Evolution of *Tribolium madens* (Insecta, Coleoptera) Satellite DNA Through DNA Inversion and Insertion

Đurdica Ugarković, Sonja Durajlija, Miroslav Plohl

Department of Molecular Genetics, Ruđer Bošković Institute, P.O. Box 1016, Bijenička 54, 41000 Zagreb, Croatia

Abstract. Two different satellite DNAs from tenebrionid species *Tribolium madens* (Insecta, Coleoptera) have been detected, cloned, and sequenced. Satellite I comprises 30% of the genome; it has a monomer size of 225 bp and a high A + T content of 74%. Satellite II, with a monomer size of 711 bp and A + T content of 70%, is less abundant, making 4% of the total DNA. Sequence variability of the monomers relative to consensus sequence is 4.1% and 1.2% for satellite I and II, respectively. Both satellites are localized in the heterochromatic regions of all chromosomes. A search for internal motifs showed that both satellites contain a related subsequences, about 100 bp long. The creation of satellite I monomer is explained by duplication of the basic subunit, followed by subsequent divergence by single point mutations, deletions, and gene conversion. Inversion of the subsequence in addition to its duplication has occurred in satellite II. The result of this inversion is possible formation of a long, stable dyad structure. The 408-bp sequence, inserted within satellite II monomer, shares no similarity with a basic subunit. Frequent direct repeats found within the inserted sequence point to its evolution by duplication of shorter motifs. It is proposed that both satellites have been derived from a common ancestral sequence whose duplication played a major role in the formation of satellite I monomer, while insertion of a new sequence together with inversion of an ancestral one induced the occurrence of satellite II.

Introduction

Insect species from family Tenebrionidae (Coleoptera) have a substantial amount of pericentromeric heterochromatin located on all chromosomes of the complement. The heterochromatic blocks preferentially contain satellite DNAs which make up 30–50% of the genome. Most of the species have a single, highly abundant satellite DNA of a unique sequence which is usually species specific (Petitpierre et al. 1995). Depending on the size of the basic repeating unit, tenebrionid satellites can be classified into two groups. The first group comprises seven different satellite DNAs of unique sequences with monomer lengths between 100 and 160 bp, while in another group a monomer sequence is about 350 bp long (Ugarković et al. 1995).

Analysis of satellites from related tenebrionid species reveals that different mutational processes are involved in their evolution. In some of the species like *Tenebrio molitor, Palorus ratzeburgii, Tribolium confusum,* and *Tribolium freemani* where only one type of highly abundant satellite DNA is found, the sequence variability is very low—up to 2%—due mainly to single base substitutions (Ugarković et al. 1989, 1992; Plohl et al. 1993; Juan et al. 1993). However, satellite DNA from *Alphitobius diaperinus* exhibits much higher divergence (Plohl and Ugarković 1994a). It is composed of three related satellite subunits organized in three higher-order satellite subfamilies. Comparison of satellite subunits re-

veals that processes of gene conversion, replication slippage, and recombination are involved in the evolution of this satellite. The presence of a long direct repeat within the satellite monomer of tenebrionid species *Misolampus goudoti* indicates that unequal crossing over has occurred both within and between repeating units (Pons et al. 1993).

Satellite DNAs often have higher-order structures either in the form of dyad structures, due to inverted repeats (Fowler and Skinner 1985; Bigot et al. 1990), or in the form of solenoid structures, induced by curvature of the DNA helix axis (Radić et al. 1987; Martinez-Balbas 1990). It is proposed that these structures perform a coding function in the binding of specific proteins and in this way influence condensation of heterochromatin (Vogt 1992). Three different satellites from tenebrionid species *T. molitor, Tenebrio obscurus,* and *P. ratzeburgii* have the same monomer length of 142 bp and form tertiary structure of very similar geometry in the form of a left-handed superhelix (Plohl et al. 1990; Ugarković et al. 1992; Plohl and Ugarković 1994b). It is suggested that conservation of such structural characteristics despite sequence divergence points to their importance for heterochromatin formation (Plohl and Ugarković 1994b).

In this paper two satellite DNAs from tenebrionid species *Tribolium madens* are characterized with respect to the nucleotide sequence, organization, and amount. Sequence variability and mutational processes acting on their evolution are studied. In addition we compare two satellites in order to determine what relationship they have with each other at the level of nucleotide sequence, higher-order structure, or internal organization. Such comparisons could give us some information concerning possible constraints operating during the evolution of satellite DNAs within a species.

Materials and Methods

Cloning and Sequence Analysis. Genomic DNA was isolated from larvae and adult insects by standard procedure (Sambrook et al. 1989) and restriction digestions were performed according to the manufacturer's instructions. After electrophoresis on 1.2% agarose gels, DNA fragments of approximately 200 bp and 500 bp after *Cla*I digestion, of approx. 220 bp after *Alu*I digestion, as well as fragments of approx. 190 and 110 bp after *Hinf*I digestion were electroeluted, filled in by Klenow DNA polymerase, and blunt end ligated with *Hind*II-linearized pUC18 plasmid vector. *E. coli* DH5 cells were transformed and cloned DNA was analyzed by hybridization under high-stringency conditions (68°C, washing at the same temperature in 0.1 × SSC; 0.1% SDS). A portion of eluted *Cla*I fragments of approx. 200 bp and of *Hinf*I fragments was labeled with digoxigenin—dUTP by the random primer labeling method (DNA Labeling and Detection Kit, Boehringer Mannheim) and used as a hybridization probe; a 500-bp *Cla*I fragment and a 220-bp *Alu*I band were detected and cloned using as a hybridization probe the 203-bp-long positive fragment obtained from *Cla*I band. Recombinants giving positive signal were directly sequenced by the dideoxy sequencing method (Sanger et al. 1977). For the Southern analysis of total genomic DNA, 5 μg of *T. madens* DNA was digested with different restriction enzymes (REs) and blotted on Boehringer Mannheim nylon

membrane, positively charged. Satellite clones labeled with digoxigenin were used as a hybridization probe under highly stringent conditions as described previously.

