



Molecular characterization of SSS139, a new satellite DNA family in sibling species of the *Drosophila buzzatii* cluster

Fernando F. Franco¹, Fabio M. Sene¹ and Maura H. Manfrin²

¹*Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brazil.*

²*Departamento de Biologia, Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brazil.*

Abstract

We characterized sequences of a novel SSS139 *RsaI* satellite DNA family in *Drosophila gouveai* and *Drosophila seriema*, two members of the *Drosophila buzzatii* cluster (*D. repleta* group). The sequences were AT-rich (69%) with a monomer unit length of about 139 bp and contained two direct subrepeats of 14 bp and 16 bp, suggesting that it might have originated by the duplication of smaller sequences. Southern and dot-blot hybridization analyses also detected SSS139 in other *Drosophila buzzatii* cluster species (*D. koepferae*, *D. antonietae*, *D. borborema* and *D. serido*) but not in *D. buzzatii*. These results agree with the marginal phylogenetic position of *D. buzzatii* within the *D. buzzatii* cluster.

Key words: *Drosophila buzzatii* cluster, molecular evolution, repetitive DNA, satellite DNA, SSS139 satellite DNA.

Received: January 25, 2007; Accepted: June 26, 2007.

Satellite DNA consists of highly repetitive and tandemly arranged DNA sequences (Charlesworth *et al.*, 1994) which are considered the main components of constitutive heterochromatin and are usually located in the centromeric regions of chromosomes and, less frequently, in telomeres (Charlesworth *et al.*, 1994; Ugarkovic and Plohl, 2002). Satellite DNAs are heterogeneous and no general function has been attributed to this DNA class. However, despite the terminology “junk DNA” frequently associated with these sequences (Orgel and Crick, 1980), several biological roles have been suggested for some satellite DNA families, such as regulation of both heterochromatin condensation and genetic expression (reviewed by Ugarkovic, 2005). Moreover, satellite DNAs are involved in the maintenance of functional centromeres in mammals (Willard, 1990) and might also be related to the late replication of centromeres (Csink and Henikoff, 1998).

The satellite DNA sequences evolve in a concerted manner (Dover, 1982) and the main molecular mechanisms involved in its evolution are slippage replication, unequal crossing-over, gene conversion and rolling circle replication, which can, in general, induce a strong intra-specific homogenization of satellite DNA sequences and inter-

specific divergence (Dover, 1982; Charlesworth *et al.*, 1994).

The monomers of a satellite DNA can be species-specific (Bachmann *et al.*, 1994; Abadon *et al.*, 1998) or shared among phylogenetically related species (Watabe *et al.*, 1997; Lopez-Flores *et al.*, 2004; Kuhn and Sene, 2005). Though there are examples of extremely conserved satellite DNA sequences among species that diverged for long evolutionary periods (de la Herrán *et al.*, 2001; Mravinac *et al.*, 2002), most satellite DNA shared among species presents high evolutionary rates, acquiring diagnostic mutations for each species.

The *Drosophila buzzatii* cluster (*D. repleta* group, *D. buzzatii* complex) is a monophyletic group composed by cactophilic and sibling species that are naturally endemic to South America (Manfrin and Sene, 2006). Currently, the *D. buzzatii* cluster is composed of seven nominal *Drosophila* species (*D. antonietae*, *D. borborema*, *D. buzzatii*, *D. gouveai*, *D. koepferae*, *D. serido* and *D. seriema*). According to aedeagus and wing morphology, the *D. buzzatii* cluster can be divided into two species groups (Tidon-Sklorz and Sene, 2001; Moraes *et al.*, 2004). One of them is composed by *D. buzzatii* populations, and another group, called the *D. serido* sibling set, is composed by the remaining *D. buzzatii* cluster species.

