Comparative analysis of different satellite DNAs in four *Mytilus* species

A. Martínez-Lage[a], F. Rodríguez[a], A. González-Tizón[a], E. Prats[b], L. Cornudella[b] and J. Méndez[a*]

[a] Dept. Biología Celular y Molecular, Universidade da Coruña, A Zapateira s/n, 15071 La Coruña, Spain.

[b] Dept. Biología Molecular y Celular, Instituto de Biología Molecular de Barcelona, Centro Superior de Investigaciones Cientificas (CSIC), Jordi Girona 18-26, 08034 Barcelona, Spain.

How to cite:

Comparative analysis of different satellite DNAs in four *Mytilus* species. A. Martínez-Lage, F. Rodríguez, A. González-Tizón, E. Prats, L. Cornudella, J. Méndez, *Genome*, 2002, 45(5), 922-929. https://doi.org/10.1139/g02-056

Abstract

We report the characterization of three satellite DNAs in four species of mussel: *Mytilus edulis*, *Mytilus galloprovincialis*, *Mytilus trossulus*, and *Mytilus californianus*. The monomers of the *Apa I* satellite DNAs were 173, 161, and 166 bp long. These satellite monomers were used to construct phylogenetic trees to infer relationships among these species. The topologies obtained clearly indicate that *M. californianus* is the most divergent species with respect to the other three. Furthermore, localization of satellite DNAs on metaphase chromosomes was performed using fluorescent in situ hybridization (FISH). Fluorescent signals revealed a different organization and distribution of these three satellite DNAs.

Keywords: satellite DNA; phylogeny; mussel; *Mytilus*

Introduction

Mussels belonging to the genus *Mytilus* are one of the most thoroughly studied marine molluscs at ecological, physiological, and genetic levels, although their origin and taxonomic status still remain unclear. Four *Mytilus* species, *Mytilus edulis*, *Mytilus galloprovincialis*, *Mytilus trossulus*, and *Mytilus californianus*, are recognized. The first three are morphologically similar and are able to cross hybridize in the areas where they coexist (McDonald et al. 1991; Gosling 1992). Analysis of mitochondrial DNA sequences has revealed that *M. californianus* is the most divergent of these species, whereas *M. edulis* and *M. galloprovincialis* are the most similar (Rawson and Hilbish 1995, 1998; Hoeh et al. 1997; Hilbish et al. 2000). Analysis of 18S rDNA sequences (Kenchington et al. 1995; Distel 2000) again showed that *M. californianus* is the most divergent species. *Mytilus galloprovincialis*, *M. edulis*, and *M. trossulus* form a star radiation topology, although this is unreliable given that the sequence alignment matrix contains insufficient informative sites (Kenchington et al. 1995).

Tandemly repeated DNA sequences, particularly satellite DNA, make up a major part of eukaryotic genomes and they are usually located in constitutive heterochromatin regions (Brutlag 1980). These sequences show extreme diversity in type, abundance, and chromosomal distribution even among closely related species (Miklos 1985; Charlesworth et al.1994). Several studies have suggested that highly repetitive sequences may play a role in heterochromatin condensation (Karpen 1994), in gene expression

(Maiorano et al. 1997), in the stability of the genome structure, and in its perpetuation (Sainz and Cornudella 1990). Distinctive features of repetitive DNA can often provide valuable information about genome organization and the evolutionary status of a given species (Wijers et al. 1993). For example, repetitive DNA has been used in phylogenetic studies in mammals (Modi et al. 1996), fishes (Nabegama et al. 2000), and insects (Mestrovic et al. 2000) because it generally shows very high intraspecific homogeneity, whereas interspecific variability increases as a direct function of phyletic distance. Mollusc repetitive sequences have only been investigated in *Crassostrea gigas* (McLean and Whiteley 1973; Clabby et al. 1996), *Mytilus edulis* (Ruíz-Lara et al. 1992; Ruíz-Lara 1993), *Donax trunculus* (Plohl and Cornudella 1996, 1997), *Argopecten irradians* (Estabrooks 1999), *Adamussium colbecki* (Canapa et al. 2000), and *Haliotis* spp. (Muchmore et al. 1998). In *Haliotis* , the characterization of satellite DNA sequences has proven to be a useful molecular marker for identification of different species (Muchmore et al. 1998). Furthermore, in the mussel species *M. edulis*, three repetitive DNA sequences have been reported by Ruíz-Lara et al. (1992) and Ruíz-Lara (1993). These sequences, named types 1, 2, and 3, differ in genomic content and in base pair length, being 0.63% and 173 bp long for type 1, 3.09% and 161 bp for type 2, and 0.07% and 166 bp for type 3.

Taking into account the potential use of repetitive DNA in evolutionary biology, we have characterized three satellite DNA sequences in the mussel species *M. edulis*, *M. galloprovincialis*, *M. trossulus*, and *M. californianus* in an attempt to analyze the divergence process of these satellite DNAs.

Materials and methods

Sample collection

Mytilus edulis mussels were collected from Yerseke (Holland), *M. galloprovincialis* from Balcobo (northwest coast of Spain), *M. trossulus* from Esquimalt Lagoon (Vancouver Island, B.C.), and *M. californianus* from Point No Point (Vancouver Island, B.C.). Adductor muscles or mantle were combined to obtain 3–6 g wet weight, and high molecular weight DNA was extracted following the protocol described in Rice and Bird (1990).

PCR amplifications and cloning

Specific primers were designed using OLIGO software (Rychlik and Rhoads 1989) from *M. edulis Apa I* repetitive DNA sequences (accession Nos. X61120, X61119, AJ249690) to isolate these repetitive DNA sequences by PCR. The primers used were as follows: S1F, 5′-TTC CAT ATC AAC CAC ACA TT-3′ and S1B, 5′-AAT ACG TAA ATG GGC AAC TT-3′ for DNA type 1; S2F, 5′-ACC AAA ACT CCC AAA ATC AA-3′ and S2B, 5′-CCC AAC AGT TTA GGA ATT AG-3′ for DNA type 2; and S3F, 5′-GTC ACC CTA CCG CTA CTT TG-3′ and S3B, 5′-CAT CCA CCT TCT TTC GTT CA-3′ for DNA type 3. The PCR amplification reactions (25 μL) were performed in a PCR Supermix (Life Technologies, Carlsbad, Calif.) containing 25 ng of template DNA, 0.5 U of *Taq* DNA polymerase and 0.2 mM of each primer. The PCR amplification profile consisted of one initial cycle of 3 min at 94°C, followed by 35 amplification cycles of 1 min at 94°C, 30 s at 45°C (55 and 58°C for type 2 and 3 sequences, respectively), 45 s at 72°C, and a final extension cycle at 72°C for 5 min.