Nucleotide sequences were analyzed and compared using MicroGenie (Beckmann) and PCGene programs. GenBank was screened using on-line programs (Benson et al. 1993).

Fluorescent in Situ Hybridization. The chromosome spreads were obtained from gonads of male adults according to Juan et al. (1991). Cloned satellites were labeled by nick translation either with biotin-16-dUTP (satellite II) or with digoxigenin-11-dUTP (satellite I). Hybridization conditions were the following: 60% formamide, 2 × SSC, 0.5 μg/μl denatured salmon sperm DNA, 0.5 μg yeast RNA and 5 ng/μl of each probe, at 37°C. Immunological detection of satellite I was achieved by anti-digoxigenin-rhodamine, Fab fragments (Boehringer Mannheim), while for the detection of satellite II an avidin-FITC system with one step of amplification was used (Pinkel et al. 1986). Chromosomes were stained with fluorochrome DAPI, 0.2 μg/ml in McIlvaine's phosphate buffer, for 15 min (Sumner 1990) and photographed with fluorescence optics in an Opton photomicroscope.

Results

Detection, Cloning, and Sequencing of T. madens Satellite DNA

T. madens DNA was digested with different restriction enzymes (Fig. 1A) in order to detect distinct bands characteristic of repetitive DNA. Digestion with *Cla*I and *Taq*I reveals a strong band of about 200 bp and a slighter ladder of dimer, trimer, and higher multimers, pointing to the tandem arrangement of repetitive units characteristic for satellite DNAs. In *Hinf*I digestion bands of about 190 and 110 bp are visible as well as shorter bands which could result from the presence of more than one *Hinf*I restriction site in the satellite monomer. A similar case exists for *Alu*I, giving a complex ladder, due probably to a mixture of two different satellite DNAs or due to the presence of more than one *Alu*I recognition site in a particular satellite monomer.

Two bands from *Hinf*I digest approximately 190 and 110 bp, are cloned separately as well as a *Cla*I band of 200 bp. Hybridization analysis of fragments cloned from an approx. 200-bp band of genomic DNA revealed, after excising from *Cla*I clones, positive inserts of two different size: six inserts of approx. 220 bp and three inserts having a size of approx. 200 bp. Sequencing results, summarized in Fig. 2A and B, reveal the following: six *Cla*I inserts (MAD1C1-6) exhibit high sequence similarity, four of the monomers are 225 bp long, and another two are 222 and 218 bp long due to nucleotide deletions. The additional clone MAD1C7 also overlaps the sequences of the above-mentioned clones but is shorter because of the occurrence of an additional *Cla*I site at position 124. Three remaining *Cla*I positives (Fig. 2B, MAD2C1-3) are 204 or 203 bp long and exhibit high mutual similarity, but cannot be overlapped with *Cla*I
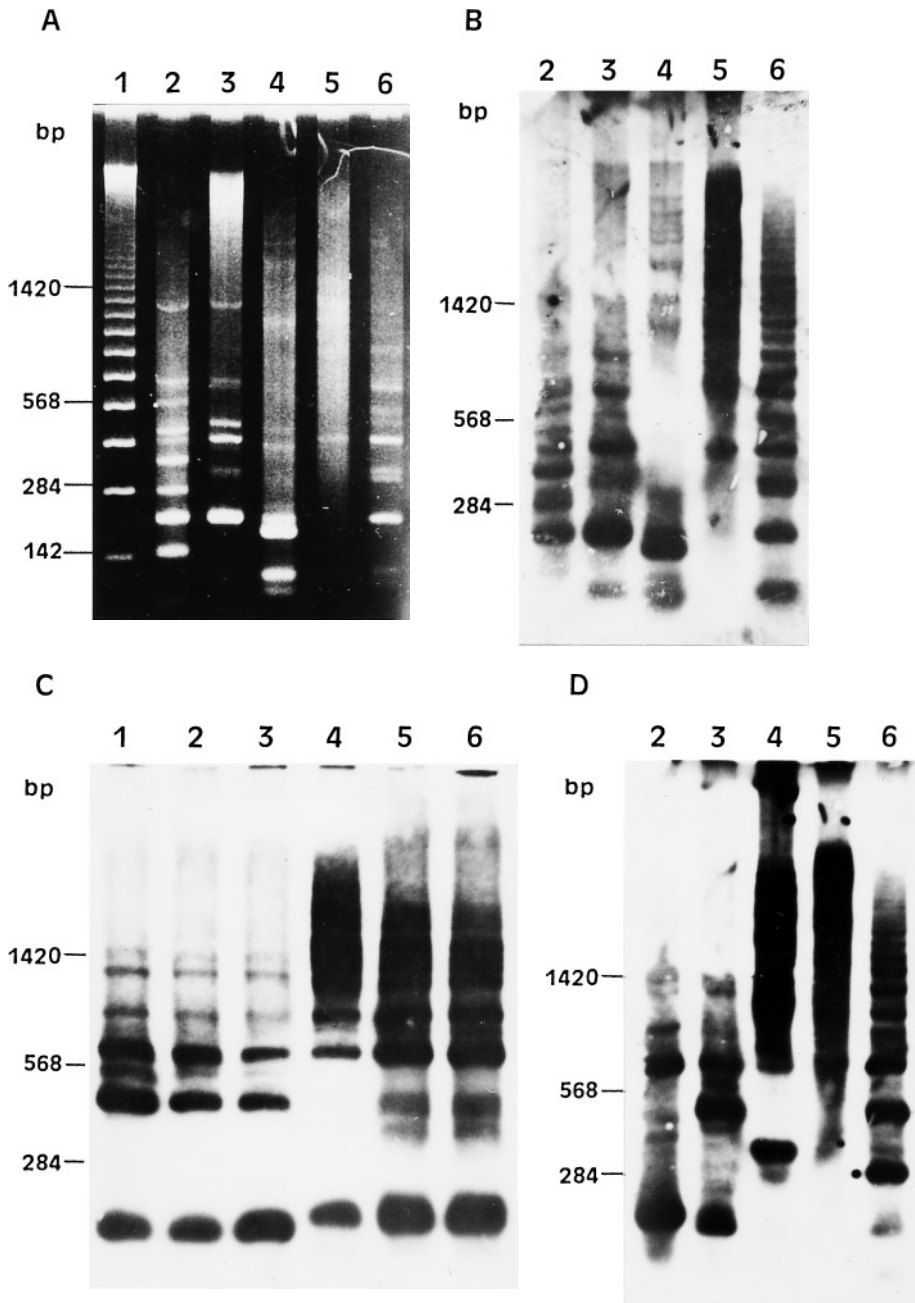
Fig. 1. **A** Electrophoretic separation of *T. madens* genomic DNA on 1.2% agarose gel; 5 μg of DNA was digested with *Alu*I (*lane 2*), *Cla*I (*3*), *Hinf*I (*4*), *Sau*3A (*5*), and *Taq*I (*6*). *T. molitor* satellite DNA, partially digested with *Eco*RI, is used as a size marker (*lane 1*). **B** The Southern blot of gel shown in A after hybridization with cloned *T. madens* satellite I. **C** The Southern blot obtained by partial digestion of 5 μg of *T. madens* DNA with 5 U of *Cla*I (*lanes 1–3*) and with 5 U of *Alu*I (*4–6*) for 30, 60, and 120 min, respectively, and subsequent hybridization with cloned satellite II probe. **D** Southern blot of gel shown in **A** after hybridization with cloned *T. madens* satellite II.