The species of the *D. buzzatii* cluster share pBuM satellite DNA, an AT-rich family which can be divided into

two subfamilies, pBuM-1, consisting of tandemly arranged repetition units of approximately 190 bp, and pBuM-2 made up of 370 bp units in the same arrangement (Kuhn *et al.*, 1999; Kuhn and Sene, 2005). The sequences of the pBuM satellite DNA family have diagnostic nucleotide substitutions for the species of the *D. buzzatii* cluster as well as quantitative differences among them (Kuhn and Sene, 2005). At the intraspecific level, a strong inter-populational similarity was found in the pBuM sequences in *D. buzzatii*, *D. seriema* and *D. gouveai* (Kuhn *et al.*, 2003; Kuhn and Sene, 2004; Franco *et al.*, 2006). All species of the *D. buzzatii* cluster also contain the CG-rich DBC-150 satellite DNA composed of 150 bp monomers and restricted to dot chromosomes (Kuhn *et al.*, 2007). An interesting characteristic of DBC-150 is the low rates of homogenization presented in relation the pBuM family, which might be related to its exclusive location in the dot chromosomes, a region with presumed suppressed meiotic recombination.

In the present paper we describe how we characterized a *RsaI* satellite DNA family in *D. gouveai* and *D. seriema* (*D. buzzatii* cluster; *D. serido* sibling set) and used Southern and dot blot analyses to detect its presence in other members of the *D. serido* sibling set. We also discuss the origin of this satellite DNA and its presence/absence pattern in the *D. buzzatii* cluster in relation to the phylogenetic hypothesis available of this group.

Seven strains representing all species of the *Drosophila buzzatii* cluster, were studied: *D. antonietae* H84P4 (Sertãozinho, SP, Brazil); *D. borborema* 1281.0 (Cafarnaum, BA, Brazil); *D. buzzatii* J79H41 (Ibotirama, BA, Brazil); *D. gouveai* H74M2 (Altinópolis-SP, Brazil); *D. koepferae* B20D2 (Tapia, Argentina); *D. serido* 1431.4 (Cafarnaum, BA, Brazil) and *D. seriema* D71C1B (Morro do Chapéu-BA, Brazil). These species are maintained in the Genetic Evolution Laboratory (Laboratório de Genética Evolutiva) at the University of São Paulo, Ribeirão Preto, Brazil.

The genomic DNA of adult flies was extracted according to Preiss *et al.* (1988) and that of *D. gouveai* (H74M2) and *D. seriema* (D71C1B) digested with the *RsaI* restriction enzyme (Invitrogen) and electrophoresed on 5% polyacrylamide gel using a 100 base-pair ladder marker (Fermentas) as standard. Prominent bands of approximately 150 bp were eluted from the gel by overnight incubation at 37 °C in an elution solution containing 500 mM sodium acetate plus 1 mM ethylenediaminetetraacetic acid (EDTA) and the recovered fragments ligated to a pUC18 plasmid vector (Amersham) and transformed in competent *E. coli* DH5- α (Amersham), recombinant clones being identified using the β -galactosidase blue-white selection system (Sambrook *et al.*, 1989). Plasmid DNA was prepared according the methodology described in Sambrook *et al.* (1989) and the DNA template reaction for sequencing was prepared according to the BigDye Terminator Cycle

Sequencing Ready Reaction Kit manual (Perkin-Elmer). Automatic DNA sequencing was performed on an ABI Prism 377 sequencer (Perkin-Elmer).

