Genet. Res., Camb. (1992), **60**, *pp.* 7–13 *With* 4 *text-figures Printed in Great Britain*

Evidence for random distribution of sequence variants in *Tenebrio molitor* satellite DNA

MIROSLAV PLOHL,[1] BRANKO BORŠTNIK,[2] VLATKA LUCIJANIĆ-JUSTIĆ[1] AND ĐURĐICA UGARKOVIĆ[1]

[1] *Ruđer Bošković Institute, Bijenička 54, PO Box 1016, 41000 Zagreb, Croatia*
[2] *Boris Kidrič Institute of Chemistry, Hajdrihova 19, PO Box 30, 61115 Ljubljana, Slovenia*

(*Received 22 November 1991 and in revised form 14 January 1992*)

Summary

Tenebrio molitor satellite DNA has been analysed in order to study sequential organization of tandemly repeated monomers, i.e. to see whether different monomer variants are distributed randomly over the whole satellite, or clustered locally. Analysed sequence variants are products of single base substitutions in a consensus satellite sequence, producing additional restriction sites. The ladder of satellite multimers obtained after digestion with restriction enzymes was compared with theoretical calculations and revealed the distribution pattern of particular monomer variants within the satellite. A defined higher order repeating structure, indicating the existence of satellite subfamilies, could not be observed. Our results show that some sequence variants are very abundant, being present in nearly 50% of the monomers, while others are very rare (0–1% of monomers). However, the distribution of either very frequent, or very rare sequence variants in *T. molitor* satellite DNA is always random. Monomer variants are randomly distributed in the total satellite DNA and thus spread across all chromosomes, indicating a relatively high rate of sequence homogenization among different chromosomes. Such a distribution of monomer variants represents a transient stage in the process of sequence homogenization, indicating the high rate of spreading in comparison with the rate of sequence variant amplification.

1. Introduction

The eukaryotic genome contains many non-coding DNA sequences with unknown function, like satellite DNAs. Sequence analysis of randomly cloned satellite repeats reveals different levels of sequence conservation among species. Some complex sequences show low levels of sequence variation (between less than 1 and 3%), as in *D. melanogaster* and feline satellite DNA (Lohe & Brutlag, 1986; Fanning, 1987). However, there are satellite sequences that vary about 30% within the species (Trick & Dover, 1984). Sequence variants can be organized in tandem arrays that constitute satellite subfamilies with clearly definable higher order repeating structures. This is the case with human alpha satellite DNA which makes 3–5% of each chromosome and is organized in at least 33 different subfamilies identified up to now (Choo *et al.* 1991). Higher order repeat variants of human alpha satellite DNA are clustered on a specific chromosomal subset (Warburton & Willard, 1990; Durfy, 1990). In the satellite DNA of the mouse (Horz & Zachau, 1977) and Guinea pig (Altenburger *et al.* 1977) different sequence variants are organized in defined regions of satellite repeat. On the other hand, all transition stages during homogenization of randomly produced repeat variants can be found in two analysed tandemly repeated families of *Drosophila* (Strachan *et al.* 1985).

An organism which is especially suitable for the study of satellite DNA is *Tenebrio molitor* (the mealworm beetle) whose genome contains a large amount (as much as 50%) of single satellite DNA (Petitpierre *et al.* 1988; Ugarković *et al.* 1989). According to the genome size (Juan & Petitpierre, 1989) this single satellite DNA is represented with approximately $1{\cdot}7 \times 10^6$ copies per haploid genome. It is evenly distributed over centromeric regions of all ($2n = 20$) *T. molitor* chromosomes (Davis & Wyatt, 1989; Juan *et al.* 1990). The satellite monomer is 142 bp long, showing high sequence homogeneity. Sequence analysis of 54 randomly cloned monomers

* Reprint requests to: Dr Miroslav Plohl, Department of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička 54, P.O. Box 1016, 41000 Zagreb, Tel. (38-41)435-111/1341, Fax (38-41)425-647 Croatia.

reveals base substitutions as the most frequent mutations. The average sequence variability is 2·5 nt/monomer, or 1·8% (Ugarković *et al.* 1989; Davis & Wyatt, 1989). The frequency of single nucleotide insertions or deletions is only 7×10^{-4} events/ position (Davis & Wyatt, 1989). No insertions or deletions of longer sequences (up to 20-mer) could be detected. Point mutations are positioned randomly along the monomer sequence.

