

Gene

September 2004; 339 : 181-188

<http://dx.doi.org/10.1016/j.gene.2004.06.049>

© 2004 Elsevier B.V. All rights reserved

Archimer <http://www.ifremer.fr/docelec/>

Archive Institutionnelle de l'Ifremer

The molecular phylogeny of oysters based on a satellite DNA related to transposons

Inmaculada López-Flores^{a*}, Roberto de la Herrán^a, Manuel A. Garrido-Ramos^a, Pierre Boudry^b, Carmelo Ruiz-Rejón^a and Manuel Ruiz-Rejón^a

^aDepartamento de Genética, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

^bIFREMER, Laboratoire de Génétique et Pathologie, 17390 La Tremblade, France

*: Corresponding author : Tel.: +34 958 243080; fax: +34 958 244073. ilopez@ugr.es

Abstract: We have analysed a centromeric satellite DNA family that is conserved in several commercial and non-commercial oyster species (*Ostrea edulis*, *O. stentina*, *Crassostrea angulata*, *C. gigas*, *C. gasar*, *C. ariakensis*, *C. virginica* and *C. sikamea*). This satellite DNA family is composed of AT-rich repeat sequences of 166±2 bp and presents a 9-bp motif similar to the mammalian CENP-B box. The homology of oyster HindIII satellite DNA with satellite DNAs from other bivalves and its relation to a part of a mobile element suggest the existence of an ancient transposable element as a generating unit of satellite DNA in bivalve molluscs. Taking advantage of its degree of conservation in oyster species, we have used this element as a taxonomic marker. This marker clearly supports a high degree of differentiation between *O. edulis* and *O. stentina*, and, conversely, upholds the contention that *C. gigas* and *C. angulata* are the same species. Finally, we have used HindIII satellite DNA as a phylogenetic marker between these species, revealing two clades, one formed by Asiatic species (*C. angulata*, *C. gigas* and *C. ariakensis*) and another by the European, American and African species (*O. edulis*, *C. virginica* and *C. gasar*, respectively).

Keywords: *Ostrea*; *Crassostrea*; CENP-B box; Genetic identification; Satellite DNA origin

1. Introduction

One of the most characteristic features of the eukaryotic genomes is the presence of a variety of repetitive sequences of several types. Among repetitive DNAs, tandemly arranged highly repeated sequences or satellite DNA (Ugarkovic and Plohl, 2002) are the main constituents of the heterochromatin. Thus, different satellite DNA families equilocally accumulate in different regions of the eukaryote chromosomes, mainly at centromeres and in subtelomeric regions (Charlesworth et al., 1994). Several questions remain unresolved concerning the role of these types of sequences within genomes, notably those related to the formation and expansion of a satellite DNA family and its possible function.

Forces governing satellite DNA appearance and amplification are not well understood. An accepted hypothesis suggests the continuous evolution of satellites from pre-existing satellites, through replication slippage and unequal crossing-over mechanisms (Ugarkovic and Plohl, 2002). However, based on recent data, alternative hypotheses are plausible. Thus, new data support the idea that some satellite DNA families originated from retroelements (Batistoni et al., 1995; Kapitonov et al., 1998). Such satellite DNAs could have originated from interspersed retrotansposons by means of unequal crossing-over (Kapitonov et al., 1998), although alternative mechanisms might be operating.

Satellite DNA sequences within genomes, as opposed to the negative view of these sequences as junk DNA, may participate in several cellular processes, as proposed for centromeric DNAs. Centromere-associated satellite DNA could establish a structural context for the action of conserved motifs, such as CENP-B box of alpha-satellite DNA (Schueler et al., 2001). On the other hand, due to the higher rates of sequence change than in any other parts of the eukaryotic genomes, satellite DNA has been used as a molecular marker for taxonomic and phylogenetic studies (Stepien and Kocher, 1997; Urganović and Plohl, 2002). In particular, among molluscs, comparative studies of satellite DNA sequences have been

made in abalone (Muchmore et al., 1998), scallop (Canapa et al., 2000) and mussels (Martínez-Lage et al., 2002).

We have analysed an oyster's satellite DNA family, concluding that it has been derived from a retroelement. Given its presence in several oyster species, we have used it as a marker to clarify several controversial taxonomic aspects. Finally, we have tried to use these sequences as a marker for phylogenetic inference.

2. Materials and methods

2.1. Sampling and DNA extraction

We analysed a total of 8 oyster species obtained from different locations (Table 1). Genomic DNA from abductor muscle of freshly captured oysters (*O. edulis* and *O. stentina*) was purified as in Winnepenninckx et al. (1993). The biological material from *C. angulata*, *C. gigas* and *C. gasar* had been classified by PCR-RFLP of mitochondrial cytochrome oxidase I gene sequence (mtCOI; Boudry et al., 1998; Lapègue et al., 2002). The remaining samples (*C. ariakensis*, *C. sikamea* and *C. virginica*) were identified according to the information provided by the suppliers. Some of these species are of considerable commercial interest: *O. edulis* (European flat, or edible, oyster), *C. virginica* (American, or eastern, oyster), *C. gigas* (Pacific, or Japanese, oyster) and *C. angulata* (Portuguese oyster).