Fig. 2. **A** Nucleotide sequences of 13 clones obtained by random cloning of *T. madens* satellite I and from them derived consensus sequence of satellite I monomer. In satellite clones only differences from consensus sequence are indicated, while missing parts are marked as *dots*. **B** Consensus sequence of satellite II monomer derived from seven randomly cloned satellite II fragments. The most prominent direct repeats found in satellite II monomer sequence are indicated with *arrows*. The direct repeats marked with *2* and *3* are 17 and 13 bp long and have a perfect similarity of 100% within each pair. Direct repeats marked with *1* are 23 bp long with a mutual similarity of 75% while the similarity of 18-bp-long direct repeats marked with *4* is 78%. The accession numbers of satellite I and II are U30598 and U30599, respectively.

A

Satellite I

```
                    10        20        30        40        50        60-       70        80        90       100       110       120
          ATCGATTTGGCTCAATTTTAAGTTGGGAATCAAGC TAAATCAAAGTATTTCAACCAAATACTGGCTCTAATAAGAGTTTTCTTCTAAAATAAAGCAATTTGAGCTATTTTCAATCCAATT
MAD1C1    ------------G--G---------C-------------------------------------------------------------------A----------------
MAD1C2    --------G-----------------------------------------------------------A----------------------------T-----
MAD1C3    -----AG------------------------------T--------------------C---------------------------C----------
MAD1C4    --------------------------A---------------------------------C------C--------------------------------
MAD1C5    -----------G-------------00000--------------TT--------T------G------CT------A-----------T--0--
MAD1C6    ------------G-----------------------0---------G---0---------G-G-------------A-----------T---
MAD1C7    --------A----G------A--A-C---AA------------A-----------AA-0--------------------------0-A-
MAD1H1    .............................---C-----------------------G----------A--------------------T---
MAD1H2    .............................---A----------------TT-------------------------G------
MAD1H3    .............................---T---------------T-----------------------------A---A---
MAD1H4    .............................----------------TT-------------G------CT------A-----------T---
MAD1H5    .............................--T--------------C-A------G---A-------------A----------T---
MAD1H6    .............................--G-----------------------T------------------A----

                   130       140       150       160       170       180       190       200       210       220
          TTGCTCGATTTTAATACAGAATTAATTCAAAACTAAT GTATTTTATTAGTATAATTGCTTCAATCTTAATTTTAGGCTGAAATAAAGCAATTTGTGATATTTTGA
MAD1C1    ----------------T------G-C----------A-----------------------------------------------------
MAD1C2    ------------------------T----------------------------------------------A-------------
MAD1C3    ------------------AT------G--A---------------------TT----------------
MAD1C4    --C------------------AT------G----------------------TT----------------C--
MAD1C5    ------------T---------C----ATA00-----------T------A----T-----------
MAD1C6    ------------T---------C----A0-----C----------------T--CATTAT-AT-TA-T--0-----AC-CA-C---C-
MAD1C7    --AA---........................................................
MAD1H1    ------------------------------A-C----------------------T------------.
MAD1H2    --AA------------------A-----------------------A---A---T----------
MAD1H3    --AT------------------A--T------------------T------------T---------C--..
MAD1H4    -------------T---........................................
MAD1H5    -------------T---........................................
MAD1H6    -------------T---........................................
```