The fragments obtained by PCR were ligated into the plasmid pGEM-T Easy Vector system (Promega, Madison, Wis.) and used to transform *Escherichia coli* JM109 competent cells. Recombinant clones were selected as white colonies on ampicillin plates containing X-gal and IPTG, and screened by electrophoresis after miniprep isolation to select recombinant clones corresponding to putative dimers. Insert sequences were determined by the dideoxy chain-termination method (Sanger et al. 1977), using

Thermo Sequenase™ Cy™ 5 Dye automatic sequencing in an ALFexpress instrument (Amersham Pharmacia Biotech, Piscataway, N.J.).

Sequence analysis and comparison

The internal repeats of satellite DNA sequences were employed to perform multiple sequence alignments using the computer program CLUSTAL X (Thompson et al. 1997), omitting the primer sequence. The sequences of the monomeric units were compared with those in EMBL and GenBank databases using the NCBI BLAST server (Altschul et al. 1997).

Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 2.1 (Kumar et al. 2001). Genetic distances were calculated according to Kimura's (1980) two-parameter model and the resulting distance matrices were subjected to UPGMA (Sneath and Sokal 1973) and neighbor-joining (Saitou and Nei 1987) analysis. Maximum-parsimony trees were constructed using a branch-and-bound search. Resolution of internal nodes was performed using 1000 bootstrap iterations with random replacement (Felsenstein 1985).

Copy number

Solutions containing defined amounts of denatured genomic DNA (200, 100, 50, and 25 ng) and denatured *M. edulis* repetitive DNA fragments (0.8. 0.4, 0.2, and 0 ng) were vacuum blotted onto nylon membranes in a Bio-Rad dot blotting apparatus. *Mytilus edulis* repetitive DNA was labeled with digoxigenin-11-dUTP using PCR and hybridized to the blotted DNA. Detection of hybridization was performed using the Dig luminescent detection kit for nucleic acids (Roche) following the manufacturer's instructions. Hybridization signals were densitometrically quantified with the Leica Q-win program ver. 2.2 from Leica Imaging Systems Ltd. The copy numbers of *Mytilus* spp. repetitive sequences were calculated from a linear plot made from dilutions of *M. edulis* repetitive DNA fragments. The haploid DNA content of *M. edulis* and *M. galloprovincialis* are 1.71 and 1.92 pg, respectively (Rodríguez-Juiz et al. 1996), and 1.51 and 1.61 pg for *M. trossulus* and *M. californianus*, respectively (González-Tizón et al. 2000).

Table 1. Copy number and percent genomic content (in brackets) for the different satellite DNAs in the four *Mytilus* mussel species.

Species	Type 1	Type 2	Type 3
M. edulis	45 000 (0.46%)	140 000 (1.31%)	1100 (0.01%)
M.galloprovincialis	79 000 (0.73%)	200 000 (1.66%)	3500 (0.03%)
M. trossulus	70 000 (0.79%)	140 000 (1.51%)	11 000 (0.12%)
M. californianus	8500 (0.10%)	80 000 (0.85%)	10 000 (0.11%)