The aim of this study was to investigate sequential organization of different monomer variants within *T. molitor* satellite DNA. To get information about the complete population of satellite monomer variants, total genomic DNA was digested with a number of restriction enzymes and the electrophoretic patterns obtained were analysed. The analysis of monomer variants with a single base substitution generating particular new restriction site shows different frequencies and random distribution of variants within the satellite. To our knowledge, we present the first example where, in an extremely abundant and homogeneous satellite DNA, one sequence variant represents almost 50% of monomers and is completely randomly distributed within the satellite DNA as a result of sequence homogenization.

2. Materials and Methods

Total DNA was isolated from *T. molitor* larvae according to the standard proteinase K/phenol/RNase A extraction procedure (Sambrook *et al.* 1989). Digestions with restriction enzymes were performed under standard conditions. To check the completeness of the digestion, fresh enzyme was added to the aliquot of already digested material and both samples were compared after electrophoresis on agarose gel.

The relative amounts of satellite DNA in bands obtained after restriction digestion and electrophoretic separation on agarose gels were determined by densitometry in a Pye Unicam SP8-100 UV/VIS spectrophotometer.

Hybridization experiments were performed after Southern transfer on Hybond N (Amersham) membrane, under stringency conditions allowing hybridization of sequences of at least 90% homology. The hybridization probe was cloned *T. molitor* satellite DNA 5-mer, labelled with digoxigenin-dUTP by the random primed DNA labelling method (DNA Labeling and Detection Kit, Boehringer).

Results of densitometric measurements were analysed in order to find out how the sequence variants were distributed along tandem repeats. The test is based on a simple geometric progression (Southern 1975) of multimer frequencies. If variants are distributed randomly, the probability of appearance $p(n)$ of fragments in each band of the electrophoretic ladder is given by the equation:

$$p(n) = P^2(1-P)^{n-1}, \quad (1)$$

where P is the probability of occurrence of a particular sequence variant and n is the segment length, expressed as number of monomers. The measured density of band $d(n)$ at a position n is proportional to the length of the segment and to the probability of its appearance:

$$d(n) \alpha np(n) \quad (2)$$

Consequently, the distribution of a particular recognition site is random when the logarithmic plot of $d(n)/n$ follows a straight line as a function of n.

If two different recognition sites S_1 and S_2 appear within each monomer with the probability P_1 and P_2, double digestion will generate fragments with the length equal to an integer multiple of the monomer length and fragments of integer length elongated for distances S_1S_2 and S_2S_1. The probability of their appearance is:

$$P_1^2(1-P_1)^{n-1}(1-P_2)^n \quad \text{and} \quad P_2^2(1-P_1)^n(1-P_2)^{n-1} \quad (3)$$

for ends with S_1, or S_2 sites, respectively. Furthermore, if the segment ends are different, the probability of appearance is:

$$P_1P_2(1-P_1)^n(1-P_2)^n \quad (4)$$

in both cases, when the segment starts with S_1 and finishes with S_2 and when it starts with S_2 and finishes with S_1.

We designed a numerical procedure in which P_1 and P_2 were looked for as a solution of a system of nonlinear equations with two unknowns. The number of the equations is equal to the number of bands detected on gel for the particular double digestion experiment. We are able to detect four to five bands and each band gives rise to one equation originating from either equations (3) or (4). In order to find such values of P_1 and P_2 which would satisfy all the equations we tried many P_1, P_2 pairs and looked for that pair which would satisfy the equations to the highest extent. The trial P_1, P_2 pairs were generated as all possible combinations of the values taken from the series 0·01, 0·02, ..., 0·99, 1·00. It turns out that the above mentioned procedure provides us with a unique solution of sufficient accuracy.

3. Results

(i) *Restriction sites in* T. molitor *satellite DNA*

The consensus sequence for *T. molitor* satellite DNA was derived from sequences of randomly cloned 54 monomers, as a succession of nucleotides occurring most frequently in the analysed sample (Ugarković *et al.* 1989; Davis & Wyatt, 1989). Because of sequence variations, monomers can be recognized according to specific restriction sites. For our purpose it is meaningful to divide restriction enzymes into two groups: in the type A we include enzymes whose recognition sites are present in the consensus sequence i.e. in the majority of satellite monomers and in

https://doi.org/10.1017/S0016672300030615 Published online by Cambridge University Press

