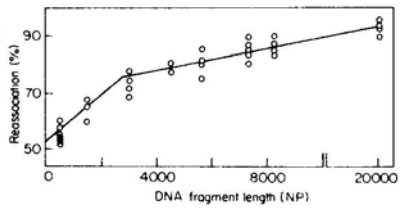

Figure 5. The percentage binding of great millet DNA to hydroxyapatite at Cot 20 M.s. as a function of DNA fragment length.

Table 6. Size distribution of S_1 resistant fractions.

Cot value M.s.	Fraction[a] of long repeats (%)	Fraction[b] of repeats of intermediate length (%)	Fraction[c] of short repeats (%)
0·1	35–41	28–34	28–34
20·0	43–48	32–38	14–25

[a, b, c] Obtained from figure 4.

Table 7. Duplex content of repetitive DNA fractions of great millet from independent experimental approaches.

Method	Duplex content
Reassociation curve of 0·4 kbp long DNA (figure 1A)	0·520
Hyperchromicity measurements (table 5)	0·586
S_1 nuclease resistance measurements (table 5)	0·550
Reassociation (%) *vs* fragment length (figure 5)	0·540

genome are essentially heterogeneous (Bendich and Anderson, 1977; Bendich, 1979; Preisler and Thompson, 1981). The increased complexity and decreased frequency of repetition, when reassociation is carried out under rigorous criteria, indicate that the family members of the highly repetitive class in the great millet genome are only broadly related. The minimum estimate of repeated sequence copy number is 193. They may represent a cryptic class of repeated sequences undetected in the reassociation kinetics of total 0·4 kbp long great millet DNA.

The reassociation kinetics data of long DNA fragments, hyperchromicity measurements of reassociated DNAs of different fragment lengths and S_1 nuclease experiments have provided a clear evidence about the occurrence of extensive interspersion of repeated and single copy DNA sequences in the great millet genome. Table 8 summarizes the available DNA sequence organization data in the family Gramineae.

Table 8. Genome organization in a few Gramineae plant species.

Species	DNA content 1C (pg)	Repetitive DNA content (%)	Length of interspersed repetitive DNA np	Length of interspersed single copy DNA np	Pattern of interspersion	Reference
Wheat *Triticum aestivum*	18·10	75·0	350– 650 (32%) 400– 800 (60%) Long repeats (10%)	1000 (88%) 1000 (12%)	Predominantly short	Flavell and Smith 1976
Rye *Secale cereale*	9·00	70·0	550– 680 (Major) 2000–3000 (Minor)	1500	Predominantly short	Smith and Flavell, 1977
Maize *Zea mays*	7·70	58·2	300–1000	2100	Short	Hake and Walbot, 1980
Finger millet *Eleucine coracana*	1·60	49·0	150– 200 (57–64%) 500–1000 (18–20%) 4000–4200 (18–23%)	1900	Mixed	Gupta and Ranjekar, 1981
Pearl millet *Pennisetum americanum*	2·50	54·0	4300–4500	1900	Long	Gupta and Ranjekar, 1982
Rice *Oryza sativa*	0·60	52·0	50– 100 (60–65%) 6000–6400 (35–40%)		Absent	Gupta *et al.*, 1981

From this table, it can be seen that the lengths of the interspersed repeated DNA sequences are diverse, varying in the range of 50–6400 nucleotide pairs (np) while that of the interspersed single copy DNA varies in a narrow range of 1000–2000 np. Great millet has essentially a mixed type of DNA sequence organization pattern. This is in keeping with the trend that plants with 1C nuclear DNA content of less than 5 pg have diverse patterns of DNA sequence organization while those with 1C nuclear DNA content greater than 5 pg have predominantly a short period interspersion pattern. The above observations lead to the question whether nuclear DNA content plays a role in determining the sequence organization pattern in plants and if so what is its significance. Attempts in this direction are in progress.

Acknowledgements

One of the authors (L.S.) thanks the Council of Scientific and Industrial Research, New Delhi for the award of a fellowship. We are grateful to Dr. V. Shankar for critically examining the manuscript and offering useful comments and suggestions.

References

Bendich, A. J. (1979) in *Molecular Biology of Plants*, (ed. L. Rubenstein) (New York: Academic Press) p. 1.
Bendich, A. J. and Anderson, R. S. (1977) *Biochemistry*, **16,** 4655.
Bonner, T. I., Brenner, D. J., Neufeld, B. R. and Britten, R. J. (1973) *J. Mol. Biol*., **81,** 123.
Braun, B. A., Schanke, K. E. and Graham, D. E. (1978) *Nucleic Acids Res*., **5,** 4283.
Britten, R. J., Graham, D. E. and Neufeld, B. R. (1974) *Methods Enzymol*., **29E,** 363.
Flavell, R. B. and Smith, D. B. (1976) *Heredity*, **37,** 231.
Graham, D. E., Neufeld, B. R., Davidson, E. M. and Britten, R. J. (1974) *Cell*, **1,** 127.
Gupta, V. S., Gadre, S. R. and Ranjekar, P. K. (1981) *Biochim. Biophys. Acta*, **656,** 147.
Gupta, V. S. and Ranjekar, P. K. (1981) *J. Biosci*., **3,** 417.
Gupta, V. S. and Ranjekar, P. K. (1982) *Indian J. Biochem. Biophys*., **19,** 167.
Hake, S. and Walbot, V. (1980) *Chromosoma*, **79,** 251.
Joshi, C. P. (1982) *Cytological and Molecular Biological Studies in Plant Genomes*, Ph. D. Thesis, University of Poona, Poona.
Mahler, H. R. and Dutton, G. (1964) *J. Mol. Biol*., **10,** 157.
Preisler, R. S. and Thompson, W. F. (1981) *J. Mol. Evol*., **17,** 78.
Ranjekar, P. K., Pallota, D. and Lafontaine, J. G. (1976) *Biochim. Biophys. Acta*, **425,** 30.
Seshadri, M. and Ranjekar, P. K. (1980) *Biochim. Biophys. Acta*, **610,** 211.
Smith, D. B. and Flavell, R. B. (1977) *Biochim. Biophys. Acta*, **474,** 82.
Ullman, J. S. and McCarthy, B. J. (1973) *Biochim. Biophys. Acta*, **294,** 405.
Wimpee, C. F. and Rawson, J. R. Y. (1979) *Biochim. Biophys. Acta*. **562,** 192.
Zimmerman, J. L. and Goldberg, R. B. (1977) *Chromosoma,* **59,** 227.