Fluorescent in situ hybridization

Samples collected from Yerseke (*M. edulis*), Balcobo (*M. galloprovincialis*), and Esquimalt Lagoon (*M. trossulus*) were fed continuously with a microalgae suspension in the laboratory. Metaphases were obtained as described by González-Tizón et al. (2000), carrying out colchicine hypotonic treatment and routine air-drying methods. The DNA probes used were pMEA 17, pMEA 18, and pMEA 16 from *M. edulis* (obtained from Cornudella's group), containing the monomer units of repetitive DNA types 1, 2, and 3, respectively. Hybridization and detection of signals were performed as previously described by González-Tizón et al. (2000). The chromosomes were counterstained with propidium iodide and photographed using a Leica RXA microscope and the appropriate filter combinations. The film used was Kodak Ektachrome Elite color print.
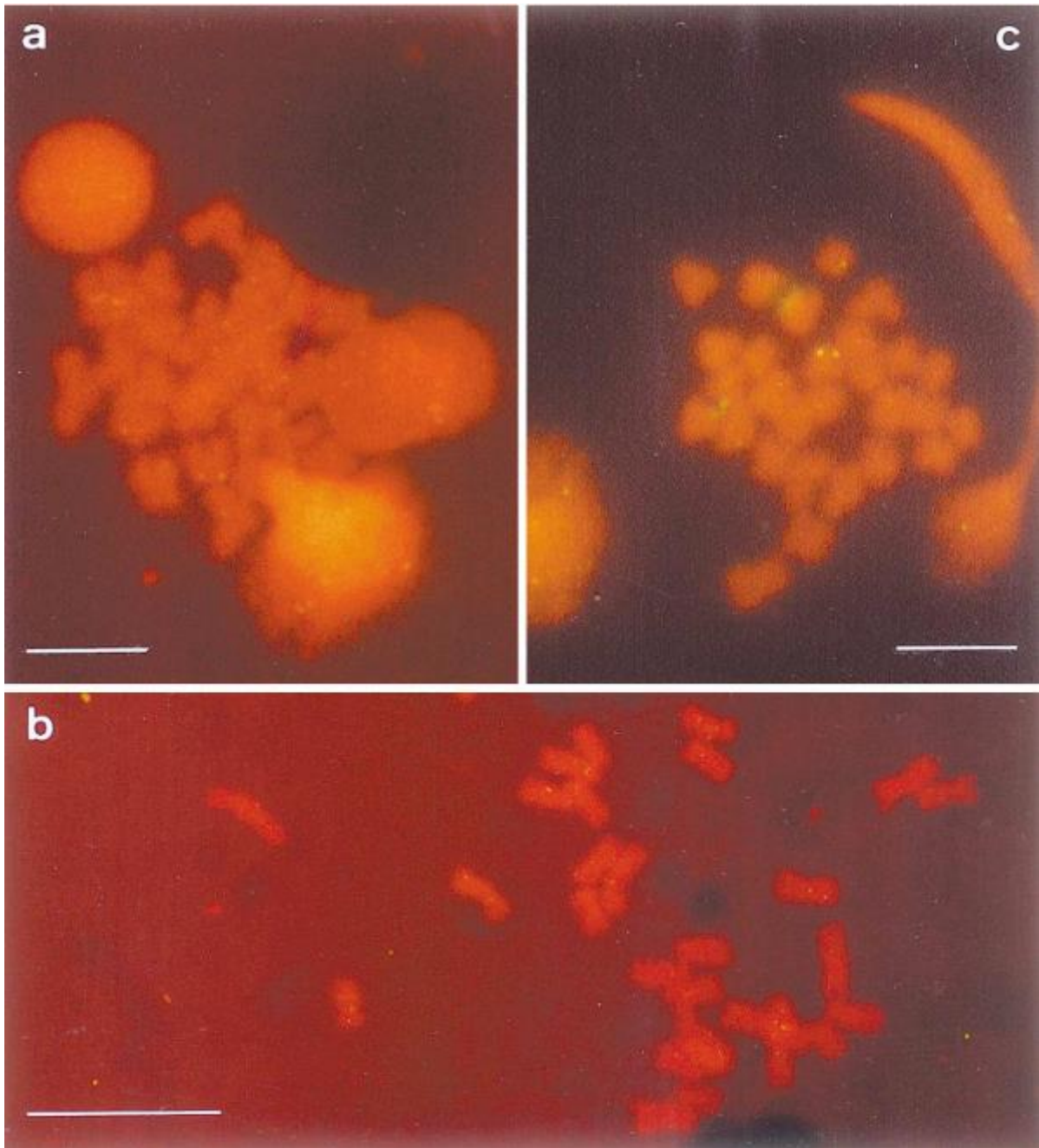
a)
```
                   CGTCTGGCCA CCACAGAGGT TCAAAGTTGC CCATTTACGT ATTTTCCATA TCAACCACAC ATTCGG.... .TACTCCTCG .ACATATCCC TACGGCAAA. . GTTCCGAAC
pEYE101            ---T------ ---------- ---------- ---------- ---------- ---------- ---------- ------TT--- -G-------- ---------- ----------
pEYE110            ---------- -----A---- ---------- ---------- ---------- ---------- ---------- ------T--- -G--C----- ---------- ----------
pEYE112            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------T--- -G-------- ---------- ----------
pGBA102            T--------- ---------- ---------- ---------- ---------- ---------- ---------- ---T------ ---------- ---------- ----------
pGBA110            ---------- --T------- ---------- ---------- ---------- ---------- ---------- --T------- ---------- ---------- ----------
pGBA111            ---------- ---A------ ---------- ---------- ---A------ ---------- ---------- ---------- ---------- ----T----- ----------
pGBA112            T--------- --T------- ---------- ---------- ---------- ---------- ---------- --T------- ---------- ----T----- ----------
pGBA113            ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
pTEL101            ---T------ ---------- ---------- ---------- ---------- ---------- ---------- ------TT--- -G-------- ---------- ----------
pTEL110            -----A---- ---------- ---------- ---------- ---------- ------T--- ---------- ------TT--- -G-----A-- ---------- ----------
pTEL112            ---------- ---------- ---------- ---------- ------T--- ---------- ---------- ------TT--- -GA------- ---------- ----------
pCaP104            .....TAAT- -GTA-AT--G CG--CT---A A-CG--..-- GG-GG-.-G- AGGGG-CGGA T-.-A-AAAA A--T--AAGT G-.-A--AGG -TG--TT--T AAC-TT---A
pCaP105            .....TAAT- -GTA-AT--C ..--CT---A A-CGC-..-- GG-GC--G- CGGG--CGGA T---A-AAAA A--T--AAGT G---A--AG- -TG--TC--T ACC-TT---A
pCaP106            .....TAAT- -GTA-AT--G CC--CT---A A-CG--..-- GG-GG-.-G- AGGGG-CGGA T---A-AAAA A--T--AAGT G---A--AG- -TG--TC--T ACC-TT---A

                   CGGACCTAGC TCATTTCAAA GGTATTGACC CAAGCTATT. CCCCACTTA. ..AATATTTT TCTGGGATCC GGGCC
pEYE101            G--------- ---------- ---------- ---------- ---------- ---------- ---------- -----
pEYE110            T--------- ------T--- ---------T- ---------- ---------- ---------- ---------- -----
pEYE112            G--------- ------T--- ---------- ---------- ---------- ---------- ---------- -----
pGBA102            ---------- ---------- ---------T ---------- ---------- ---------- -T-------- -----
pGBA110            ---------- ----A----- --------TT ---------- TG-------- ---------- ---------- -----
pGBA111            T--------- ------T--- ---------T- ---------- ---------- ---------- -T-------- -----
pGBA112            ---------- ----A----- --------TT ---------- TG-------- ---------- ---------- -----
pGBA113            ---------- ---------- --------TT ---------- ---------- ---------- ---------- -----
pTEL101            G--------- ---------- ---------- ---------- ---------- ---------- ---------- -----
pTEL110            T--------- ---------- ---C------ --G------- ---------- ---------- ---------- -----
pTEL112            G--------- -T-------- ---------- ---------- ---------- ---------- ---------- -----
pCaP104            T-AGGTAG-T CAGG--.GG- AC-T-G-CGT AGG.A---GT TGATGAG--A G---G-G-G GT--AT--GG AAA..
pCaP105            T-AG-TAG-T C-GG---GG- AC-T-GCCGT AGG-A---GT -GATGAG--C CG---G-G-G GT--AT--GG AAA..
pCaP106            T-AG-TAG-T C-GG---GG- AC-T-GCCGT AGG-A---GT -GATGAG--C CG---G-G-G GT--AT--GG AAA..
```