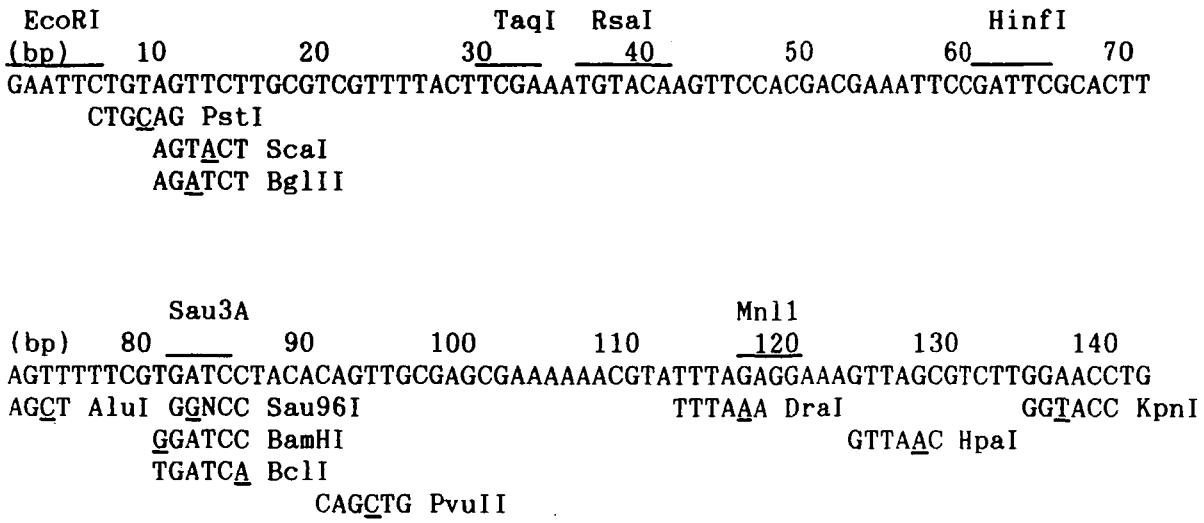

Fig. 1. Consensus sequence of *T. molitor* satellite DNA (Ugarković *et al.* 1989; Davis & Wyatt, 1989). Restriction enzymes with recognition site in the consensus are listed above the sequence (type A). Recognition sequences for enzymes whose restriction site can be generated with a single point mutation (type B) are presented under the consensus sequence. The nucleotide necessary for creating a restriction site is underlined.

the type B enzymes whose recognition sites can be generated from the consensus sequence with a single defined point mutation. The electrophoretically separated *T. molitor* DNA, digested with the two types of enzymes mentioned above, provides us with important information about sequential diversity of tandemly repeated monomers. We have been working with 5 type A and 11 type B enzymes (Fig. 1).

(ii) *Digestion with type B restriction enzymes*

The single nucleotide change necessary for the formation of the restriction site for type B enzymes can occur at only one position along the satellite monomer. Digestion of *T. molitor* genomic DNA with *Hpa* I produces a ladder of 142 bp multimers, clearly visible after gel electrophoresis on ethidium – bromide stained agarose gels. These bands hybridize strongly with cloned *T. molitor* satellite DNA. On the contrary, in the digests of *Sau*96, *Alu* I, *Dra* I, *Pvu* II, *Pst* I, *Kpn* I, *Bcl* I and *Bam*H I, a weak ladder of 142 bp multimers is visible only after hybridization. Finally, ladder pattern could not be detected at all when *T. molitor* DNA was digested with *Bgl* II or *Sca* I (Fig. 2). A different intensity or absence of bands obtained after digestion clearly shows a different abundance of particular monomer variants in the satellite. Bands of maximum intensity are positioned in the ladder according to the frequency and distribution of a particular restriction site, reflecting the organization of monomer variants within the satellite (Fig. 2). When a monomer variant is abundant, the most intense bands in the ladder are those belonging to shorter multimer length (e.g. after *Hpa* I digestion). On the contrary, when monomer variants are rare, as after *Alu* I digestion, the most intense bands in the ladder are shifted towards longer multimers. In that case the majority of hybridizing material is in almost undigested fragments of genomic DNA. This distribution is possible if different sequence variants are randomly interspersed within the satellite, and not organized in defined blocks. If a sequence variant is rare, the probability that identical variants can be found in proximity to each other is low, resulting in a ladder of relatively long multimers.