2.2. Identification and amplification of satellite DNA

Monomeric units of satellite DNA were located within genomic *O. edulis* DNA by digestion of mixed DNA from 2-3 individuals with a variety of restriction endonucleases and electrophoresis in agarose gel with ethidium bromide. The most prominent band visible under UV light was excised from the gel and purified using GFX™ PCR DNA and

Gel Band Purification Kit (Amersham Biosciences). Purified fragments were then ligated to the pUC19 vector and cloned in competent DH5- α cells (Gibco BRL). Minipreparations were made using Perfectprep[®] Plasmid Mini (Eppendorf) and recombinant plasmid carrying the monomeric sequences were identified after screening with the purified band by dot-blot hybridisation. The probe was labelled using the non-radioactive method ECL[™] Direct Nucleic Acid Labelling and Detection Systems (Amersham Biosciences). The same probe was used in Southern-blot analysis to check the tandem organization of the repetitive sequence in the genome of *O. edulis*.

The cloned monomeric sequences were aligned by Megalign program of DNASTAR package (LASERGENE). The alignment was used to design the primers OsedAdir (5'-TCGGATGGGGCCACAATA-3') and OsedArev (5'-TGGTTCTTGAGAAGAAGA-3') which were used in PCR amplification of the satellite DNA family in the remaining species. Amplification was carried out from 25-100 ng oyster DNA from 2-3 individuals as a template in 50 μ l volume containing 100 ng of each primer, 10 mM dNTPs, and 1.25 units of Taq polymerase (Amersham Biosciences) in Taq buffer. The PCR was conducted with an initial denaturation at 94°C for 5 min, 40 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 1 min, and a final step at 72°C for 10 min in a GeneAmp 2700 Applied Biosystems thermocycler. The amplified products were separated electrophoretically, purified from agarose gels and cloned in a pGEM-T Easy Vector System (Promega). The protocols were carried out following the manufacturer's recommendations for each kit.

2.3. Sequence analysis

Multiple-sequence alignments were performed again using Megalign program of DNASTAR package (LASERGENE). Neighbour-joining phylogenetic analyses were carried

out using the MEGA package (Kumar et al. 2001). The distance method was used due to the features of satellite DNA families: short length, high variability between the monomeric units of the different species and the unequal rates of sequence change among lineages (Kumar et al., 2001). Genetic distances were calculated according to Kimura's 2-parameters (Kimura, 1980). The study of the transition stages of concerted evolution of the *HindIII* tandem repeat family was made according to Strachan et al. (1985). This analysis compares the variation pattern at each nucleotide position (considered independently) of a DNA shared by two species; the sequences from two species are compared and each nucleotide position is classified in one of six different homogenisation stages (Transition Stage 1-6). For a given nucleotide position, the TS1 shows complete homogeneity across all clones of a pair of species while the new mutations gradually spread throughout the sequences of one species, these being termed TS2, TS3, TS4, TS5 (mutations in a minority of positions in clones of one species while the other species remains homogeneous for the base in the position to positions diagnostic for each species), and TS6 which groups all subsequent mutations beyond this point.

3. Results and discussion

3.1. Characterization and conservation of the HindIII satellite DNA

We searched for satellite DNA sequences within the genome of *O. edulis* by means of genomic-DNA digestions with 20 restriction endonucleases (*AluI*, *BamHI*, *BclII*, *BglII*, *CfoI*, *DraI*, *EcoRI*, *HaeIII*, *HindII*, *HindIII*, *HinfI*, *HpaII*, *MspI*, *NdeII*, *PstI*, *PvuII*, *RsaI*, *SacI*, *SspI* and *XbaI*) and subsequent electrophoresis on agarose gels of the generated DNA fragments. Among the different enzymes used, the digestion performed with *HindIII*, *DraI*, *BclII*, *HinfI* and *HaeIII* generated a canonical pattern of satellite DNA sequences visible in the gels as an intense band of about 170 bp over the rest of the DNA smear. We followed a

conventional cloning approach for the 170-bp fragments and we obtained the sequences of a total of 8 monomeric units of the *Hind*III satellite DNA from *O. edulis*. After a search of the EMBL/GenBank databases, we found that these sequences are closely similar to the *Hae*III satellite DNA previously cloned from the oyster species *C. gigas* (Clabby et al., 1996). This satellite DNA is located at the centromeres of the chromosomes of that species (Wang et al., 2001).