B

Satellite II

```
                    10        20        30        40        50        60        70        80        90       100       110
          ATCGATTTGGCTCGATTTTAAATCAGAAATCGCTCTAAATCAATTGTATTTCAAC GATGAAATAACTGTTTTCAGTCAAAATAAAGCAATTTGAGCTATTTCCAATTCCA
MAD2C1    ---------------------------------------------------------------------------------------------------------
MAD2C2    ---------------------------------------------------------------------------------------------------------
MAD2C3    ----------------------------------------0---------C----------A-------------------------------------------
MAD2A1    .............................................................----------------------------
MAD2A2    .............................................................----------------------------
MAD2A3    .............................................................----------------------------
MAD2C4    ............................................................................................

                   120       130       140       150       160       170       180       190       200       210       220
          TTTTGCTTTATTTTTGCAGAAAACT GTTATTTAAGCCGTCTATTTTGTTGAAATACACTGATTTGGTT CAATTACTGGTTTAAAATTGAGCCAAATCGATGAGAAGTTGAG
MAD2C1    -----------------------A----A---------G-----------------------A-   ....................
MAD2C2    -----G---------------------------C--------A---------T--------------A--------------   ....................
MAD2C3    -----------------------------------------A---------------A-T----A----------A-----   ....................
MAD2A1    -----------------------------------------------------------------------------------------CA-----
MAD2A2    -C-----------------------------A----------------------------------------------------
MAD2A3    ---------------------------------A-----------------------------T----------------------
MAD2C4    ..........................................................------------------C-

                                         1
                   230       240       250       260       270       280       290       300       310       320       330
          TAAATTGTATTATAGTGGCTGAAAACGGGTATTTAAACACGAATTGTGTTGAAATGCATTGATGTGTGATGTTTTCTGGCATAAAATCCAAGAAAAGCTGTTGGAATATC
MAD2C1    .......................................................................................................
MAD2C2    .......................................................................................................
MAD2C3    .......................................................................................................
MAD2A1    ----------------------C-----------------------------------------------------
MAD2A2    ------------------------------------------------------------------------------
MAD2A3    ---------------------------------------------------T--------------------------
MAD2C4    ------T----------------C-----------------------------------

                                                                                                          2
                   340       350       360       370       380       390       400       410       420       430       440
          ATCAAATCGTCTCATTTTGTCTTGAAACCGCAATTTTTAATACTTATTTGACAAAACACATTATTTTGCAAGGTTTTCCTAAACAAAAACAACGAAATGTGTGTTTATTCG
MAD2C4    ----------------------------------------------------------------------------------------------------------
          ------>              3                4                                                                2
                   450       460       470       480       490       500       510       520       530       540       550
          TCGTTTACGGCCAAACAGCTATTTTGATTAATTTCTAATCGTTTTGGCTGAATTTTAGGCCAAAACTCGGCAATTTTGCTCGATTTGAACACAGAGAGTGTTTATTCGT
MAD2C4    ----------------------------------------------------------------------------------------------------------
          ------>              3                4
                   560       570       580       590       600       610       620       630       640       650       660
          CGTTTCTGACTTAAATAAGACGATTTGATTAATTTCCAATGCATTTGGATCACTTTTAAGCAGGAAATCAACCTAAATCAAAGTATTACAACAATATAATAGTTTAAATA
MAD2C4    ----------------------------------------------------------------------------------------------------------

                   670       680       690       700       710
          TTGGTTCACTTTCAAAATAAAGCAATTTGAGTTATTTTCAATTCAACATTG
MAD2C4    --------------------------------------------------
```

354

clones from the first group and therefore are representatives of another repetitive DNA family.

Sequence analysis was also performed on six *Hinf*I clones (Fig. 2A), three of them obtained by cloning a band approx. 190 bp long (MADH1-3) and another three by cloning a band of 110 bp long (MAD2H4-6). The sequencing results show that three longer clones overlap *Cla*I clones of the first group in the region between *Hinf*I sites at positions 27 and 224, while three shorter clones span the region between positions 27 and 139 where an additional *Hinf*I site can appear due to the single nucleotide substitution.

Cloned fragments of each group were used as hybridization probes after Southern transfer of digested *T. madens* genomic DNA (Fig. 1B–D). The hybridization was performed under highly stringent conditions, allowing annealing of sequences with 90% or higher similarity. Hybridization with the cloned probe belonging to the first group of repetitive DNA reveals a regular ladder in *Cla*I and *Taq*I digestions characteristic of satellite DNA (Fig. 1B). We designate it as *T. madens* satellite I. The ladder is composed of satellite monomer 225 bp long and multimeric bands up to the 4-mer, with decreasing intensity, characteristic of a type A digestion pattern of tandem repeats. Also, intermediate bands can be observed, resulting from an additional *Taq*I site at position 125 in the monomer sequence and possible creation of a *Cla*I site at the same position by single point mutation. Two recognition sites for *Alu*I in the satellite monomer result in a complex ladder pattern while digestion with *Hinf*I results in fragments shorter than a satellite monomer. However, *Sau*3A partially cleaves *T. madens* satellite I, producing the type B restriction enzyme pattern (Fig. 1B). Densitometric analysis of agarose gel and spot hybridization using cloned satellite monomer as a probe shows that this satellite DNA is very abundant, constituting 30% of the whole *T. madens* DNA.

Hybridization of *T. madens* DNA, partially digested with *Cla*I and *Alu*I, using a cloned 203-bp *Cla*I fragment (Fig. 2B) as a probe, reveals the pattern shown on Fig. 1C. Besides the 203-bp fragment the most prominent bands in the *Cla*I digest are about 500 and 700 bp long. In *Alu*I digestion the same band of approx. 700 bp occurs together with a ladder composed of 1,000 bp, 1,400 bp, and longer bands, indicating tandem arrangement characteristic of satellite DNA. In addition, a prominent band of about 220 bp is also observed in the *Alu*I digest after hybridization with the cloned 203-bp *Cla*I probe. In order to further characterize this satellite DNA and to define its monomeric repeating unit we have cloned a band of approx. 500 bp obtained by *Cla*I digestion and a band of approx. 220 bp obtained by *Alu*I digestion of *T. madens* DNA (Fig. 1A). For the detection of positives a cloned 203-bp *Cla*I fragment was used as a hybridization probe.