(iii) *Quantitative analysis of type B digestion profiles*

Distribution of satellite DNA in the *Hpa* I ladder is shown in Fig. 3. The centre of mass is in the range of first 5 bands containing about 80% of the DNA material in the ladder (Fig. 3*a*). Hybridization (Fig. 2) shows that all *T. molitor* satellite DNA is included in the *Hpa* I ladder, with no undigested fragments which could be interpreted as regions of satellite DNA without *Hpa* I monomer variant. Abundance of *Hpa* I restriction sites in the satellite DNA was determined from densitometric measurements of digestion profile (Fig. 3*a*). The plot $\log(d(n)/n)$ versus n is linear and equation 1 is fulfilled with the probability of occurrence of this sequence variant $P = 0{\cdot}5$ (Fig. 4), suggesting that the appearance of analysed restriction site within the satellite is random and that there is 50% probability that the position 128 in the sequence of satellite monomer is occupied with adenine, generating a *Hpa* I restriction site. According to this we can conclude that sequence variants, represented with *Hpa* I restriction site, are distributed randomly among other sequence variants within the entire satellite DNA mass in the *T. molitor* genome.

The additional information about the frequency and correlation in the distribution of restriction sites can be revealed by double digestion experiments. We performed several double digestion experiments and

https://doi.org/10.1017/S0016672300030615 Published online by Cambridge University Press