Although different, a related satellite DNA was also found within the genome of *Donax trunculus* (Plohl and Cornudella, 1996) (Fig. 1A). To analyse the degree of conservation of this satellite DNA, we searched the genomes of several oyster species for its presence. Thus, by means of dot-blot and of Southern-blot hybridisations, using the *Hind*III monomeric units of *O. edulis* as a probe, we demonstrated the presence of the *Hind*III satellite DNA within the genomes of several oyster species of the genera *Ostrea* (*O. stentina*) and *Crassostrea* (*C. angulata*, *C. gigas*, *C. gasar*, *C. ariakensis* and *C. virginica*) (all species studied except *C. sikamea*; see below for this one). Therefore, we designed a pair of specific primers for PCR amplification of the *Hind*III satellite DNA from these species. The PCR products were cloned and 4-8 monomers from each species were sequenced and analysed. We have obtained the sequence of a total of 46 monomers from the total of the 7 species.

A remarkable feature of this satellite DNA in all the species is that the repeat units possess a list of several molecular specifics, which are in common with satellite DNAs located in eukaryotic centromeres. Thus, the sequence length of the monomers was 166 ± 2 bp and they were rich in AT base pairs (63% on average). Several runs of A/T bp are visible. These two characteristics are related to the capacity of the repeat unit to acquire stable curvatures, this being an important feature for centromeric DNAs (Radic et al., 1992). Furthermore, we found that a highly characteristic sequence motif of 9 bp (TATGAAAGC) is conserved (Fig.1A). It has been proposed that this motif could have an

important role in the centromeric function as it has been found to be conserved in several vertebrate and invertebrate centromeric DNA satellites and is related to the core sequence of the CENP-B centromeric motif of primates and to the CDEIII region of *Saccharomyces cerevisiae* (Canapa et al., 2000). These characteristics together with the degree of conservation of this satellite DNA and the location at the centromeres of the chromosomes of the oysters (Wang et al., 2001) might indicate a role of this satellite DNA in the centromeric function.

A second remarkable homology was found for the *Hind*III sequences of the oysters with a MITE element of the *Zea mays* genome (56%; Fig. 1B). In the same way, the comparison of the *Hind*III sequences with a recently reported MITE-like transposable element found within the genome of the American oyster *C. virginica* (Gaffney et al., 2003) revealed the homology between the two elements (Fig. 1C). Specifically, the *O. edulis* *Hind*III monomer is 58% similar to the CvA element of *C. virginica* (Fig. 1D). Gaffney et al. (2003) found this homology also between CvA and the centromeric satellite DNA of *C. gigas*. This transposable element was also found in the bivalve *Anadara trapezia* (Gaffney et al., 2003).

The homology between the oyster *Hind*III satellite DNA and the satellite DNAs of other bivalve species such as *Donax trunculus*, and the homology of all of them with the transposable element CvA characterised in *C. virginica* (Gaffney et al., 2003), supports the existence of an ancient transposable element which would have acted as a generating element of satellite DNAs in bivalve molluscs. A homologous transposable element to CvA might be present in the genome of oysters. In fact, when we used *C. sikamea* genomic DNA as a template for PCR with the *Hind*III primers, we amplified a homologous element to CvA instead of *Hind*III satellite (data not shown).

Many reports have demonstrated that diverse mobile elements are stable structural components of the heterochromatic regions of the eukaryotic genomes (reviewed in

Dimitri and Junakovic, 1999). In fact, formation of heterochromatin appears to have some sort of sequence requirement, which is met by at least some transposable elements (Dimitri and Junakovic, 1999). However, until now, little evidence has supported the idea that retroelements may serve as a basic core of satellite monomers, and, therefore, a relationship between retroelements and satellite DNA formation. In fact, recent reports indicate that satellite DNAs are derived from retroelements such as SINEs (Batistoni et al., 1995) or LINEs (Kapitonov et al., 1998). Although satellite DNAs could be derived from short fragments "captured" when retroelements could be acting in repair mechanisms, the most plausible explanation (and certainly this can be applied to the oyster satellite DNA) would be that these types of satellite DNAs originated by means of unequal crossing-over from interspersed retroelements (Kapitonov et al., 1998).

3.2. The use of *HindIII* satellite DNA as a molecular marker

Figure 2 shows the relationships between all the 46 *HindIII* monomeric units analysed in the present paper. As can be seen in this figure, the satellite monomers group together by taxonomic affinity in most cases. However, the sequences of *Ostrea edulis* are not only differentiated from the sequences of the rest of sequences of *Crassostrea* species but also of *O. stentina*. At the larva stages *O. edulis* (the European flat oyster) is very difficult to distinguish from *O. stentina*, a species without commercial interest but coexisting in the wild with the former (Amezcuca et al., 1999). This similarity poses a serious problem when wild spat of *O. edulis* are collected for farming. Some molecular data (allozymes and mitochondrial DNA) (Amezcuca et al., 1999; Comesaña et al., 2001) have begun to provide tools to solve the problem. The *HindIII* satellite DNA provides an additional marker for their taxonomic identification because these sequences are highly differentiated between *O. edulis* and *O. stentina*. In fact, the most unexpected result presented in the Fig. 2 is that the species of *Ostrea* are not grouped together, and this result can be explained by the

classification in which *O. edulis* and *O. stentina* are grouped in different genera—*O. edulis* in *Ostrea* and *O. stentina* in *Ostreola* (Harry, 1985). However, the apparent close relationship between *O. stentina* and *C. gigas* and *C. angulata* species remains unexplained.