Southern hybridization of all *D. buzzatii* cluster species was carried out using 5 μ g of genomic DNA from each of the strains tested. The DNA was digested with *RsaI* and the fragments separated in 1% agarose gel and blotted onto nylon membranes (Amersham) according to Sambrook *et al.* (1989). Labeling of probe DNA (clone D71C1B-150/10) was done using the Random Primers DNA Labeling System (Gibco BRL) and membranes were hybridized at 60 °C overnight in a hybridization solution containing 7% sodium dodecyl sulfate (SDS) and 0.5 M Na₂HPO₄ and washed twice at about 25 °C for 15 min in a solution containing 5% SDS, 40 mM Na₂HPO₄ and 1 mM EDTA. For dot-blot analyses 50 ng, 100 ng and 250 ng of genomic DNA from each of the *D. buzzatii* cluster species tested were placed on nylon membranes (GE Healthcare) and denatured with a solution containing 0.5M NaOH and 1.5 M NaCl, neutralized with a solution containing 1 M Tris-Base and 1.5 M NaCl (pH 7.4) and then rinsed with 2x saline-sodium citrate (SSC) buffer. The membranes were then dried and UV crosslinked. The labeling of the probe and hybridization used the conditions described above for Southern hybridization. Genomic DNA from *Drosophila coroica* (*D. repleta* group, *D. fasciola* subgroup) was used as negative control and the sequence of the clone D71C1B-150/10 as the positive control.

The alignment of the sequences was carried out using the CLUSTAL W program v1.8 (Thompson *et al.*, 1994) and sequences homology searches were performed on the NCBI GenBank using BLASTn (Altschul *et al.*, 1990). Detection of direct repeats in the satellite DNA monomers was performed on the consensus sequence using the online search program OligoRep.

We obtained six sequences, four from *D. seriema* (GenBank EF035014 to EF035017) and two from *D. gouveai* (GenBank EF035018 and EF035019). The high similarity found in all sequences indicating that they belonged to the same DNA family (Figure 1). All of the sequences obtained were AT-rich (~69% on average) and their sizes range from 138 bp (D71C1B-150/14) to 145 bp (H74M2-150/3).

Southern hybridization revealed a ladder pattern, characteristic of highly repetitive sequences (Plohl *et al.*, 2004). The presence of additional and larger than expected bands (~140 bp) of decreasing intensity is explainable by the variability presented at the restriction enzyme recognition site in some copies in the tandem array (Figure 2a).

No homolog sequence obtained in this research was found in the NCBI GenBank (searched on May 15, 2007). These results indicate that the sequences described here constitute an uncharacterized satellite DNA family, which we called SSS139 (the letters are a reference for serido sibling set and the numbers representing the average monomer