b)
```
                   CCTTTTTGGC CCCTAATTCC TAAACTGTTG GGACCAAAAC TCCCAAAATC AATCCCAACC TTCCTTTTAT GGTCATAAAC CTTGTGTTTA AATTTCATAG A.TTTCTATT
pEYE206            ---------- ---------- ---------- AA-------- ---------- --A-----A- ---------- ---------- ---------- ---------- ----------
pEYE207            ---------- ---------- ---------- A-------- ---------- ---A------ ---------- ---------- ---------- ---------C ----------
pEYE208            ---------- ---------- ---------- -A-------- ---------- ---------A- ---------- ---------- ---------- ---------- ----------
pGBA201            -T-------- ---------- --------C ---------- ---------- -----A---- --------G- --------C- .------CA- ---C------ ----------
pGBA202            -T-----T-- ---------- ----AA-C- -A---TC--- ---------- ---A------ ---------- ---------- ---------- ---------- ----------
pGBA203            -T-------- ---------- -----A--C ------C-- ---------- ---A------ --------G- ---------- .------CA- ---C------ --A-------
pTEL207            .--------- ---------- ---------- ---------- ---------- ---A------ ---------- ---------- ---------- ---------- ----------
pTEL208            .-C------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
pTEL209            .--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
pCaP213            ---------- ---------- ---------- ---------- ---------- ---G-A ---C---G- AA-ATGG--- ------G-AC ---G---G-- --GA--C--A
```
```
                   TACTTAT.AC TAAAGTTATG GTG.CGAAAA CCAAGAATAA TGCTTATTTG ..GGTC.
pEYE206            ---------- ---C------ -----A---- ----A--A-- ---------- -------
pEYE207            ---------- ---C------ ----T----- ------A--- ---------- -------
pEYE208            ---------- ---------- -------A-- ------A--- ---------- -------
pGBA201            ------A--- ---------T ---------- ------C--- ----C-C
pGBA202            ------A--- ---------T ---------- ------C--- ----C--
pGBA203            ------A--- ---------T ---------- ---.------ ------C--- ----C-C
pTEL207            ---C------ ---------- ---CGA---- ---------- ---------- -------
pTEL208            -G-----C-- ---C------ ---------- ---------- ---------- -------
pTEL209            ---C------ ---------- ----G----- ---------- ---------- -------
pCaP213            C---A-C--- AC-------- TCT-.-G-G TT-.C---A-- ---------T TC--C--
```

c)
```
                   CAGGCAACCT TAGGCAAGTG TCTATCATTT GTACCTAGTG GGCAAAAAAG CTTTGCCGTA GGGAAAATTT TATGAACGAA AGAAGGTGGA TGGTAGGGTC ACCCTACCGC
pEYE312            ---------- --A------- C--------- --G------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
pEYE313            ---------- --A------- C--------- --G------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
pGBA314            -------G-- -A-------- C--------- --T------- --------G- -----A---T ---------- ---------- ---------- ---------- ----------
pGBA315            -------G-- ----C----- C-------G- --T------- ---------- ---------- ---------- ---------- ------G--- ---------- ----------
pGBA316            -------.== ==C======= C==-C=C=== ========= C--------- ---------- T===G==== ---------- ---------- ---------- -------C-
pTEL305            ----T----- ---------- ---------- ---------- ---------- ---------- ---------- -G----G--- ------G--- ---------- ----------
pTEL306            ----T----- ---------- ---------- ---------- ---------- ---------- ---------- -G-------- ---------- ---------- -------C-
pTEL307            ----T----- ---------- ---------- ---------- ---------- ---------- ---------- -G-------- ---------- ---------- -------C-
pCaP301            -------T- ----T----- ---------- --T------- --G-C----- ---------- ----GTT--- AT-T--TTG- --T---G--- ---------- ----------
pCaP302            -------T- ---------- --T------- --G-C--G-- ---------C- ----TTT--- ---------- ------A-----CA
pCaP308            -------T- ----T----- ---------- --T------- --G-C----- ---------- ----GTT--- AT-T--TTG- --T---G--- ---------- ----------
```
```
                   TACTTTGGAC CTCAAAAAAA AAAGTTCCGT AACTTTACCC ACCTGGGTAT TTGGGCC
pEYE312            ---------- ---------- ---------- ---------- ---------- -----T-
pEYE313            ---------- ---------- ---------- ---------- ---------- -----T-
pGBA314            ---------- --.,,.----- ---------- ---------- ---------- -------
pGBA315            ---------- --.,.------ ---------- ---------- ---------- -------
pGBA316            A--------- --A------- -.,.------- ---------- ---------- -------
pTEL305            ---------- ---------- ---------- ---------- ---------- -------
pTEL306            A--------- ---C------ ---------- ---------- ---------- -------
pTEL307            ---------- ---------- ---------- ---------- ---------- -------
pCaP301            ---------- T----TT-T- -------T-- ---------- ---------C- ---A-T-
pCaP302            A-G------- -----TT--- ---------- ----T---C- -----T-
pCaP308            ---------- -----TT-T- -------T-- ---------- --------C- ---A-T-
```

Figure 1. Alignment of the nucleotide sequences of the sequenced monomers from *Mytilus* spp. satellite DNAs and the consensus sequence (bold letters) derived from them. **(a)** Satellite DNA type 1. **(b)** Satellite DNA type 2. **(c)** Satellite DNA type 3. In the monomer sequences, bases showing no variation from their respective consensus are shown with dashes (–). Nucleotide differences from the consensus are listed and gaps are shown with dots. Clones are named as follows: pEYE for *M. edulis* from Yerseke; pGBA for *M. galloprovincialis* from Balcobo; pTEL for *M. trossulus* from Esquimalt Lagoon; and pCaP for *M. californianus* from Point No Point. The sequences reported in this article have been deposited in the GenBank database with the following accession numbers: for satellite DNA type 1, AJ420289–AJ420302; for satellite DNA type 2, AJ420303–AJ420312; and for satellite DNA type 3, AJ420758–AJ420768.

Figure 2. Fluorescent in situ hybridization with satellite DNA probes. **(a)** Type 1 DNA located at interstitial and subtelomeric regions in *M. edulis.* **(b)** Type 2 DNA hybridization showing a large number of dispersed dots in *M. galloprovincialis.* **(c)** Type 3 DNA located on three chromosome clusters at subtelomeric regions in *M. trossulus.* Bar=10μm.