The sequencing results show that three *Alu*I clones which are 221 bp long (MAD2A1-3) partially overlap the 203-bp *Cla*I fragment. They also span the first part of the 508-bp *Cla*I fragment (MAD2C4) which is adjacent to the 203-bp *Cla*I fragment (Fig. 2B). All three types of clones therefore characterize another highly repetitive DNA having a basic repeating unit of 711 bp. This DNA is designated as *T. madens* satellite II.

The Southern blot of *T. madens* DNA, after hybridization with cloned satellite II fragments, shows the presence of a 711-bp band in *Alu*I, *Cla*I, *Hinf*I, and *Taq*I digestion (Fig. 1D). Due to the presence of more than one recognition site for these enzymes in most of the 711-bp monomeric units, additional, shorter bands occur on Southern blot, e.g., bands of 203 and 508 bp in *Cla*I and *Taq*I digestion. Presence of an additional *Taq*I site at position 520 in the satellite II sequence results in a band of 317 bp, clearly visible on Southern blot. The ladder arrangement composed of multimers of a 711-bp repeating unit, as well as of intermediate bands, is visible in *Taq*I and *Hinf*I digestion. Quantification of the satellite II with spot hybridization and subsequent densitometry shows that it constitutes 4% of *T. madens* genomic DNA. Screening of gene bank reveals no similarity of *T. madens* satellites with any other known satellite DNA, indicating their species specificity.

Sequence Variability of Satellite DNAs

The consensus sequence of satellite I is 225 bp long, with 74% A + T content (Fig. 2A). Analysis of sequence variability of satellite I relative to consensus sequence shows that single point mutations mainly contribute to divergence of this satellite. They are spread randomly over the whole sequence. However, the same substitutions appear frequently at the same positions in different clones, pointing to the possibility that gene conversion has played a role in spreading sequence variants between satellite monomers. Deletions are rare and mostly in the form of single nucleotides except for the 5-nt deletion in the clone MAD1C5. The calculated variability of the satellite monomers relative to consensus sequence is 4.1%, which is comparable to other analyzed tenebrionid satellites (Petitpierre et al. 1995; Ugarković et al. 1995).

The consensus sequence of satellite II is 711 bp long, with 70% A + T content (Fig. 2B). The single nucleotide substitutions are again the most represented mutations; only 3 single nt insertions and one deletion are found. The resulted sequence variability is only 1.2%. Transversions are, in both satellites, more frequent than transitions. The ratio between nucleotide changes to A or T relative to G or C is 1.6 and 1.0 for satellite I and II, respectively, indicating no significant shift of these sequences to higher A + T content.

Internal Sequence Organization of T. madens *Satellites*

A search for direct repeats in *T. madens* satellites shows that the monomer of satellite I can be divided in two halves, the first one spanning region 3–115 nt and the

A

```
TRMADSAT11  CGATTTGGCTCAATTTTAAGTTGGGAATCAAGCTAAATCAAA-GTATTTCAACCA    54
TRMADSAT21  CGATTTGGCTCGATTTTAAATCAGAAATCGCTCTAAATCAATTGTATTTCAACGA    55
TRMADSAT12  CAATTTTGCTCGATTTTAATACAGAA-TCAATTCAAAACTAATATATTTTATTAG    54
TRMADSAT22  --ATTTGGCTCAATTTTAAACCAGTAATTGATCCAAATCAG-TGTATTTCAACAA    52
            ****.****.*******.  .*.* * ..  ***.*..  .***** *.  .

TRMADSAT11  AATACTGGCTCTAATAA-GAGTTTTCTTCTAAAATAAAGCAATTTGAGCTATTTT   108
TRMADSAT21  TGAAATAACT----------GTTTTCAGTCAAAATAAAGCAATTTGAGCTATTTC   100
TRMADSAT12  TATAATTGCTTCAATCT-TAATTTTAGGCTGAAATAAAGCAATTTGTGATATTTT   108
TRMADSAT22  AATCGACGCTTAATTAACTAGTTTTCTGCAAAATAAAGCAAAATG---------    98
            ... . .**       .**** .. .***********..**

TRMADSAT11  CAAT-C   113
TRMADSAT21  CAATTC   106
TRMADSAT12  GAAT--   112
TRMADSAT22  ------    98
```

B

```
5                           T      ----GATGA-----                    TTTGAGCTAT
ATTTGGCTCGATTTTAAATCAGAAATCGCTCTAAATCAA TGTATTTCAAC     AATAACT-GTTTTCAGTCAAAATAAAGCAA      T
       *         *   *   * * *      *                  *         * **                    T
TAAACCGAGTTAAAATTTGGTCATTAACTAGGTTTAGTC-ACATAAAGTTG     TAATTGA CAAAAGACGTTTTTATTTCGTT      C
205                                 TTTTAGCTGCGAAT      T                    TTACCTTAAC
```

Fig. 3. **A** Alignment of subunits of satellite I (TRMADSAT11, TRMADSAT12) and satellite II (TRMADSAT21, TRMADSAT22). Nucleotides identical in all four subunits are indicated with an *asterisk*. **B** Possible dyad structure within *T. madens* satellite II between nt 5 and 205. The positions of imperfect complementarity are marked with an *asterisk*.

second one region 116–2 nt. These halves are designated TRMADSAT11 and TRMADSAT12 (Fig. 3A) and their mutual sequence similarity is 66.1% (Table 1). No significant inverted repeats are detected in this satellite. Analysis of potential retardation characteristics by electrophoresis on native polyacrylamide gels showed normal migrational characteristics of circulary permutated satellite monomers, indicating that the satellite does not exhibit sequence-induced DNA curvature (not shown). This result was confirmed by computational analysis based on three different algorithms (Ugarković et al. 1992).