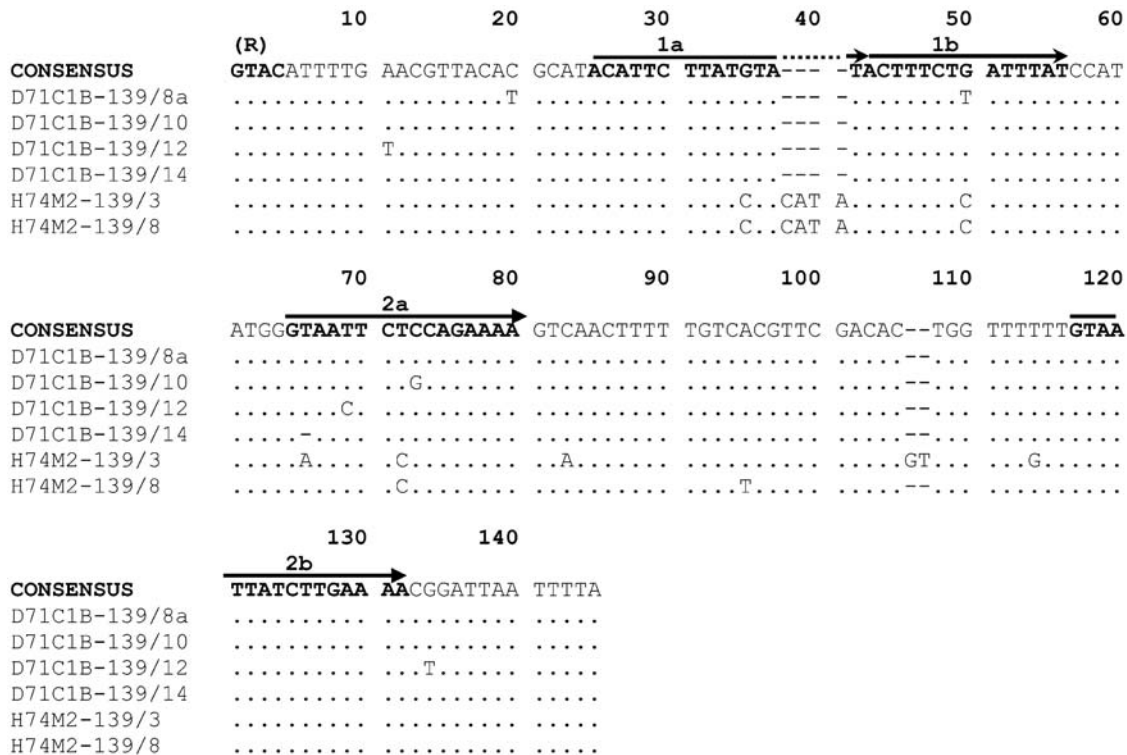


Figure 1 - Alignment containing the sequences of the SSS139 satellite DNA from *Drosophila gouveai* (H74M2) and *Drosophila seriema* (D71C1B) aligned with the SSS139 consensus sequence. The direct subrepetitions 1a and 1b (78.6% similarity) and 2a and 2b (81.25% similarity) are in bold and indicated with arrows. (R) Indicates the location of the *RsaI* restriction site (GTAC). (.) Indicates similar bases, (-) indicates *indels*.

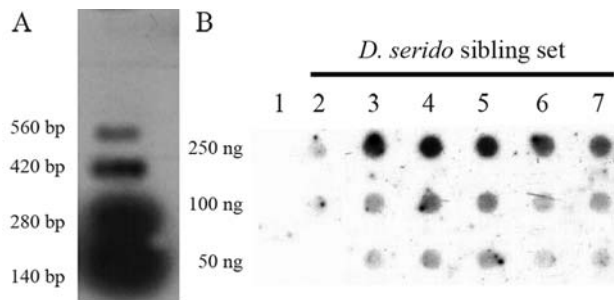


Figure 2 - A. Southern blotting membrane hybridization. The ladder pattern characteristic of sequences organized in tandem obtained by digestion of genomic DNA (10 µg) from *Drosophila seriema* with the *RsaI* restriction enzyme and submitted to hybridization with D71C1B-139/10 probe. B. Dot-blot membrane hybridization using 50 ng, 100 ng and 250 ng of genomic DNA from each species applied to a nylon membrane and submitted to hybridization with the D71C1B-139/10 probe. The following *Drosophila* species are shown: (1) *D. buzzatii*; (2) *D. koepferae*; (3) *D. antonietae*; (4) *D. gouveai*; (5) *D. seriema*; (6) *D. borborema*; and (7) *D. serido*.

size). This satellite DNA is unrelated to the other two satellite families (pBuM and DBC-150) described for the *D. buzzatii* cluster species (Kuhn and Sene, 2005; Kuhn *et al.*, 2007).

Despite the high nucleotide similarity (~95% on average), the SSS139 copies of *D. gouveai* and *D. seriema* have different nucleotides at positions 35, 50 and 72, suggesting that these sites might be species-specific. Moreover, the

CATA sequence in positions 37 to 41 of the consensus sequence also might be specific for *D. gouveai* (Figure 1), suggesting that this satellite DNA evolved independently in *D. gouveai* and *D. seriema*. Indeed, concerted evolution can lead to fast divergence of satellite sequences in reproductively isolated species through independent horizontal spreading of variant monomers in a tandem array in each lineage (Dover, 1982; Charlesworth *et al.*, 1994).