Results and discussion

DNA amplifications using the primers designed for the three repetitive mussel DNA sequences yielded multimers in all the individuals tested. One clone from each individual was sequenced and the results are shown in Fig. 1. As observed, type 1 repetitive DNA sequences are 171 bp long, except for *M. californianus* clones; type 2 are 161 bp long, except for the pTEL209 of *M. trossulus* clones; and type 3 repetitive DNA sequences are 167 bp long, excluding *M. galloprovincialis* clones. The A + T content of the repetitive DNAs is approximately 55% for types 1 and 3, and 65% for type 2. These values range within those obtained in other bivalve repetitive DNAs (Ruiz-Lara et al. 1992; Plohl and Cornudella 1996; Clabby et al. 1996; Canapa et al. 2000).

The genomic contents of the three *Apa I* repetitive sequences were determined by comparison of the corresponding hybridization signal intensities on appropriate dot blots (data not shown). As observed in Table 1, type 2 satellite DNA is the most abundant in the four mussel species, whereas type 3 shows low content. These results are very similar to those described in other bivalve species (Plohl and Cornudella, 1996; Canapa et al. 2000). However, although among our mytilids the content of type 2 satellite DNA is the highest, it is still considered low if compared with other species containing this type of sequence, as in the flour beetle *Tenebrio molitor* (Plohl et al. 1992) or in the hermit crab species (*Pagurus pollicaris*) (Fowler and Skinner 1985).

Fluorescent in situ hybridization was performed to locate these repetitive sequences on the chromosomes of *M. edulis*, *M. galloprovincialis*, and *M. trossulus*. Although the fluorescent signal is very weak (owing to low satellite DNA content in these mussel genomes), we have observed that type 1 localizes to interstitial and subtelomeric regions on at least nine chromosome clusters (Fig. 2a); for type 2, the large number of dispersed dots observed reveals a random distribution of several relatively small-sized clusters in all the chromosomes of the three mussel species (Fig. 2b). Type 3 repetitive DNA locates on three chromosome clusters at subtelomeric regions (Fig. 2c). The distribution of these repetitive DNAs on the chromosomes of each of the mussel species examined, together with their respective copy numbers, confirm that repetitive DNAs types 1, 2, and 3 are satellite DNAs. Karyotypes of *M. edulis*, *M. galloprovincialis*, and *M. trossulus* (Martínez-Lage et al. 1995, 1996) show great similarity in their chromosome morphology. This fact, the existence of weak fluorescent signals, and the similar chromosome locations of these satellite DNA clusters impede differentiation of these three mussel species on the basis of the chromosomal distribution of the satellite DNAs types 1, 2, and (or) 3.

Table 2. Summary of base changes in the satellite DNAs of the four species of mussels.

	Change				Ratio	
	(a) G,C $\rightarrow$ X	(b) A,T $\rightarrow$ X	(c) ts	(d) tv	a:b	c:d
Satellite 1						
M. edulis	8	2	9	1	4.00	9.00
M.galloprovincialis	11	4	10	5	2.75	2.00
M. trossulus	10	3	9	4	3.33	2.25
M. californianus	56	48	62	42	1.17	1.48
Satellite 2						
M. edulis	8	3	5	6	2.67	0.83
M.galloprovincialis	8	12	9	11	0.67	0.82
M. trossulus	3	4	4	3	0.75	1.33
M. californianus	9	26	15	20	0.35	0.75
Satellite 3						
M. edulis	2	2	4	0	1.00	∞
M.galloprovincialis	8	11	5	14	0.73	0.36
M. trossulus	3	4	2	5	0.75	0.40
M. californianus	15	18	11	22	0.83	0.50

Note: **X**, any nucleotide; **ts**, transition; **tv**, transversion.

In some organisms, the existence of a close correlation between heterochromatin distribution and chromosome location of satellite DNAs has been demonstrated, suggesting that satellite DNA is one of the major constituents of heterochromatin (John and Miklos 1979; Fanning et al. 1988). However, in these mussel species there are small amounts of heterochromatin that distributes at telomeric regions (Martínez-Lage et al. 1995, 1996), so that the correlation of satellite DNA – heterochromatin would be, in mussels, directly related to satellite DNA types 1 and 3. In this sense, equilocal distribution of heterochromatin and satellite DNAs suggests that telomeres are the presumptive sites where heterochromatin amplification tends to be initiated, and that the telomeres are the regions from which the heterochromatin sequences are dispersed (Schweizer and Loidl 1987; John 1988). Likewise, it was proposed that chromosome evolution may be associated with the activity of rapidly evolving repetitive DNA sequences, such as tandem repeats (Wichman et al. 1991), and specifically that tandemly repeated sequences may facilitate chromosome rearrangements (Bradley and Wichman 1994; Garagna et al. 1997). However, only "active" satellite DNA, which is actively involved in the processes of expansion, contraction, and mobilization, could promote changes in the karyotype morphology (Slamovits et al. 2001). Taking this assumption into account, type 2 satellite DNA, with its dispersed distribution, could be more directly related with the chromosome rearrangements observed in these mussels.

Sequence analysis of the cloned satellite monomers reveals species homology of 91% in all cases, indicative of the high degree of conservation of these sequences. The distribution of type 2 satellite DNA along the mussel chromosomes could be explained as the outcome of the high efficiency of the spreading mechanism (gene conversion and unequal crossing-over) relative to mutation events, as other authors have suggested in other species (Plohl et al. 1992; Bruvo et al. 1995; Mestrovic et al. 2000). On the other hand, interspecific homologies among *M. edulis*, *M. galloprovincialis*, and *M. trossulus* display values exceeding 90%. In *M. californianus*, satellite DNA types 2 and 3 show homology of 76 and 85%, respectively. These values are slightly lower than those from the other three mussel species, but similar to those reported for other bivalve species (Clabby et al. 1996; Muchmore et al. 1998). However, type 1 satellite DNA in *M. californianus* displays a homology value of 37%, owing to the existence of approximately 100 variable sites. Sequence comparison with EMBL and GenBank nucleic acid databases revealed significant homology only with the repetitive DNA sequence of *M. edulis* reported by Ruiz-Lara et al. (1992).