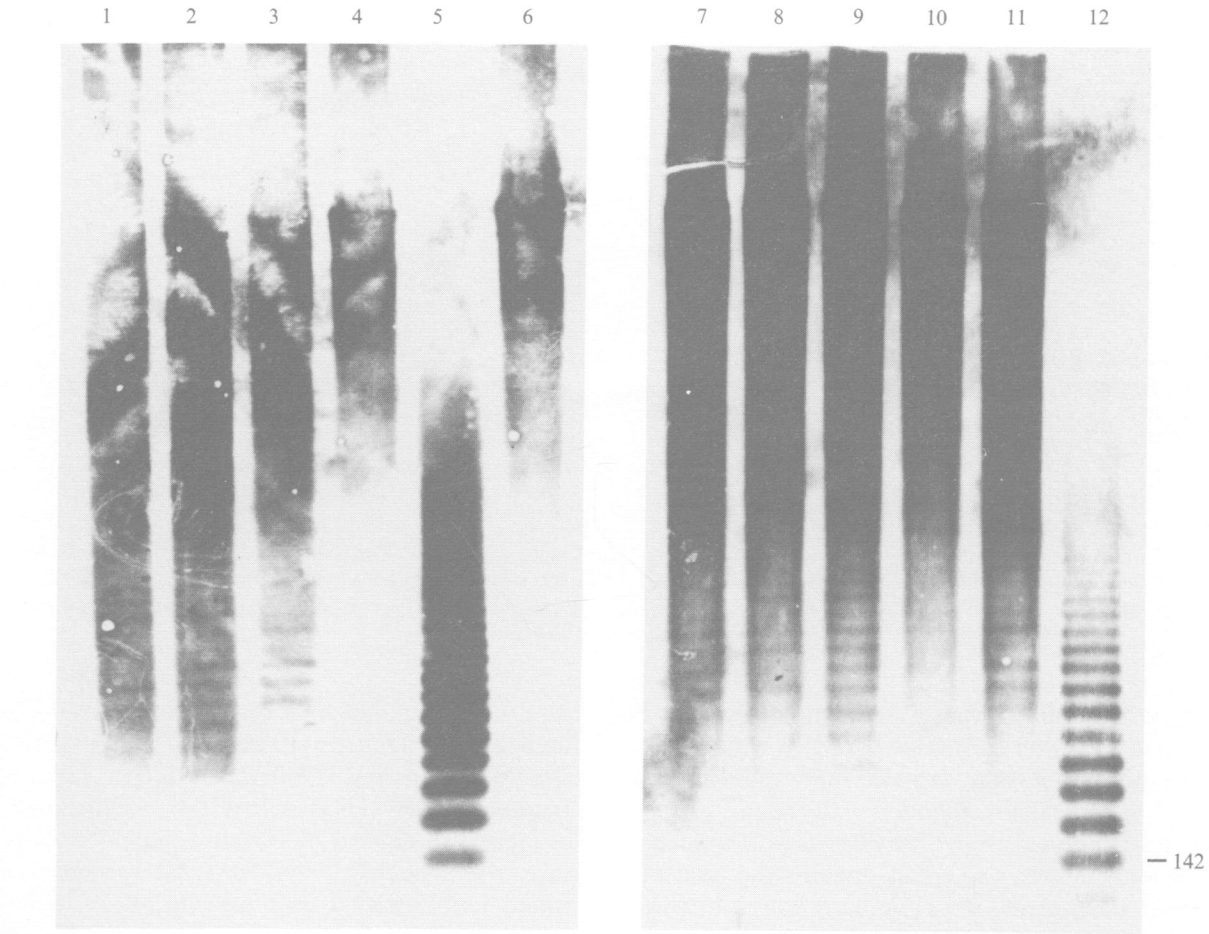

Fig. 2. Southern blots of *T. molitor* satellite DNA digested with restriction enzymes from type B: lane 1, *Sau*96 I; 2, *Alu* I; 3, *Dra* I; 4, *Bgl* II; 5, *Hpa* I; 6, *Sca* I; 7, *Pvu* II; 8, *Pst* I; 9, *Kpn* I; 10, *Bcl* I; 11, *Bam*H I and lane 12, partial digestion of *T. molitor* satellite DNA with *Eco*R I as size marker. Each lane contains 5 μg of DNA, except lane 12, which contains 50 ng. Cloned *T. molitor* satellite DNA was used as a hybridization probe. Only *Hpa* I profile is visible even before hybridization. The amount of variants with restriction site generating weak ladder, visible only after hybridization, is estimated to be under 1% of satellite monomers.

the following were treated quantitatively: *Hpa* I/*Rsa* I; *Hpa* I/*Hinf* I and *Hpa* I/*Sau*3A (Fig. 3*b*, *c*). The fragments produced as a result of double digestion are of the expected length, in accordance with the position of the restriction sites (Fig. 1). Random distribution would be corroborated if (i) the solution of the system of equations mentioned in Materials and methods gives a unique value of P_1 and P_2 and (ii) the values of P_1 and P_2 match with the corresponding values obtained from single digestion experiments. Analysis of the results shows that both criteria (i) and (ii) are fulfilled. Let us discuss the *Hpa* I/*Rsa* I example: the five peaks belonging to the fragments of different length as depicted in Fig. 3*c* have areas 0·23, 0·51, 0·048, 0·051 and 0·073, from left to right, respectively. Described numerical procedure approaches closest to the above mentioned values with $P_1 = 0{\cdot}45$, $P_2 = 0{\cdot}9$ what should be compared with the single digestion results $P_1 = 0{\cdot}5$ and $P_2 = 0{\cdot}92$ (sections iii and iv, respectively). The five areas which emerge from equations (3) and (4) with $P_1 = 0{\cdot}45$ and $P_2 = 0{\cdot}9$ are 0·25, 0·47, 0·031, 0·036 and 0·051. Similar values were obtained from the two other double digestion experiments. In all cases the P_1 and P_2 values are slightly below the values (up to 10%) obtained by single digestion experiments. These slight disagreements in probabilities, as well as in area values, are probably due to low resolution of double digestion profiles. Furthermore, quantification of double digestion profiles can give more accurate information about distribution of restriction sites if not so common restriction enzymes (e.g. type A) are used. In the case of *T. molitor* satellite DNA we could not find any other enzyme, except *Hpa* I, appearing with comparable frequency which would be more suitable for double digestion analysis.

Distribution of satellite DNA in the ladder obtained after digestion with some other type B restriction enzymes (Fig. 2) can be visualized only after hybridization, as described in the section (ii). Because the signals are very weak and higher bands obscured with the growing background of undigested material, it is difficult to determine the digested amount in the ladder accurately. The frequency of *Sau*96, *Alu* I, *Dra* I, *Pvu* II, *Pst* I, *Kpn* I, *Bcl* I and *Bam*H I restriction sites is estimated to be 0·5–1%. Linear

https://doi.org/10.1017/S0016672300030615 Published online by Cambridge University Press