One of the hypotheses for the origin of satellite DNA monomers suggests the amplification, rearrangement and subsequent differentiation of small motifs (Ugarkovic and Plohl, 2002). Lee *et al.* (1997) proposed that satellite DNA I, shared among deer and cattle, emerged from a 31 bp motif that was initially amplified 25 times to form a complex monomer. It has also been proposed that *HindIII* satellite DNA family sequences (shared among species of sturgeons from the *Ancipenser* genus) with repetition units of from 169 bp to 172 bp evolved from various duplications of shorter sequences (de la Herrán *et al.*, 2001). We found the direct subrepetitions 1a/1b (78.6% similarity) and 2a/2b (81.25% similarity) in the consensus sequence of SSS139 satellite DNA (Figure 1). This suggests that the SSS139 monomers may have evolved by duplication of smaller sequences and the later accumulated mutations, giving rise to a longer sequence that may have been amplified by the molecular mechanisms such as unequal crossing over or rolling circle replication which are often responsible for the

increase of copy number in tandemly arranged sequences (Charlesworth *et al.*, 1994). An alternative hypothesis is that the 1a/1b and 2a/2b subrepetitions emerged from a pre-existing SSS139 monomer and this repetition unit duplication carrier was spread horizontally by concerted evolution events.

The occurrence of SSS139 in other species of the *D. buzzatii* cluster was detected with the Southern blot (data not shown) and dot-blot analyses and there were signs of hybridization with the SSS139 probe for all the *D. buzzatii* cluster species forming the *D. serido* sibling set, but not for *D. buzzatii* strains (Figure 2b). However, although our data suggests the absence of SSS139 in *D. buzzatii* species, the hypothesis that SSS139 is highly divergent or is present in a residual quantity in *D. buzzatii* cannot be discarded because the Southern and dot blot analyses were performed under high stringency conditions.

According to phylogenies based on mitochondrial (Manfrin *et al.*, 2001) and nuclear genes (Rodríguez-Trelles *et al.*, 2000), *D. buzzatii* belongs to the most basal lineage of the *D. buzzatii* cluster. Furthermore, among the species of the cluster, *D. buzzatii* has a primitive pattern of chromosome inversions (Ruiz and Wasserman, 1993) and is the most reproductively isolated species (Marin *et al.*, 1993; Madi-Ravazzi *et al.*, 1997; Machado *et al.*, 2006). Morphologically, the aedeagus and wing morphometry of this species is different from that of the other species of the cluster (Tidon-Sklorz and Sene 2001; Moraes *et al.*, 2004; Manfrin and Sene, 2006). In this context, it is not surprising that other species in the *D. buzzatii* cluster share characteristics not found in *D. buzzatii*.

With regard to the satellite DNA, the pBuM (Kuhn *et al.*, 1999; Kuhn and Sene, 2005) and the SSS139 families have yielded some insights into the evolutionary relationships of the species from the *D. buzzatii* cluster. Regarding the pBuM family, *D. buzzatii* presents the most divergent sequences and shows several species-specific substitutions. Nevertheless, sequences of this satellite DNA from other species from the *D. buzzatii* cluster present relatively fewer diagnostic mutations (Kuhn *et al.*, 1999; Kuhn and Sene, 2005). The presence or absence pattern of the SSS139 satellite DNA family also subdivides *D. buzzatii* cluster species into the *D. buzzatii* group (absence of SSS139) and the *D. serido* sibling set (presence of SSS139) (Figure 2b). The main conclusion from these data is that *D. buzzatii* is the most differentiated species of the *D. buzzatii* cluster, in agreement with its marginal phylogenetic position within the cluster (Manfrin *et al.* 2001; Rodríguez-Trelles *et al.*, 2000).