One more serious exception with respect to the taxonomic affinity of the *Hind*III sequences is that of *C. gigas* and *C. angulata*, for which the tree in Fig. 2 shows two clades. The reason for this is the presence of two *Hind*III satellite DNA subfamilies within these two species. Thus, we divided the sequences from these two species into two subgroups: those of the subfamily 1 and those of the subfamily 2. The difference between the two sequence types is based on the existence of 16 fixed differences (Fig. 3) that appear as diagnostic sites for each subfamily. However, even between these two subfamilies, there is no difference between the two species. In fact, within each clade, the sequences of the two species are intermixed. Current data on *C. angulata* and *C. gigas* do not clearly differentiate between them. Morphological studies in adults and spats (Ranson, 1960), fertilization experiments (Huvet et al., 2001, Huvet et al., 2002), karyotype analysis (Leitão et al., 1999) and allozyme markers (Buroker et al., 1979; Mattiucci and Villani, 1983) place the two taxa in the same species. Nevertheless, other data (ecophysiological characteristics) support the idea that they are two different species, such as the high capacity for nutrient assimilation and growth (see Haure et al., 2003), low parasitic-disease susceptibility, and high salinity and temperature tolerance of *C. gigas* with respect to *C. angulata* (Haure et al., 2003), combined with genetic differences in mtCOI (Ó Foighil et al., 1998; Boudry et al., 1998). However, these latter results are not inconsistent with the possibility of only one species. In fact, in *C. virginica* populations from the Gulf of Mexico and the Atlantic coast of America show differences in ecophysiological characteristics and mitochondrial haplotypes, and they are not considered two different species (McDonald et al., 1996). Another factor to consider is the lack of reproductive barriers between the two

taxa. Again, this is not completely conclusive. For example, mussels *M. edulis*, *M. galloprovincialis* and *M. trossulus* can hybridise and they are considered three different species (López et al., 2002). In our case, *Hind*III satellite DNA cannot differentiate these two taxa. Hence, our results are consistent with previous morphological, genetic and hybridisation data cited above showing the very close relationship between *C. angulata* and *C. gigas* and indicating that they are not in fact two different species.

With respect to the tree of the Fig. 2, its topology is noteworthy, since the two subfamilies from *C. angulata* and *C. gigas* are grouped in two different clusters, associated with the sequences from *C. ariakensis* the subfamily 1, and with *O. stentina* sequences the subfamily 2 (Fig. 2). The sequences of *O. edulis*, *C. virginica* and *C. gasar* are grouped in a sub-clade which constitute a clade with the sequences of *C. ariakensis* and subfamily 1 from *C. angulata* and *C. gigas*. In fact, the positions that define the subfamilies in *C. angulata* and *C. gigas* are also present in the sequences of the remaining species: *O. stentina* have the ones corresponding to subfamily 2 and the remaining species present mainly the nucleotides corresponding to subfamily 1 of *C. angulata* and *C. gigas*. We could draw some phylogenetic conclusions from this tree. However, the bootstrap values are low for many of the tree branches, making any phylogenetic conclusions from this tree tentative. Moreover, the clustering might be initially influenced by the existence of two *Hind*III satellite DNA subfamilies and the presence of paralogous sequences in the tree shown in this figure. Taking into account the orthology and paralogy between these sequences, we have analysed the differences between sequences and we have found genetic levels of intraspecific variability lower than levels of interspecific genetic distances, which is reflected in the tree of the Fig. 2 where the sequences are grouped by taxonomic affinity, with the exception of *C. angulata* and *C. gigas*, due probably to their scant genetic differentiation. With the exception of these two species, the analysis shows

than *Hind*III sequences evolve in a concerted way and can be used as a molecular marker for phylogenetic analysis.

In this sense, we used a consensus sequence representative of each species to give information regarding the phylogenetic relationships between the oyster species analysed. Fig. 3 shows the *Hind*III consensus sequences obtained for each of the seven species. For *C. gigas* and *C. angulata*, we defined two consensus sequences, each one for each *Hind*III subfamily. The phylogeny obtained using orthologous sequences pertaining to subfamily 1 (subfamily 2 is present only in *C. gigas* and *C. angulata* themselves and also in *O. stentina*) is illustrated in the Fig. 4. We found that the sequences type 1 of *C. gigas* (Pacific oyster) and *C. angulata* (Portuguese oyster) are closely related to *C. ariakensis*, an Asiatic species. The other clade found in the tree reveals that *O. edulis* (European oyster), *C. gasar* (African oyster) and *C. virginica* (American oyster) constitute a strongly supported monophyletic group of species. These groups are consistent with previous results based on mitochondrial and nuclear markers. The grouping of *Hind*III sequences of the Portuguese oyster (*C. angulata*) together with *C. gigas* and *C. ariakensis* confirm the Asiatic origin of the former species, in accord with Ó Foighil *et al.* (1998). The closest relationship between Asiatic species within *Crassostrea* genera obtained by us, was probed using the mtCOI gene and the 16S rRNA gene (Ó Foighil *et al.*, 1998; Lapègue *et al.*, 2002) and the nuclear 28S rRNA gene sequences (Littlewood, 1994). Still, the paraphyletic situation of *Crassostrea* species when the Asiatic and Pacific oyster are analysed together with species belonging to both *Ostrea* and *Saccostrea* genera has also been described (Littlewood, 1994; Ó Foighil and Taylor, 2000). These two works are based on partial 28S rRNA gene sequence. Therefore, our results based on rapidly evolving satellite DNA confirm the closer relationship of *C. gasar* and *C. virginica* with *O. edulis* than with the Asiatic species, apparently due to the relative isolation of the Atlantic oysters from the Pacific ones.