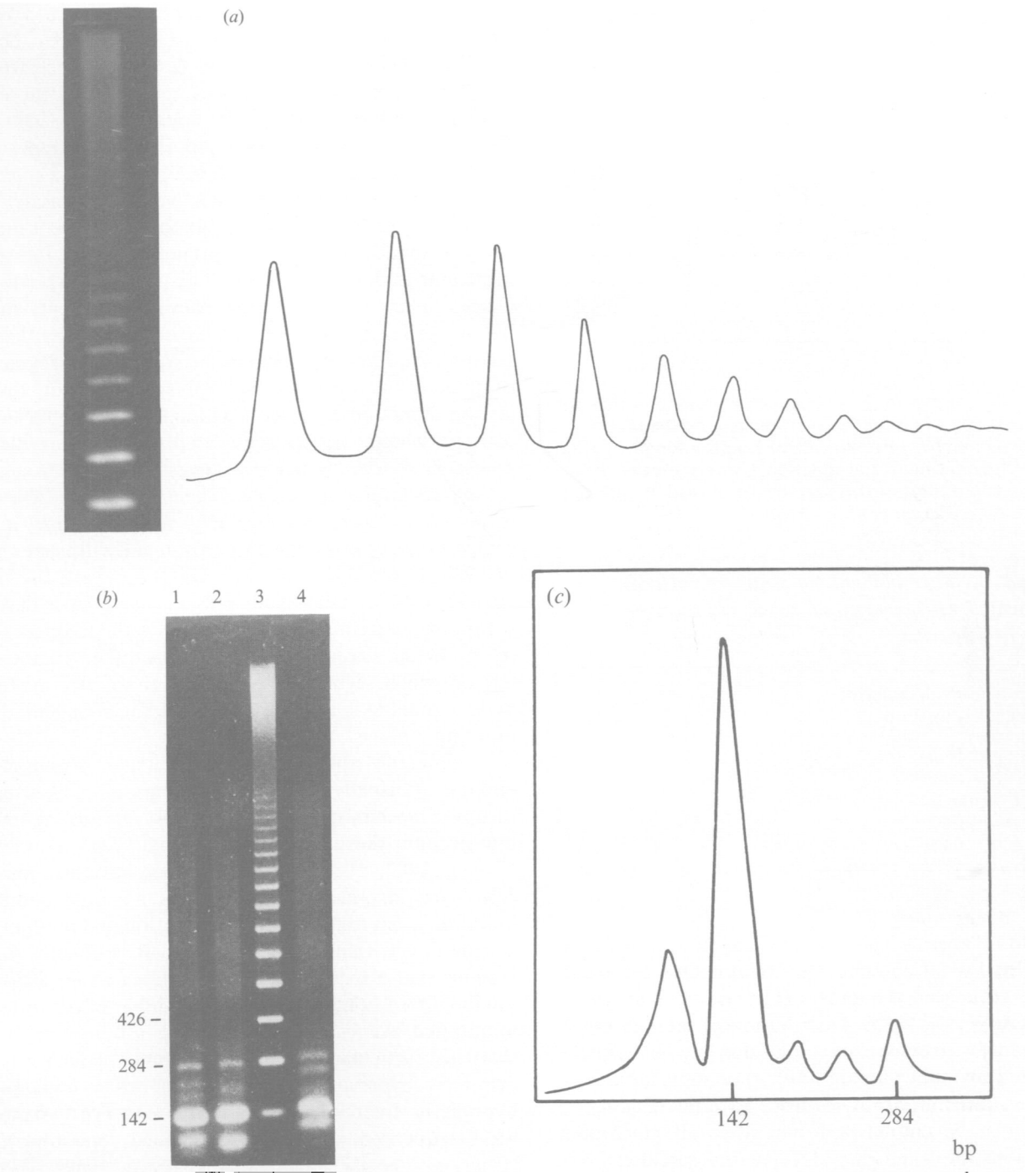

Fig. 3. Distribution of satellite DNA in *Hpa* I ladder. (*a*) Electrophoresis of genomic DNA digested with *Hpa* I on ethidium bromide stained agarose gel and densitometric profile. (*b*) Double digestions of *T. molitor* genomic DNA with *Hpa* I and: *Rsa* I (lane 1); *Hinf* I (lane 2) and *Sau*3A (lane 4). In the lane 3 is genomic DNA partially digested with *Eco*R I, as size marker. The abundance of monomer variants is determined from densitometric measurements of single and double digestions. (*c*) Densitogram of one representative (*Hpa* I + *Rsa* I) double digestion profile.

relationship in the plot $\log(d(n)/n)$ versus n with $P \sim 0{\cdot}01$ indicates random distribution of sequence variants, as concluded for the distribution of *Hpa* I variant.

(iv) *Digestion with type A restriction enzymes*

Digestion with 5 restriction enzymes with recognition sites in the consensus sequence (type A) produces one prominent band belonging to the satellite monomer and a few weak bands corresponding to multimers, with rapidly decreasing amounts of material (not shown). This profile is characteristic for recognition sites present in almost all satellite monomers. The digestion profile obtained shows that the frequency of mutated variants is rather low. Relative amounts of observed multimers are slightly different in each digestion, indicating different abundance of sequence

https://doi.org/10.1017/S0016672300030615 Published online by Cambridge University Press