As suggested by morphological characters, the *D. serido* sibling set species form a monophyletic group in a phylogeny based on the xanthine dehydrogenase nuclear gene (Rodríguez-Trelles *et al.*, 2000). On the other hand, according to phylogeny based on mitochondrial DNA data the *D. serido* sibling set is polyphyletic due the phylogen-

etic position of *D. koepferae* as a sister group of *D. buzzatii* in the basal lineage (Manfrin *et al.*, 2001). The pBuM and SSS139 satellite DNA data suggest that *D. koepferae* could be more closely related to the remaining *D. serido* sibling set species than to *D. buzzatii* species, thus supporting morphological (Tidon-Sklorz and Sene 2001; Moraes *et al.*, 2004) and nuclear gene data (Rodríguez-Trelles *et al.*, 2000) and in disagreement with mitochondrial DNA data (Manfrin *et al.*, 2001). In this context, we suggest that the phylogenetic position of *D. koepferae* must be reconsidered by further studies using more molecular markers.

Acknowledgments

We thank Professor Nilce M. Martinez-Rossi for the use of her laboratory for some of the experiments. We also are grateful to Paulo R. Epifanio and Roseli A.P. Ferreira for technical assistance, Gustavo C.S. Kuhn for reading the first draft of the manuscript and to anonymous referees for useful and critical comments. This work was supported by several grants from the Brazilian agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP, grant number 03/05031-0), Financiadora de Estudos e Projetos (FINEP) and Universidade de São Paulo (USP) to MHM. FAPESP made personal fellowships to FMS (grant number 04/08565-8) and FFF (grant number 05/51780-0).

References

- Abadon M, Grenier E, Laumond C and Abad P (1998) A species-specific satellite DNA from the entomopathogenic nematode *Heterorhabditis indicus*. *Genome* 41:148-153.
- Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-10.
- Bachmann L, Venanzetti F and Sbordoni V (1994) Characterization of a species-specific satellite DNA family of *Dolichopoda schiavazzi* (Orthoptera, Rhaphidophoridae) cave crickets. *J Mol Evol* 39:274-281.
- Charlesworth B, Sniegowski P and Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215-220.
- Csirik AK and Henikoff S (1998) Something from nothing: The evolution and utility of satellite repeats. *Trends Genet* 14:200-204.
- de la Herrán R, Fontana F, Lanfredi M, Congiu L, Leis M, Rossi R, Ruiz Rejon C, Ruiz Rejon M and Garrido-Ramos MA (2001) Slow rates of evolution and sequence homogenization in an ancient satellite DNA family of sturgeons. *Mol Biol Evol* 18:432-436.
- Dover G (1982) Molecular drive: A cohesive mode of species evolution. *Nature* 299:111-117.
- Franco FF, Kuhn GCS, Sene FM and Manfrin MH (2006) Conservation of pBuM-2 satellite DNA sequences among geographically isolated *Drosophila gouveai* populations from Brazil. *Genetica* 128:287-295.