Searching for internal structural characteristics of satellite II shows the potential formation of a long, stable dyad structure between nucleotides 5 and 205 (Fig. 3B) due to the presence of a long inverted repeat in this part of the satellite. The dyad structure contains two stems, 50 and 29 nt long, each exhibiting a complementarity of 86%, and two loop structures of 23 and 19 nt. Similar to the case for satellite I, no indication for sequence-induced DNA curvature is found after analysis on polyacrylamide gel as well as computational analysis.

Sequence comparison between satellites I and II reveals that the first part of satellite II (3–108 nt) exhibits sequence similarity with the two halves of satellite I: 78.3% homology with TRMADSAT11 and 63.7% homology with TRMADSAT12 (Table 1). The first part of satellite II is designated TRMADSAT21 and is compared with subunits of satellite I in Fig. 3A. Due to the long inverted repeat in the satellite II, its second part (109–205 nt), in reverse orientation, again shows sequence similarity with both subunits of satellite I and with the first unit of satellite II (Table 1). The reverse orientation of the second half of satellite II is marked as

Table 1. Sequence homology of subunits from *T. madens* satellites I and II

Sequence pair	Homology (%)
TRMADSAT11-TRMADSAT12	66.1
TRMADSAT11-TRMADSAT21	78.3
TRMADSAT11-TRMADSAT22	72.5
TRMADSAT11-TRMADSAT23	79.8
TRMADSAT12-TRMADSAT21	63.7
TRMADSAT12-TRMADSAT22	60.2
TRMADSAT12-TRMADSAT23	61.8
TRMADSAT21-TRMADSAT22	76.5
TRMADSAT21-TRMADSAT23	69.7
TRMADSAT22-TRMADSAT23	70.8

TRMADSAT22 and is compared with other three subunits (Fig. 3A, Table 1). The part of satellite II composed of 89 nt between positions 614 and 702 again shares a similarity of 79.8% with subunit TRMADSAT11 and of 70.8% with TRMADSAT22, while similarity with the other two subunits is insignificant. This part of satellite II is designated TRMADSAT23 and a comparison with other subunits is shown in Table 1. The rest of 408 nt of satellite II (206–613 nt) has no similarity to the above-mentioned subunits, characteristic for both satellites. However, the frequent direct repeats present in this part of the sequence (Fig. 2B) indicate its occurrence by duplication of shorter motifs.

Localization of Satellites on *T. madens Chromosomes*

T. madens has a diploid chromosome number of $2n = 20$ (9 + XY) (Juan and Petitpierre 1991). The chromosomes
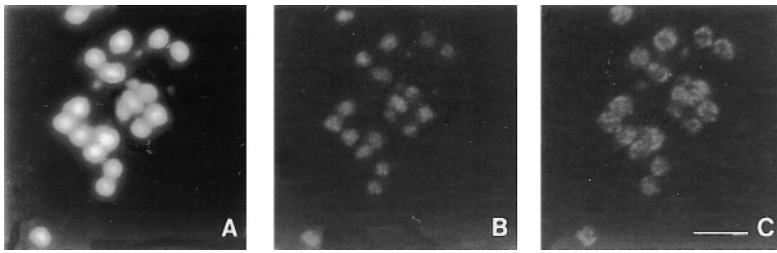
Fig. 4. Mitotic metaphase chromosomes of *T. madens* after staining with fluorochrome DAPI **(A),** fluorescent in situ hybridization with cloned satellite I **(B),** and satellite II probe **(C).** The pattern reveals colocalization of satellites in the heterochromatic regions of all 20 chromosomes. Three small *dots* represent supernumerary chromosomes. Bar = 3 μm.

have small sizes (1–3 μm), resulting in a poor morphology. In addition, the presence of small supernumerary chromosomes has been disclosed in some *T. madens* populations (Smith 1956).

Staining of condensed *T. madens* mitotic, metaphase chromosomes with fluorochrome DAPI which preferentially binds to A + T-rich regions induces intense fluorescent signals on all of the chromosomes (Fig. 4A). The bright fluorescence results from binding of DAPI to A + T-rich satellites I and II which constitute large heterochromatic blocks. In order to determine the distribution of each of the satellite, double-labeled fluorescent in situ hybridization has been performed. Hybridization with more abundant satellite I, labeled with digoxigenin and detected by rhodamine-labeled antibody, reveals distribution of satellite I on all chromosomes, coinciding with the regions of heterochromatin brightly stained with DAPI (Fig. 4B). Biotin-labeled satellite II, used as a hybridization probe, was detected by an avidin-FITC system with one-step signal amplification. Yellow fluorescent signal is again present on all of the chromosomes, coinciding with the regions of rhodamine fluorescence (Fig. 4C). However, the signal of satellite II appears to be in the form of dispersed dots, suggesting the possible distribution of minor satellite II within the larger blocks of more abundant satellite I.

Discussion

Most tenebrionid satellites are highly homogeneous within the species, similar to satellites found in some species of Diptera (Schmidt 1984; Bachmann and Sperlich 1993) and Hymenoptera (Tares et al. 1993). The satellites variants which are products of single base substitutions are randomly distributed within the satellite and uniformly spread on all chromosomes as revealed for *T. molitor* (Plohl et al. 1992). The efficient homogenization is due to the gene conversion and unequal crossing over, whose rate is probably effected by the large amount of satellite DNA in the genomes of tenebrionid species.