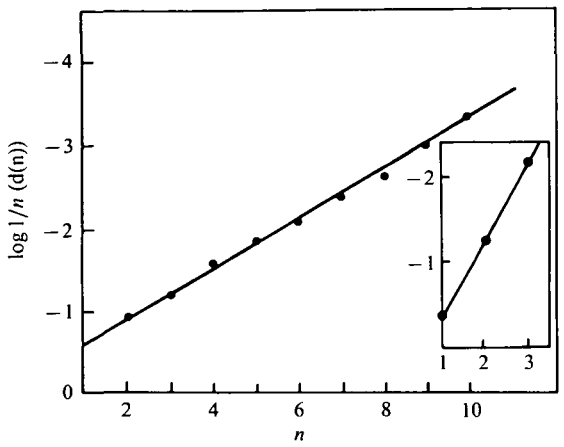

Fig. 4. Test of distribution of sequence variants within the satellite. Linear plot $\log(d(n)/n)$ versus n, where $d(n)/n$ represents number of fragments of length n obtained after *Hpa* I digestion fulfils the equation 1 with the value $P = 0{\cdot}5$. Inset: plot $\log(d(n)/n)$ versus n for digestion with one of type A restriction enzyme, *Eco*R I.

Table 1. *Relative amounts of sequence variants represented with recognition sequences for type A restriction enzymes*

Enzyme	Probability of digestion
Sau3A	0·96
Hinf I	0·95
*Eco*R I	0·92
Rsa I	0·92
Mnl I	0·86

variants (Table 1). This result is in agreement with the expected loss of consensus restriction sites according to the sequenced monomers (Ugarković *et al.* 1989; Davis & Wyatt, 1989). Distribution of satellite DNA fragments with consensus recognition sites gives linear relationship in the plot log $(d(n)/n)$ versus n, indicating random distribution of sequence variants (Fig. 4).

It must be emphasized that after all performed digestions the fragments of *T. molitor* satellite DNA are exact multimers of 142 bp, without any irregular fragment observed. Also, defined higher order repeating structure, which would indicate the existence of defined satellite subfamilies, could not be observed.

4. Discussion

Cloning and sequencing of monomers can reveal only limited information about the polymorphism of sequence variants in satellite DNA. In this way one can only detect a certain number of variants, but can get no information about their sequential arrangements unless long satellite fragments are cloned and sequenced. Sequence analysis of *T. molitor* cloned satellite dimers shows the same difference between adjacent monomers as within randomly cloned satellite sequences (Davis & Wyatt, 1989). On the contrary, sequence analysis of tsetse fly repetitive DNA revealed that contiguous repeats and those cloned at random are divergent (Trick & Dover, 1984). Recently, a new method for mapping variant repeat units in individual arrays was developed (Jeffreys *et al.* 1991). Sequence variation in hypervariable minisatellites is assayed by PCR amplification on alleles, or on total genomic DNA, using primers specific for a particular variant repeat unit. The individual specificity in interspersion pattern of variant repeats in human MS32 minisatellite alleles is revealed by this method. Another approach in the study of the degree of monomer sequences polymorphism and the organization of monomer variants within the satellite DNA can be obtained with simple restriction enzyme analysis and quantification of digestion products.

An outstanding feature of *T. molitor* satellite DNA is its abundance, homogeneity and presence in centromeric regions of all chromosomes (Petitpierre *et al.* 1988; Ugarković *et al.* 1989; Davis & Wyatt, 1989; Juan *et al.* 1990). Our results present the evidence that within the entire mass of satellite DNA the frequency of monomer variants is different. Another characteristic which seems to be of general validity in *T. molitor* satellite is the randomness in their sequential ordering.

A different abundance of particular sequence variants is usually explained in terms of a bias in turnover mechanism favouring a new variant, taking into account the time of a mutational event, as well (Dover, 1982, 1986, 1989). Sequence variants, like *Hpa* I for instance, could arise as a single point mutation event followed by amplification and random distribution within a whole satellite. It is difficult to imagine that a sequence variant building about 50% satellite DNA could result from a number of separate mutational events. In that case one should assume that this particular site in the monomer sequence is extremely hypervariable, which is a rather unlikely hypothesis. Interestingly, no *Hpa* I sequence variant was sequenced among 54 cloned monomers (Ugarković *et al.* 1989; Davis & Wyatt, 1989). The reasons for this cloning artifact are not clear, but similar preferences can obscure determination of consensus sequence.