To confirm the accuracy of the consensus sequences estimated for each species, we developed a study of the transitional stages of concerted evolution for the oyster *HindIII* sequences (Strachan et al., 1985) (Table 2). Most of the comparisons between species gave a percentage of TS 4, 5 and 6 (sites for which the process of intraspecific sequence homogenisation and interspecific divergence is complete or nearly so) correlated with the number of differences found between the consensus sequences established for each species. However, we found that the differences between *C. angulata* and *C. gigas* are minimal and that all the sites fall within stages 2 or 3. This low degree of differentiation led to only four nucleotide differences when comparing *C. gigas* 1 with *C. angulata* 1 (subfamily 1 sequences), and two when comparing *C. gigas* 2 with *C. angulata* 2 (subfamily 2 sequences) (Fig. 3). In any case, these differences between consensus sequences represent sites not clearly homogenized and differentiated between the two taxa. This result supports our conclusion using *HindIII* satellite sequences concerning low genetic differentiation between *C. angulata* and *C. gigas*.

3.3. Conclusions

In summary, we conclude that: (1) The satellite DNA family that we have studied in oysters, the *HindIII* family, is conserved in all the species analysed and preserves several molecular specifics common to centromeric satellite DNAs. (2) The presence of the *HindIII* satellite DNA in the oyster genomes, the homology with satellite DNAs from other bivalves and the homology of all of them with a transposable oyster element appears to indicate the existence of an ancient transposable element, which acted as a generating element of satellite DNAs in bivalve molluscs. (3) The *HindIII* satellite DNA is a useful marker for taxonomic identification to uncover phylogenetic relationships between species of the *Ostrea* and *Crassostrea* genera. (4) The results that we have found using the *HindIII* satellite DNA relegate *C. angulata* and *C. gigas* to the same species.

Acknowledgements

This research was supported by a grant from the Plan Andaluz de Investigación (Group No. CVI0200). We are greatly indebted to Dr. José Ignacio Navas (Centro de Investigación y Formación Acuícola y Pesquera “Agua del Pino”, Huelva, Spain) for his help in sampling and technical assistance.

References

- Amezcuca, O. 1999. Caracterización genética de poblaciones naturales de *Ostrea stentina* en las costas del Golfo de Cádiz. Tesis de Licenciatura. Universidad de Cádiz.
- Batistoni, R., Pesole, G., Marracci, S., Nardi, I. 1995. A tandemly repeated DNA family originated from SINE-related elements in the European plethodontid salamanders (Amphibia, Urodela). *J. Mol. Evol.* 40, 608-615.
- Boudry, P., Heurtebise, S., Collet, B., Cornette, F., Gerard, A. 1998. Differentiation between populations of the Portuguese oyster, *Crassostrea angulata* (Lamarck) and the Pacific oyster, *Crassostrea gigas* (Thunberg), revealed by mtDNA RFLP analysis. *J. Exp. Mar. Biol. Ecol.* 226, 279-291.
- Buroker, N.E., Hershberger, W.K., Chew, K.K. 1979. Population genetics of the family Ostreidae. I. Intraspecific studies of *Crassostrea gigas* and *Saccostrea commercialis*. *Mar. Biol.* 54, 157-169.
- Canapa, A., Barucca, M., Cerioni, P.N., Olmo, E. 2000. A satellite DNA containing CENP-B box-like motifs is present in the Antarctic scallop *Adamussium colbecki*. *Gene.* 247, 175-180.
- Clabby, C., Goswami, U., Flavin, F., Wilkins, N.P., Houghton, J.A., Powell, R. 1996. Cloning, characterization and chromosomal location of a satellite DNA from the Pacific oyster, *Crassostrea gigas*. *Gene* 168, 205-209.
- Comesaña, A.S., Fossum, A., Sanjuan, A. 2001. Distinguishing the commercial flat oyster *Ostrea edulis* from *Ostreola stentina* by PCR-RFLP of the mitochondrial 16S rRNA gene. I Congreso Internacional de Ciencia y Tecnología Marina. Pontevedra (Spain).
- Charlesworth, B., Sniegowski, P. and Stephan, W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371, 215-220.