Analysis of sequence variability of satellite DNAs relative to the consensus sequences shows that single-point mutations are the major contributors to divergence of these satellites. The point mutations spread randomly throughout the whole sequence, although, in some cases, the same substitution occurs at a particular position in more than one monomer. We compared each monomeric sequence with the consensus sequence to determine the number and type of substitutions since satellite amplifications or homogenization events have occurred (Table 2). Only one sample was counted when two or more identical substitutions were encountered at the same position in different clones from the same species. The others could have arisen from localized homogenization events and not from DNA replication errors, as pointed out by other authors (Smith 1976; Fanning et al. 1989; Ugarkovic et al. 1992). The changes involving A or T nucleotides are fewer than expected considering the A + T richness of the sequences (55–65%), especially for satellite DNA type 1. The ratio between transitions and transversions in satellite type 1 is higher than 1.0, in satellite type 2 it is near 1.0, and in satellite type 3 it is lower than 1.0 (except for the *M. edulis* clones). The ratios obtained for satellite type 3 are similar to those from primates (Fanning et al. 1989), canids (Fanning 1989) and *Drosophila* (Strachan et al. 1985), whereas those obtained for satellite type 2 are similar to the flour beetle satellites (Plohl et al. 1992; Ugarkovic et al. 1996). Taking into account the work of Strachan et al. (1985), our results suggest that selection involving mutations could be absent in type 3 satellite DNA, because there is an excess of transversions over transitions. As these authors point out, each nucleotide position is susceptible to mutation, and the

resulting variant can spread as a consequence of family turnover without strong selective or genomic constraints. Otherwise, in type 2 satellite DNA there may exist low selection pressure or, alternatively, no selection at all, whereas in satellite type 1 the selection process could be active.
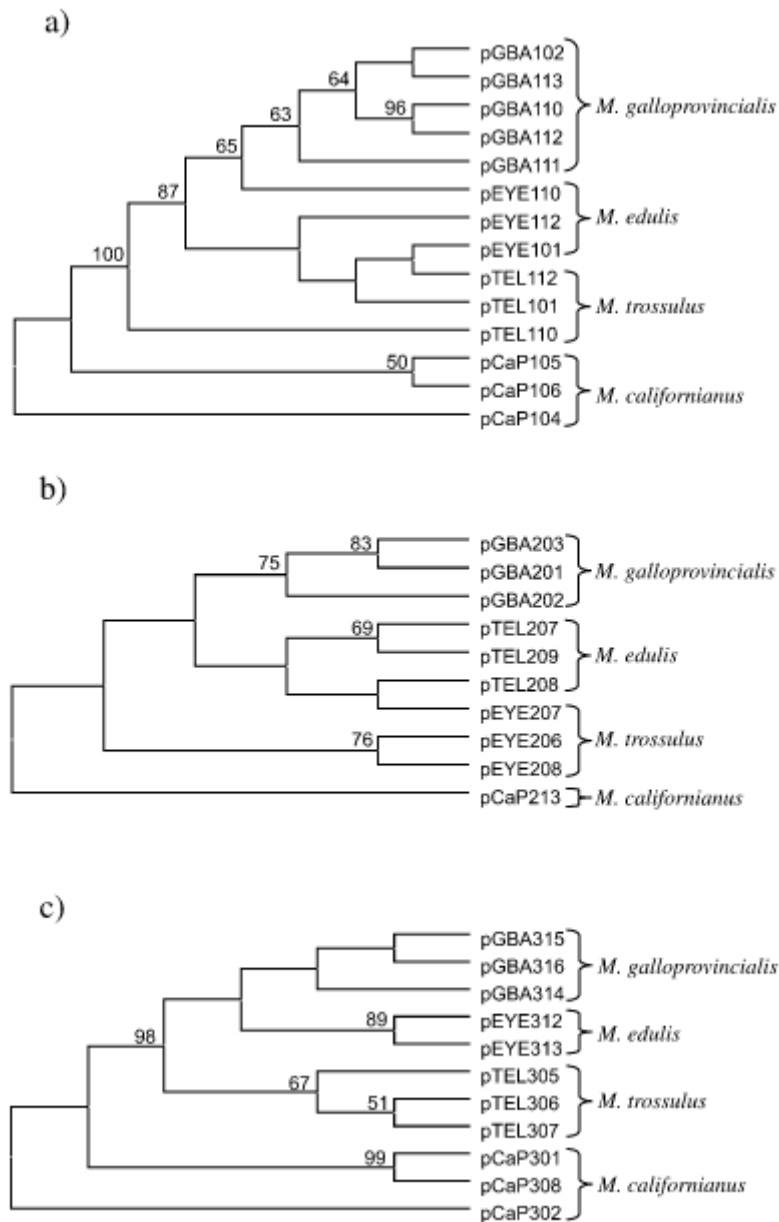

Figure 3. Single maximum-parsimony trees from the satellite DNAs. Bootstrap values over 50% are shown above the branches. **(a)** Satellite DNA type 1 (L = 128; CI = 0.930; RI = 0.558; RCI = 0.890). **(b)** Satellite DNA type 2 (L = 74; CI = 0.838; RI = 0.647; RCI = 0.542). **(c)** Satellite DNA type 3 (L = 60; CI = 0.850; RI = 0.830; RCI = 0.706). **L**, tree length; **CI**, consistency index; **RI**, retention index; **RCI**, rescaled consistency index.