The tenebrionid satellites are usually unique to the particular species, although there are cases when the same satellite is shared between two related species. Satellite having a monomer length of 142 bp makes up 50% of the *T. molitor* genome and 23% of the *Tenebrio obscurus* genome, respectively. Sequence divergence be-

tween the two related satellites is 20% due to single base substitutions distributed randomly within the sequence. To the ''142-bp satellite family'' belongs also the satellite from *P. ratzeburgii* which, despite high sequence divergence, exhibits a tertiary structure of very similar geometry to those of two other 142-bp satellites (Ugarković et al. 1992).

Comparison of satellite DNAs from two parasitic wasps, *Diadromus pulchellus* and *Diadromus collaris,* shows that they derive from the same basic sequence motif which was amplified and subsequently diverged. In addition, insertion of a 169-bp fragment has invaded the satellite in *D. collaris,* creating a new repeating unit. Since the basic and the inserted unit contain palindromes with stable thermodynamic features it is suggested that this could be a positive selective value necessary for the insertion and its maintenance within the satellite (Rojas-Rousse et al. 1993). Two satellite DNAs from *T. madens* described here constitute a similar case. We propose that they evolved from a common ancestral sequence which was about 100 bp long. The amplification process was followed by mechanisms of sequence divergence: single point mutations, deletions, and gene conversion resulted in the creation of a new, longer repeating unit as the one characteristic for satellite I, which is composed of two basic subunits (Fig. 5A). In the evolution of satellite II, besides the previously mentioned processes of sequence divergence, an inversion of basic satellite subunit, caused by site-specific recombination, has taken place. Also, the insertion of a new sequence and its at-least-partial duplication has occurred, resulting in the creation of a complex, 711-bp-long monomer unit of satellite II (Fig. 5B). Satellite I is much more abundant than satellite II, constituting 30% of *T. madens* genomic DNA. It is A + T rich (74%) and exhibits sequence divergence of 4.1%, similar to other tenebrionid satellites. Satellite II has a similar, high A + T content of 70%, but is more rare and makes up 4% of the whole genome. From the comparison of subunits, presented in Table 1, it is evident that the first subunits in both satellites (TRMADSAT11 and TRMADSAT21) share a high sequence similarity of 78%, as well as subunits TRMADSAT11 and TRMADSAT23 with 79.8% mutual similarity. This could be explained by the origin of both satellites from the common ancestral subunit. Parallel but independent evolution of two satellites has resulted in divergence of subunits TRMADSAT12 and TRMADSAT22, having
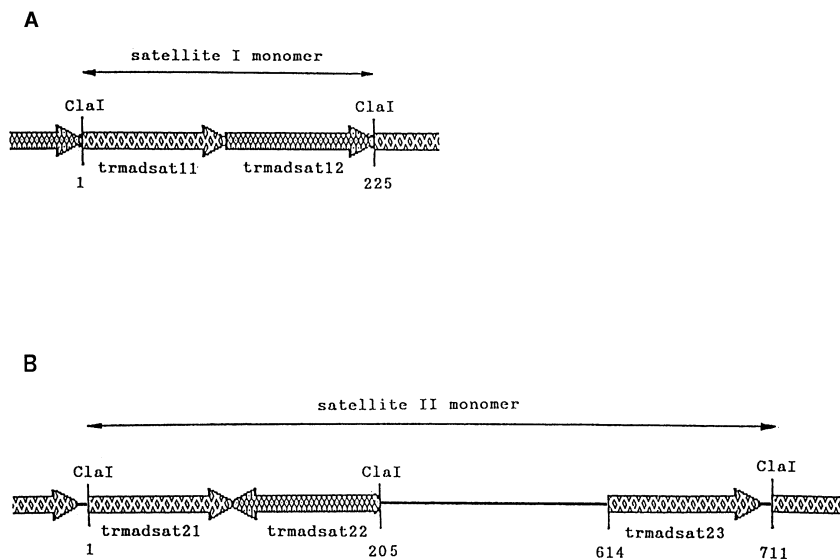
A

B

Fig. 5. Schematic representation of internal organization of *T. madens* satellites. **A** 225-bp-long monomer of satellite I composed of two subunits. **B** Monomer of satellite II, 711 bp long, composed of three related subunits indicated by *arrows* and of inserted DNA sequence between nt 206 and 613.

low mutual similarity of only 60.2%. In addition, inversion of the subunit within satellite II has also taken place.

The inversion of a satellite subunit has been previously observed in the simple satellite of *Drosophila simulans*. The repeat unit of this satellite is 15 bp long and comprised of three 5-bp subunits, two being arranged as direct repeats with the third in an inverted orientation (Lohe and Roberts 1988). Although the rearrangement of DNA segments is characteristic for bacteria, where it operates as gene-regulation mechanism (Simon et al. 1980), and for immunoglobulin genes inducing their somatic diverstification (Sakano et al. 1979), the results presented show that it can act as a mechanism effecting satellite DNA evolution within a species.

The inversion of a subunit within *T. madens* satellite I results in the appearance of a new repeating unit, characteristic for satellites II, which contains a stable palindromic structure similar to the termini of some transposable elements. However, the inverted repeat in satellite II could induce formation of a dyad structure which is considered to have a coding function for specific protein binding necessary for organization of higher order heterochromatin structure. It could be proposed that inversion of DNA sequence, known to be a very rare event, has occurred only once in the evolution of the *T. madens* satellite. However, the newly created sequence, due to acquired stable palindromic structure, has accrued some selective advantages, enabling its amplification and spread.

Mapping of simple satellite DNAs in heterochromatin of *D. melanogaster* reveals that satellites closely related in sequence are often located near one another on the same chromosomes (Lohe et al. 1993). The cytogenetic analysis of *T. madens* satellites shows their distribution on all chromosomes of the complement. Moreover, they are both positioned in the same chromosomal region characterized as pericentromeric constitutive heterochro-

matin. The more abundant satellite I is spread homogeneously over the whole heterochromatic region, while the small blocks of satellite II seem to be interspersed between the larger areas made up of satellite I. Such colocalization of the satellites and equal distribution on all of the chromosomes favor their proposed common origin.