- Kuhn GCS, Bollgönn S, Sperlich D and Bachmann L (1999) Characterization of a species-specific satellite DNA of *Drosophila buzzatii*. *J Zoolog Syst Evol Res* 37:109-112.
- Kuhn GCS, Franco FF, Silva-Jr WA, Martinez-Rossi NM and Sene FM (2003) On the pBuM189 satellite DNA variability among South American populations of *Drosophila buzzatii*. *Hereditas* 139:161-166.
- Kuhn GCS and Sene FM (2004) Characterization and inter-population variability of a complex *HpaI* satellite DNA of *Drosophila seriema* (*repleta* group). *Genetica* 121:241-249.
- Kuhn GCS and Sene FM (2005) Evolutionary turnover of two pBuM satellite DNA subfamilies in the *Drosophila buzzatii* species cluster (*repleta* group): From *alpha* to *alpha/beta* arrays. *Gene* 349:77-85.
- Kuhn GCS, Franco FF, Manfrin MH, Moreira-Filho O and Sene FM (2007) Low rates of homogenization of the DBC-150 satellite DNA family restricted to a single pair of microchromosomes in species from the *Drosophila buzzatii* cluster. *Chrom Res* 15:457-469.
- Lee C, Court DR, Cho C, Haslett JL and Lin CC (1997) Higher-order organization of subrepeats and the evolution of cervid satellite I DNA. *J Mol Evol* 44:327-335.
- Lopez-Flores I, de la Herran R, Garrido-Ramos MA, Boudry P, Ruiz-Rejon C and Ruiz-Rejon M (2004) The molecular phylogeny of oysters based on a satellite DNA related to transposons. *Gene* 339:181-188.
- Machado LP, Madi-Ravazzi L and Tadei WJ (2006) Reproductive relationships and degree of synapsis in the polytene chromosomes of the *Drosophila buzzatii* species cluster. *Braz J Biol* 66:279-293.
- Madi-Ravazzi L, Bicudo HE and Manzato JA (1997) Reproductive compatibility and chromosome pairing in the *Drosophila buzzatii* complex. *Cytobios* 89:21-30.
- Manfrin MH, Brito ROA and Sene FM (2001) Systematics and evolution of the *Drosophila buzzatii* (Diptera, Drosophilidae) cluster using mtDNA. *Ann Entomol Soc Am* 94:333-346.
- Manfrin MH and Sene FM (2006) Cactophilic *Drosophila* in South America: A model for evolutionary studies. *Genetica* 126:57-75.
- Marin I, Ruiz A, Pla C and Fontdevila A (1993) Reproductive relationships among ten species of the *Drosophila repleta* group from South America and the West Indies. *Evolution* 47:1616-1624.
- Moraes EM, Spressola VL, Prado PRR, Costa LF and Sene FM (2004) Divergence in wing morphology among sibling species of the *buzzatii* cluster. *J Zoolog Syst Evol Res* 42:154-158.
- Mravinac B, Plohl M, Mestrovic N and Ugarkovic D (2002) Sequence of PRAT satellite DNA "frozen" in some coleopteran species. *J Mol Evol* 54:774-783.
- Orgel LE and Crick FHC (1980) Selfish DNA: The ultimate parasite. *Nature* 284:604-607.
- Preiss A, Hartley DA and Artavanis-Tsakonas S (1988) Molecular genetics of enhancer of split, a gene required for embryonic neural development in *Drosophila*. *EMBO J* 12:3917-3927.
- Plohl M, Bruvo B, Mestrovic N, Mravinac B, Petrovic V, Durajlija-Zinic S and Ugarkovic D (2004) Satellite DNA sequences in centromeric heterochromatin. *Period Biol* 106:95-102.
- Rodriguez-Trelles F, Alarcon L and Fontdevila A (2000) Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): A maximum-likelihood approach. *Mol Biol Evol* 17:1112-1122.
- Ruiz A and Wasserman M (1993) Evolutionary cytogenetics of *Drosophila buzzatii* species complex. *Heredity* 70:582-596.
- Sambrook J, Fritsh EF and Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Tidon-Sklorz R and Sene FM (2001) Two new species of the *Drosophila serido* sibling set (Diptera, Drosophilidae). *Iheringia* 90:141-146.
- Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680.
- Ugarkovic D and Plohl M (2002) Variation in satellite DNA profiles - Causes and effects. *EMBO J* 21:5955-5959.
- Ugarkovic D (2005) Functional elements residing within satellite DNAs. *EMBO Rep* 11:1035-1039.
- Watabe H, Bachmann L, Haring E and Sperlich D (1997) Taxonomic and molecular studies on *Drosophila sinobscura* and *D. hubeiensis*, two sibling species of the *D. obscura* group. *J Zoolog Syst Evol Res* 35:81-94.
- Willard HF (1990) Centromeres of mammalian chromosomes. *Trends Genet* 6:410-416.

Internet Resource

OligoRep at <http://wwwmgs.bionet.nsc.ru/mgs/programs/oligorep/>)

Associate Editor: Louis Bernard Klaczko