To examine evolutionary relationships among these satellite DNAs, we carried out a phylogenetic analysis. The Kimura two-parameter distances (data provided upon request) among monomer clones of the three satellite DNAs are between 0.02 and 0.11, except for the *M. californianus* clones with respect to the other *Mytilus* spp. clones, which show a genetic distance of approximately 1.25 for satellite type 1, 0.28 for type 2, and 0.17 for type 3. A MEGA analysis was employed to construct different dendograms using several procedures: UPGMA (Sneath and Sokal 1973), neighbor-joining (Saitou

and Nei 1987), and maximum-parsimony methods. The topologies of all trees for each satellite DNA were congruent, except for certain unresolved nodes, which were produced by the arrangement of some monomeric units. For this reason, we only show the maximum-parsimony trees (Fig. 3). As expected from the comparison of sequence data, there are two clusters for satellite type 1 and type 3, one of the cluster groups being the *M. californianus* clones and the other the rest of *Mytilus* spp. clones. This is supported by at least 98 out of the 100 bootstrapped trees. The bootstrap values obtained for the rest of the nodes were generally low. Clones of *M. galloprovincialis* cluster together for the three satellite DNAs, whereas clones of *M. edulis* and *M. trossulus* are not separated as clearly as those of *M. galloprovincialis*. However, data from the phylogenetic analysis using mitochondrial 16S RNA and cytochrome-*c* oxidase grouped *M. californianus* and *M. trossulus* separately, but did not separate *M. edulis* and *M. galloprovincialis* clones (Geller et al. 1993; Rawson and Hilbish 1995, 1998; Hilbish et al. 2000). On the other hand, the phylogenetic analysis of 18S RNA sequences did not define a bootstrap-supported cluster for *M. trossulus* , and only *M. californianus* sequences were differentiated (Kenchington et al. 1995; Distel 2000). In this sense, our results reveal that satellite DNA type 3 is the best species-specific marker and shows the smallest transition–transversion ratio among clones, this value being close to 0.5. This could suggest that concerted evolutionary processes are separating satellite type 3 but do not separate types 1 and 2.

In conclusion, the sequence analyses suggest the existence of a mechanism of concerted evolution for the three satellite DNAs studied, where selection is acting in different ways. The high frequency of a nucleotide substitution in the same position in the different clones analyzed reveals a process of molecular expansion of these monomeric variants. Furthermore, the phylogenetic relationships among clones reflect that *M. californianus* is the most divergent, appearing in a cluster that is different from that of *M. edulis*, *M. galloprovincialis*, and *M. trossulus.*

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*: 3389–3402.

Bradley, R.D., and Wichman, H.A. 1994. Rapidly evolving repetitive DNAs in a conservative genome: a test of factors that affect chromosomal evolution. *Chromosome Res. 2*: 354–360.

Brutlag, D.L. 1980. Molecular arrangement and evolution of heterochromatic DNA. *Ann. Rev. Genet. 14*: 121–144.

Bruvo, B., Plohl, M., and Ugarkovic, D. 1995. Uniform distribution of satellite DNA variants on the chromosomes of tenebrionid species *Alphitobius diaperinus* and *Tenebrio molitor. Hereditas, 123*: 69–75.

Canapa, A., Barucca, M., Cerioni, P.N., and Olmo, E. 2000. A satellite containing CENP-B box-like motifs is present in the Antarctic scallop *Adamussium colbecki*. *Gene, 247*: 175–180.

Charlesworth, B., Sniegowski, P., and Stephan, W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* (London), *371*: 215–220.

Clabby, C., Goswami, U., Flavin, F., Wilkins, N.P., Houghton, J.A., and Powell, R. 1996. Cloning, characterization and chromosomal location of satellite DNA from the Pacific oyster, *Crassostrea gigas*. *Gene, 168*: 205–209.

Distel, D.L. 2000. Phylogenetic relationships among Mytilidae (Bivalvia): 18S rRNA data suggest convergence in mytilid plans. *Mol. Phyogenet. Evol 15*: 25–33.

Estabrooks, S.L. 1999. The telomeres of the bay scallop, *Argopecten irradians* (Lamarck*). J. Shellfish Res. 18*: 401–404.

Fanning, T.G. 1989. Molecular evolution of centromere-associated nucleotide sequences in two species of canids. *Gene, 85*: 559–563.

Fanning, T.G., Modi, W.S., Wayne, R.K., and O'Brien, S.J. 1988. Evolution of heterochromatin-associated satellite DNA loci in felids and canids (Carnivora). *Cytogenet. Cell Genet., 48*: 214–219.

Fanning, T.G., Seuanez, H.N., and Forman, L. 1989. Satellite DNA sequences in the neotropical marmoset *Callimico goeldii* (Primates, Platyrrhini). *Chromosoma, 98*: 396–401.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using bootstrap. *Evolution, 39*: 783-791.

Garagna, S., Pérez-Zapata, A., Zuccotti, M., Marcheretti, S., Marziliano, N., Redi, C.A., Aguilera, M., and Capanna, E. 1997. Genome composition in Venezuelan spinyrats of the genus *Proechimys* (Rodentia, Echimydae). I. Genome size, C-heterochromatin and repetitive DNAs *in situ* hybridization patterns. *Cytogenet. Cell Genet. 78*: 36–43.

Geller, J.B., Carlton, J.T., and Powers, D.A. 1993. Interspecific and intrapopulation variation in mitochondrial ribosomal DNA sequences of *Mytilus* spp. (Bivalvia: Mollusca). *Mol. Mar. Biol. Biotechnol. 2*: 44–50.

González-Tizón, A., Martínez-Lage, A., Rego, J.I., Ausió, J., and Méndez, J. 2000. DNA content, karyotypes and chromosomal location of 18S-5.8S-28S ribosomal loci in some species of bivalve molluscs from the Pacific Canadian coast. *Genome, 43*: 1065–1072.

Gosling, E.M. 1992. Genetics of *Mytilus*. In *The mussel Mytilus: ecology, physiology, genetics and culture*. Edited by E.M. Gosling. Elsevier Press, Amsterdam, The Netherlands. pp. 309–382.

Hilbish, T.J., Mullinax, A., Dolven, S.I., Meyer, R.K., Koehn, R.K., and Rawson, P.D. 2000. Origin of the antitropical distribution pattern in marine mussels (*Mytilus* spp.): routes and timing of transequatorial migration. *Mar. Biol. 136*: 69–77.

Hoeh, W.R., Stewart, D.T., Saavedra, C., Sutherland, B.W., and Zouros, E. 1997. Phylogenetic evidence for role-associated mitochondrial DNA in *Mytilus* (Bivalvia: Mytilidae). *Mol. Biol. Evol. 14*: 959–967.

John, B. 1988. The biology of heterochromatin. In *Heterochromatin: molecular and structural aspects*. Edited by R.S. Verma. Cambridge University Press, Cambridge, U.K. pp. 1–147.

John, B., and Micklos, G.L.G. 1979. Functional aspects of satellite DNA and heterochromatin. *Internatl. Rev. Cytol. 58*: 1–114.