References

Bachmann L, Sperlich D (1993) Gradual evolution of a specific satellite DNA family in *Drosophila ambigua, D. tristis* and *D. obscura*. Mol Biol Evol 10:647–659

Benson D, Lipman DJ, Ostell J (1993) Genbank. Nucleic Acids Res 21:2963–2965

Bigot Y, Hamelin MH, Periquet G (1990) Heterochromatin condensation and evolution of unique satellite DNA families in two parasitic wasp species: *Diadromus pulchellus* and *Eupelmus vuilleti* (Hymenoptera). Mol Biol Evol 7:351–364

Fowler RF, Skinner DM (1985) Cryptic satellites rich in inverted repeats comprise 30% of the genome of a hermit crab. J Biol Chem 260:1296–1303

Juan C, Petitpierre E (1991) Chromosome number and sex determining system in Tenebrionidae (Coleoptera). In: Zunino M, Belles X, Blas M (eds) Advances in coleopteorology. AEC, Barcelona, pp 167–176

Juan C, Gosalvez J, Mezzanotte R, Petitipierre E (1991) Cytological and biochemical characterization of the *in situ* endonuclease digestion of fixed *Tenebrio molitor* chromosomes. Chromosoma 100: 432–438

Juan C, Vazquez P, Rubio JM, Petitpierre E, Hewitt GM (1993) Presence of highly repetitive sequences in *Tribolium* flour-beetles. Heredity 70:1–8

Lohe AR, Roberts PA (1988) Evolution of satellite DNA sequences in *Drosophila.* In: Verma RS (ed) Heterochromatin. Cambridge University Press, Cambridge, MA pp 148–186

Lohe AR, Hilliker AJ, Roberts PA (1993) Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster.* Genetics 134:1149–1174

Martinez-Balbas A, Rodrigeuz-Campos A, Garcia-Ramirez M, Sainz J, Carrera P, Aymami J, Azorin F (1990) Satellite DNAs contain sequences that induce curvature. Biochemistry 29:2342–2348

Petitpierre E, Juan C, Pons J, Plohl M, Ugarković D (1995) Satellite DNA and constitutive heterochromatin in tenebrionid beetles. In: Kew Chromosome Conference IV. (In press)

Pinkel D, Gray J, Trask B, van den Engh G, Fuscoe J, van Dekken H (1986) Cytogenetic analysis by *in situ* hybridization with fluorescently labeled nucleic acid probes. Cold Spring Harb Symp Quant Biol 51:151–157

Plohl M, Borštnik B, Ugarković Đ, Gamulin V (1990) Sequence-induced curvature of *Tenebrio molitor* satellite DNA. Biochimie 72:665–670

Plohl M, Borštnik B, Lucijanić-Justić V, Ugarković Đ (1992) Evidence for random distribution of sequence variants in *Tenebrio molitor* satellite DNA. Genet Res 60:7–13

Plohl M, Lucijanić-Justić V, Ugarković Đ, Petitpierre E, Juan C (1993) Satellite DNA and heterochromatin of the flour beetle *Tribolium confusum.* Genome 36:467–475

Plohl M, Ugarković Đ (1994a) Analysis of divergence of *Alphitobius diaperinus* satellite DNA—roles of recombination, replication slippage and gene conversion. Mol Gen Genet 242:297–304

Plohl M, Ugarković Đ (1994b) Characterization of two abundant satellite DNAs from the mealworm *Tenebrio obscurus.* J Mol Evol 39:489–495

Pons J, Petitpierre E, Juan C (1993) Characterization of the heterochromatin of the darkling beetle *Mysolampus goudoti:* cloning of two satellite DNA families and digestion of chromosomes with restriction enzymes. Hereditas 119:179–185

Radić MZ, Lundgren K, Hamkalo BA (1987) Curvature of mouse satellite DNA and condensation of heterochromatin. Cell 50:1101–1108

Rojas-Rousse D, Bigot Y, Periquet G (1993) DNA insertions as a component of the evolution of unique satellite DNA families in two genera of parasitoid wasps: *Diadromus* and *Eupelmus* (Hymenoptera). Mol Biol Evol 10:383–396

Sakano H, Huppi K, Heinrich G, Tonegawa (1979) Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 280:288–293

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 9.16–9.19

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5476

Schmidt ER (1984) Clustered and interspersed repetitive DNA sequence family of *Chironomus.* The nucleotide sequence of the Cla-elements and of various flanking sequences. J Mol Biol 178:1–15

Simon M, Zeig J, Silverman M, Mandel G, Doolittle R (1980) Phase variation: evolution of controlling element. Science 209:1370–1374

Smith SG (1956) The status of supernumerary chromosomes in diabrotica. J Heredity 47:157–164

Sumner AT (1990) Banding with fluorochromes other than quinacrine. In: Hyman U (ed) Chromosome banding. Unwin Hyman, London, pp 163–165

Tares S, Cornuet JM, Abad P (1993) Characterization of an unusually conserved AluI highly reiterated DNA sequence family from the honeybee, *Apis mellifera.* Genetics 134:1195–1204

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

Ugarković Đ, Plohl M, Lucijanić-Justić V, Borštnik B (1992) Detection of satellite DNA in *Palorus ratzeburgii:* analysis of curvature profiles and comparison with *Tenebrio molitor* satellite DNA. Biochimie 74:1075–1082

Ugarković Đ, Petitpierre E, Juan C, Plohl M (1995) Satellite DNAs in tenebrionid species: structure, organization and evolution. Croat Chem Acta 68:627–638

Vogt P (1992) Code domains in tandem repetitive DNA sequence structures. Chromosoma 101:585–589