- Dimitri, P., Junakovic, N. 1999. Revising the selfish DNA hypothesis new evidence on accumulation of transposable elements in heterochromatin. *Trends Genet.* 15, 123-124.
- Gaffney, P.M., Pierce, J.C., Mackinley, A.G., Titchen, D.A., Glenn, W.K. 2003. *Pearl*, a novel family of putative transposable elements in bivalve molluscs. *J. Mol. Evol.* 56, 308-316.
- Harry, H.W. 1985. Synopsis of the supraspecific classification of living oysters (Bivalvia: Gryphaeidae and Ostreidae). *Veliger* 28, 121-158.
- Haure, J., Huvet, A., Palvadeau, H., Nourry, M., Penisson, C., Martin, J.L.Y., Boudry, P. 2003. Feeding and respiratory time activities in the cupped oysters *Crassostrea gigas*, *Crassostrea angulata* and their hybrids. *Aquaculture* 218, 539-551.
- Huvet, A., Balabaud, K., Bierne, N., Boudry, P. 2001. Microsatellite analysis of 6-hour-old embryos reveals no preferential intraspecific fertilization between cupped oysters *Crassostrea gigas* and *Crassostrea angulata*. *Mar. Biotechnol.* 3(5): 448-453.
- Huvet, A., Gérard, A., Ledu, C., Phélipot, P., Heurtebise, S., Boudry, P. 2002. Is fertility of hybrids enough to conclude that the two oysters *Crassostrea gigas* and *Crassostrea angulata* are the same species?. *Aquat. Living Resour.* 15: 45-52.
- Kapitonov, V.V., Holmquist, G.P., Jurka, J. 1998. L1 repeat is a basic unit of heterochromatin satellites in cetaceans. *Mol. Biol. Evol.* 15, 611-612.
- Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111-120.
- Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M. 2001. *MEGA2: Molecular Evolutionary Genetics. Analysis software*, Arizona State University, Tempe, Arizona, USA.
- Lapègue, S., Boutet, I., Leitão, A., Heurtebise, S., Garcia, P., Thiriou-Qiévieux, C., Boudry, P. 2002. Trans-Atlantic distribution of a mangrove oyster species revealed by 16S mtDNA and karyological analyses. *Biol. Bull.* 202: 232-242.
- Leitão, A., Boudry, P., Labat, J.P., Thiriou-Qiévieux, C. 1999. Comparative karyological study of cupped oyster species. *Malacologia* 41 (1) : 175-186.
- Littlewood, D.T.J. 1994. Molecular phylogenetics of cupped oysters based on partial 28S rRNA gene sequences. *Mol. Phyl. Evol.* 3(3): 221-229.
- López, J.L., Marina, A., Álvarez, G., Vázquez, J. 2002. Application of proteomics for fast identification of species-specific peptides from marine species. *Proteomics.* 2: 1658-1665.

- Martínez-Lage, A., Rodríguez, F., González-Tizón, A., Prats, E., Cornudella, L., Méndez J. 2002. Comparative analysis of different satellite DNAs in four *Mytilus* species. *Genome*. 45(5): 922-929.
- Mattiucci, S., Villani, F. 1983. Allozyme study in oysters classified as *Crassostrea gigas* (Thunberg, 1973) and *Crassostrea angulata* (Lamarck, 1819) (Mollusca: Ostreidae). *Parasitologia*. 25: 21-27.
- McDonald, J.H., Verrelli, B.C., Geyer, L.B. 1996. Lack of geographic variation in anonymous nuclear polymorphisms in the American oyster, *Crassostrea virginica*. *Mol. Biol. Evol.* 13, 1114-1118.
- Muchmore, M.E., Moy, G.W., Swanson, W.J., Vacquier, V. 1998. Direct sequencing of genomic DNA for characterization of a satellite DNA in five species of Eastern Pacific abalone. *Mol. Mar. Biol. Biotechnol.* 7(1): 1-6.
- Ó Foighil, D.O., Taylor, D.J. 2000. Evolution of parental care and ovulation behaviour in oysters. *Mol. Phyl. Evol.* 15(2): 301-313.
- Ó Foighil, D.O., Gaffney, P.M., Wilbur, A.E., Hilbish, T.J. 1998. Mitochondrial cytochrome oxidase I gene sequences support an Asian origin for the Portuguese oyster *Crassostrea angulata*. *Mar. Biol.* 131, 497-503.
- Plohl, M., Cornudella, L. 1996. Characterization of a complex satellite DNA in the mollusc *Donax trunculus*: analysis of sequence variations and divergence. *Gene* 169, 157-164.
- Radic, M.Z., Saghbini, M., Elton, T.S., Reeves, R., Hamkalo, B.A. 1992. Hoechst 33258, distamycin A, and high mobility group protein I (HMG-I) compete for binding to mouse satellite DNA. *Chromosoma* 101, 602-608.
- Ranson, G. 1960. Les prodissoconques (coquilles larvaires) des ostréides vivants. *Bulletin l'Institut Océanographique (Monaco)* 1, 1-41.
- Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., Willard, H. F. 2001 Genomic and genetic definition of a functional human centromere. *Science* 294, 109-115.
- Stepien, C.A., Kocher, T.D. 1997. Molecules and morphology in studies of fish evolution. In: Kocher, T.D. and Stepien, S.D. (Eds.) *Molecular Systematics of Fishes*. Academic Press, San Diego, pp. 1-11.
- Strachan, T., Webb, D., Dover, G. 1985. Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J.* 7, 1701-1708.
- Ugarkovic, D., Plohl, M. 2002. Variation in satellite DNA profiles-causes and effects. *EMBO J.* 21, 5955-5959.
- Wang, Y.P., Xu, Z., Guo, X.M. 2001. A centromeric satellite in the Pacific oyster (*Crassostrea gigas* Thunberg) identified by fluorescence in situ hybridization. *Mar. Biotechnol.* 3, 486-492.