Karpen, G.H. 1994. Position-effect variegation and the new biology of heterochromatin. *Curr. Opin. Genet. Dev. 4*: 281–191.

Kenchington, E., Landry, D., and Bird, C.J. 1995. Comparison of taxa of the mussel *Mytilus* (Bivalvia) by analysis of the nuclear small-subunit rRNA gene sequence. *Can. J. Fish. Aquat. Sci. 52*: 2613–2620.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Biol. 16*: 111–120.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinfomatics, 17*: 1244–1245.

Maiorano, D., Cece, R., and Badaracco, G. 1997. Satellite DNA from the brine shrimp *Artemia* affects the expression of a flanking gene in yeast. *Gene, 189*: 13–18

Martínez-Lage, A., González-Tizón, A., and Méndez, J. 1995. Chromosomal markers in three species of the genus *Mytilus* (Mollusca: Bivalvia). *Heredity, 74*: 369–375.

Martínez-Lage, A., González-Tizón, A., and Méndez, J. 1996. Chromosome differences between European mussel populations (genus *Mytilus*). *Caryologia, 49*: 343–355.

McDonald, J.H., Seed, R., and Koehn, R.K. 1991. Allozyme and morphometric characters of three species of *Mytilus* in the northern and southern hemispheres. *Mar. Biol. 111*: 323–335.

McLean, K.W., and Whiteley, A.H. 1973. Characteristics of DNA from the oyster, *Crassostrea gigas*. *Biochim. Biophys. Acta, 335*: 35–41.

Mestrovic, N., Mravinac, B., Juan, C., Ugarkovic, D., and Plohl, M. 2000. Comparative study of satellite sequences and phylogeny of five species from the genus *Palorus* (Insecta, Coleoptera). *Genome, 43*: 776–785.

Miklos, G.L. 1985. Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. In *Molecular evolutionary genetics*. Edited by J.R. McIntyre. Plenum Press, New York. pp. 241–321.

Modi, W.S., Gallagher, D.S., and Womack, J.E. 1996. Evolutionary histories of highly repeated DNA families among the Artiodactyla (Mammalia). *J. Mol. Evol. 42*, 337–349.

Muchmore, M.E., Moy, G.W., Swanon, W.J., and Vacquier, V.D. 1998. Direct sequencing of genomic DNA for characterization of a satellite DNA in five species of eastern pacific abalone. *Mol. Mar. Biol. Biotech. 7*: 1–6.

Nabegama, M., Kubota, S., and Kohno, S. 2000. Concerted evolution of a highly repetitive DNA family in Eptatretidae (Cyclostoma, Agnatha) implies specifically differential homogenization and amplification events in their germ cells. *J. Mol. Evol. 50*: 154–169.

Plohl, M., Borstnik, B., Lucijanic-Justic, V., and Ugarkovic, D. 1992. Evidence for random distribution of sequence variants in *Tenebrio molitor* satellite DNA. *Genet. Res. 60*: 7–13.

Plohl, M., and Cornudella, L. 1996. Characterization of a complex satellite DNA in the mollusc *Donax trunculus*: analysis of sequence variations and divergence. *Gene, 169*: 157–164.

Plohl, M., and Cornudella, L. 1997. Characterization of interrelated sequence motifs in four satellite DNAs and their distribution in the genome of the mollusc *Donax trunculus*. *J. Mol. Evol. 44*: 189–198.

Rawson, P.D., and Hilbish, T.J. 1995. Evolutionary relationships among the male and female mitochondrial DNA lineages in the *Mytilus edulis* species complex. *Mol. Biol. Evol. 12*: 893–901.

Rawson, P.D., and Hilbish, T.J. 1998. Asymmetric introgression of mitochondrial DNA among European populations of blue mussels (*Mytilus* spp.). *Evolution, 52*: 100–108.

Rice, E.L., and Bird, C.J. 1990. Relationships among geographically distant population of *Gracilaria verrucosa* (Gracilariales, Rhodophyta) and related species. *Phycologia, 29*: 501–510.

Rodríguez-Juíz, A.M., Torrado, M., and Méndez, J. 1996. Genome-size variation in bivalve molluscs determined by flow cytometry. *Mar. Biol. 126*: 489–497.

Ruíz-Lara, S. 1993. *Análisis del ADN satélite en el mejillón Mytilus edulis*. PhD thesis, Universidad Central de Barcelona, Spain.

Ruíz-Lara, S., Prats, E., Sainz, J., and Cornudella, L. 1992. Cloning and characterization of a highly conserved satellite DNA from the mollusc *Mytilus edulis*. *Gene, 117*: 237–242.

Rychlik, W., and Rhoads, E. 1989. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucleic Acids Res. 17*: 8543–8551.

Sainz, J., and Cornudella, L. 1990. Preservation of a complex satellite DNA in two species of Echinoderms. *Nucleic Acids Res. 18*: 885–890.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol. 4*: 406–425.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A. 74*: 5463–5467.

Schweizer, D., and Lloid, J. 1987. A model for heterochromatin dispersion and the evolution of C-band patterns. In *Chromosomes today*. Vol. 9. Edited by D.L. Hayman, R.H. Rofe, and P.J. Sharp. Allen & Unwin, London, U.K. pp. 61–74.

Smith, G.P. 1976. Evolution of repeated DNA sequences by unequal crossovers. *Science* (Washington, D.C.), *191*: 528–535.

Sneath, P.H., and Sokal, R.R. 1973. *Numerical taxonomy*. W.H. Freeman, San Francisco, Calif.

Strachan, T., Webb, D., and Dover, G. 1985. Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J. 4*: 1701–1708.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res. 24*: 4876–4882.

Ugarković, D., Plohl, M., Lucijaniæ-Justiae, V., and 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.

Wichman, H.A., Payne, C.T., Ryder, O.A., Hamilton, M.J., Maltbie, M., and Baker, R.J. 1991. Genomic distribution of heterochromatic sequences in equids: implication to rapid chromosomal evolution. *J. Hered. 82*: 369–377.

Wijers, E.R., Zijlstra, C., and Lenstra, J.A. 1993. Rapid evolution of horse satellite DNA. *Genomics, 18*: 113–117.

---

[*] fina@udc.es