Clustering of sequence variants was observed in repetitive DNA families of tsetse fly (Trick & Dover, 1984), as well as in satellite DNAs from Guinea pig (Altenburger *et al.* 1977), mouse (Horz & Zachau, 1977), tilapia (Wright, 1989) and human (Warburton & Willard, 1990). Highly homologous satellite DNA is present in several species of *Ctenomys*, with large interspecies differences in the amount of monomer variants. According to digestion patterns, variants seem to be randomly distributed (Rossi *et al.* 1990). Random distribution within the satellite could be explained as a result of sequence homogenization in

the process of molecular drive (Dover, 1982). The process of spread (molecular drive) is rapid relative to the mutation rate. It is assumed that the rate of fixation of variants is characteristic for each nongenic family (Strachan *et al.* 1985). Because monomer variants are randomly distributed in total *T. molitor* satellite DNA they should be spread on all chromosomes, as well. Such a distribution can be a consequence of gene conversion and unequal exchange which play major roles in the spreading of satellite variant throughout the family (Strachan *et al.* 1985). It seems that in the case of *T. molitor*, the rates of homogenization among different chromosomes and within chromosomes are similar and high in relation to many other satellite DNAs which contain clustered monomer variants.

This work was supported by the Research Fund of Republic Croatia, project No. 1-08-269 and by The Ministry of Science and Technology of Republic Slovenia.

References

Altenburger, W., Horz, W. & Zachau, H. G. (1977). Comparative analysis of three guinea pig satellite DNAs by restriction nucleases. *European Journal of Biochemistry* **73**, 393–400.

Choo, K. H., Vissel, B., Nagy, A., Earle, E. & Kalitsis, P. (1991). A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Research* **19**, 1179–1182.

Davis, C. A. & Wyatt, G. R. (1989). Distribution and sequence homogeneity of an abundant satellite DNA in the beetle, *Tenebrio molitor*. *Nucleic Acids Research* **17**, 5579–5586.

Dover, G. (1982). Molecular drive: a cohesive model of species evolution. *Nature* **299**, 111–117.

Dover, G. A. (1986). Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends in Genetics* **168**, 159–165.

Dover, G. A. (1989). Linkage disequilibrium and molecular drive in the rDNA gene family. *Genetics* **122**, 249–252.

Durfy, S. J. (1990). Concerted evolution of primate alpha satellite DNA. Evidence for an ancestral sequence shared by gorilla and human × chromosome alpha satellite. *Journal of Molecular Biology* **216**, 555–566.

Fanning, T. G. (1987). Origin and evolution of a major feline satellite DNA. *Journal of Molecular Biology* **197**, 627–634.

Horz, W. & Zachau, H. G. (1977). Characterization of distinct segments in mouse satellite DNA by restriction nucleases. *European Journal of Biochemistry* **73**, 383–392.

Jeffreys, A. J., MacLeod, A., Tamaki, K., Neil, D. L. & Monckton, D. G. (1991). Minisatellite repeat coding as a digital approach to DNA typing. *Nature* **354**, 204–209.

Juan, C., Gosalvez, J. & Petitpierre, E. (1990). Improving beetle karyotype analysis: restriction endonuclease banding of *Tenebrio molitor* chromosomes. *Heredity* **65**, 157–162.

Juan, C. & Petitpierre, E. (1989). C-banding and DNA content in seven species of *Tenebrionidae* (Coleoptera). *Genome* **32**, 834–839.

Lohe, A. R. & Brutlag, D. L. (1986). Multiplicity of satellite DNA sequences in *Drosophila melanogaster*. Proceedings of the National Academy of Science USA **83**, 696–700.

Petitpierre, E., Gatewood, J. M. & Schmid, C. W. (1988). Satellite DNA from the beetle *Tenebrio molitor*. *Experientia* **44**, 498–499.

Rossi, M. S., Reig, O. A. & Zorzopulos, J. (1990). Evidence for rolling-circle replication in a major satellite DNA from the South American rodents of the genus *Ctenomys*. *Molecular Biology and Evolution* **7**, 340–350.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbour Laboratory.

Southern, E. M. (1975). Long range periodicities in mouse satellite DNA. *Journal of Molecular Biology* **94**, 51–69.

Strachan, T., Webb, D., Dover, G. A. (1985). Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO Journal* **4**, 1701–1708.

Trick, M. & Dover, G. A. (1984). Unexpectedly slow homogenization within a repetitive DNA family shared between two subspecies of tsetse fly. *Journal of Molecular Evolution* **20**, 322–329.

Ugarković, Đ., Plohl, M. & Gamulin, V. (1989). Sequence variability of satellite DNA from the mealworm *Tenebrio molitor*. *Gene* **83**, 181–183.

Warburton, P. E., Willard, H. F. (1990). Genomic analysis of sequence variation in tandemly repeated DNA. *Journal of Molecular Biology* **216**, 3–16.

Wright, J. M. (1989). Nucleotide sequence, genomic organization and evolution of a major repetitive DNA family in tilapia (*Oreochromis mossambicus/hornorum*). *Nucleic Acids Research* **17**, 11–12.