Winnepenninckx, B., Backeljau, T., De Wachter, R. 1993. Extraction of high molecular weigh DNA from molluscs. Trends Genet. 9, 407.

Table 1. Information concerning the oyster species analysed, including the GenBank accession numbers of the *HindIII* sequences.

Species	Common name	Clones	Accession number	Origin	Population
<i>O. edulis</i>	European flat oyster	Oed8, Oed9, Oed30, Oed37, Oed38, Oed41, Oed86, Oed105	AJ601406 to AJ601413	Huelva, Spain	Hatchery
<i>O. stentina</i>	Provence oyster	Ost4, Ost5, Ost6, Ost22, Ost37	AJ604556 to AJ604560	Huelva, Spain	Wild
<i>C. angulata</i>	Portuguese oyster	CaCad22, CaCad87, CaCad93, CaCad23, CaCad21, CaCad24	AJ601422 to AJ601427	Cádiz, Spain	Wild
		CaKao19, CaKao23, CaKao42	AJ601428 to AJ601430	Kaoshiung, Taiwan	Wild
<i>C. gigas</i>	Pacific oyster	CgArc46, CgArc47, CgArc83, CgArc104, CgArc37, CgArc99	AJ601431 to AJ601436	Arcachon, France	Wild
		CgHir42, CgHir81, CgHir97	AJ601437, AJ604547, AJ604548	Hiroshima, Japan	Wild
<i>C. gasar</i>	Mangrove oyster	Cgas39, Cgas52, Cgas65, Cgas89	AJ601418 to AJ601421	Casamance, Senegal	Wild
<i>C. ariakensis</i>	Suminoe oyster	Car1, Car2, Car3, Car4, Car65, Car70, Car81	AJ604549 to AJ604555	New Jersey, USA	Hatchery
<i>C. sikamea</i>	Kumamoto oyster	Cs20, Cs35, Cs43, Cs98	-	California, USA	Hatchery
<i>C. virginica</i>	American oyster	Cv26, Cv44, Cv52, Cv55	AJ601414 to AJ601417	New Brunswick, Canada	Wild

Table 2. Percentage of transition stages (TS) (Strachan et al., 1985) calculated from the pairwise comparisons of orthologous *Hind*III monomeric sequences. Note that the closest species show neither ST5 nor ST6 in the analysis (marked with *). The more differentiated the species, the higher the TS5 and TS6 values reflected in the comparison between their sequences. Species are denoted by their generic and specific initials (Oed, *O. edulis*; Ost, *O. stentina*; Ca, *C. angulata*; Cg, *C. gigas*; Cga, *C. gasar*; Car, *C. ariakensis*; Cv, *C. virginica*), and for *C. angulata* and *C. gigas* also by a number indicating the subfamily type. N= number of nucleotides compared.

	TS1	TS2	TS3	TS4	TS5	TS6	N
Oed-Ca1	50.29	20.36	6.58	2.39	7.18	3.59	167
Oed-Cg1	38.09	27.97	0.59	2.97	2.97	1.19	168
Oed-Cga	57.14	19.64	1.19	2.98	4.17	3.57	168
Oed-Car	37.50	20.83	4.17	4.76	1.79	3.57	168
Oed-Cv	58.93	19.05	1.19	3.57	4.17	2.98	168
Ca1-Cg1*	45.50	32.33	2.37	1.19	0.00	0.00	168
Ca1-Cga	57.73	5.95	7.14	2.97	11.90	2.38	168
Ca1-Car	45.83	23.21	11.31	0.59	0.59	0.00	168
Ca1-Cv	61.90	4.16	8.33	4.16	11.90	1.78	168
Cg1-Cga	47.33	23.07	0.59	4.73	5.32	2.36	169
Cg1-Car*	34.91	30.77	2.36	0.00	0.00	0.00	169
Cg1-Cv	46.15	24.26	0.59	4.14	7.10	1.18	169
Cga-Car	46.43	19.64	6.55	6.55	5.95	1.19	168
Cga-Cv	86.14	10.84	1.20	0.60	0.60	0.00	166
Car-Cv	48.21	22.62	7.14	6.55	5.95	1.79	168
Ost-Ca2	49.70	29.58	4.73	3.55	0.00	1.18	169
Ost-Cg2	58.92	22.02	4.76	3.57	1.78	1.19	168
Ca2-Cg2*	44.91	33.53	4.18	0.00	0.00	0.00	168

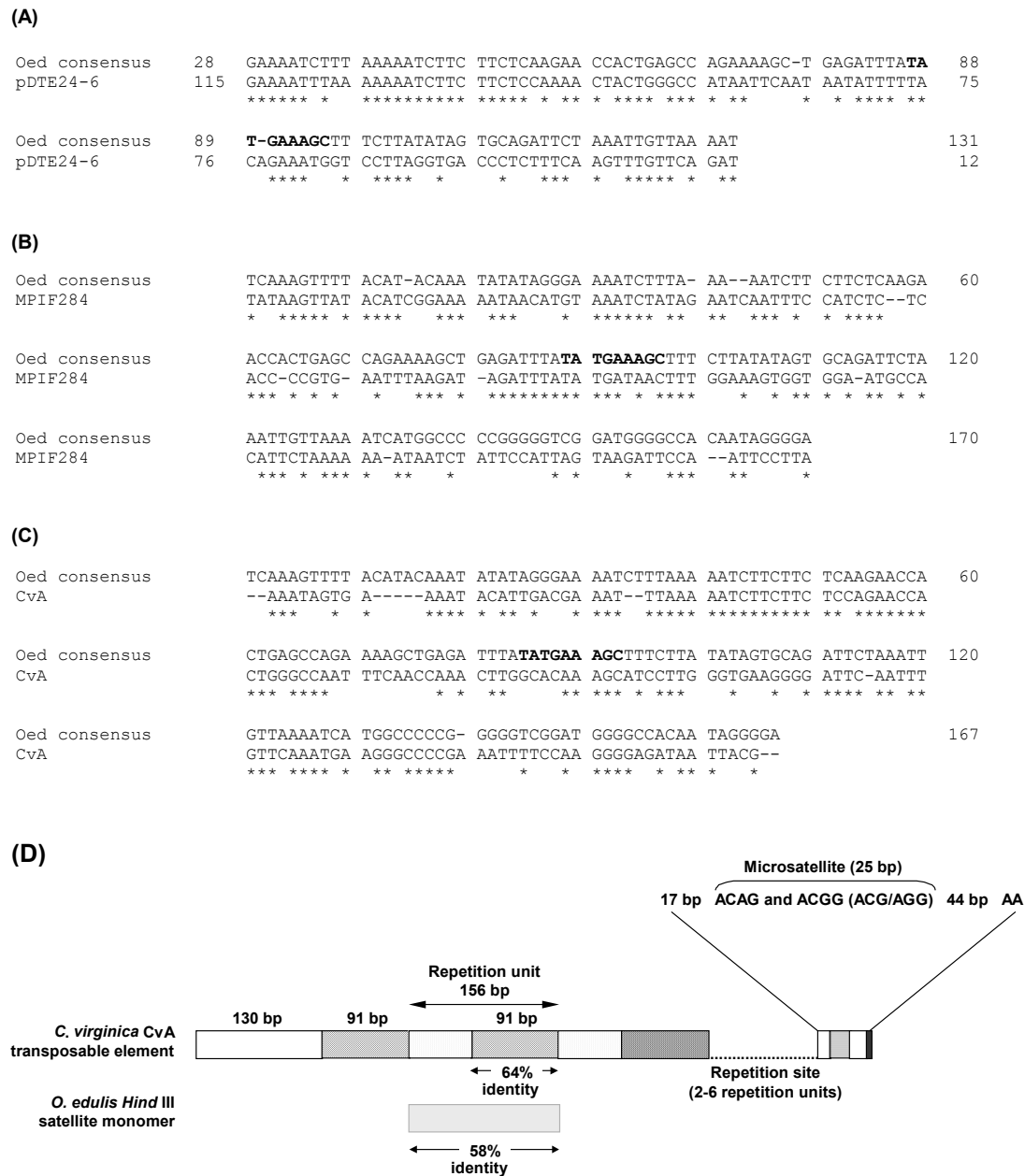


Fig. 1. Alignment of *O. edulis* HindIII satellite DNA with homologous sequences. The figure shows the alignment of the 103 homologous nucleotide positions of *O. edulis* consensus sequence and *D. trunculus* (pDTE24-6) monomeric unit of satellite DNA (A), the alignment of complete HindIII monomeric unit with a MITE of *Z. mays* (MPIF284) (B) and with the transposable element CvA of *C. virginica* (C). In bold, the nine-bp motif similar to the mammalian CENP-B box. Last diagram (D) shows the location of homologous zone between *O. edulis* HindIII satellite and *C. virginica* CvA transposable element.

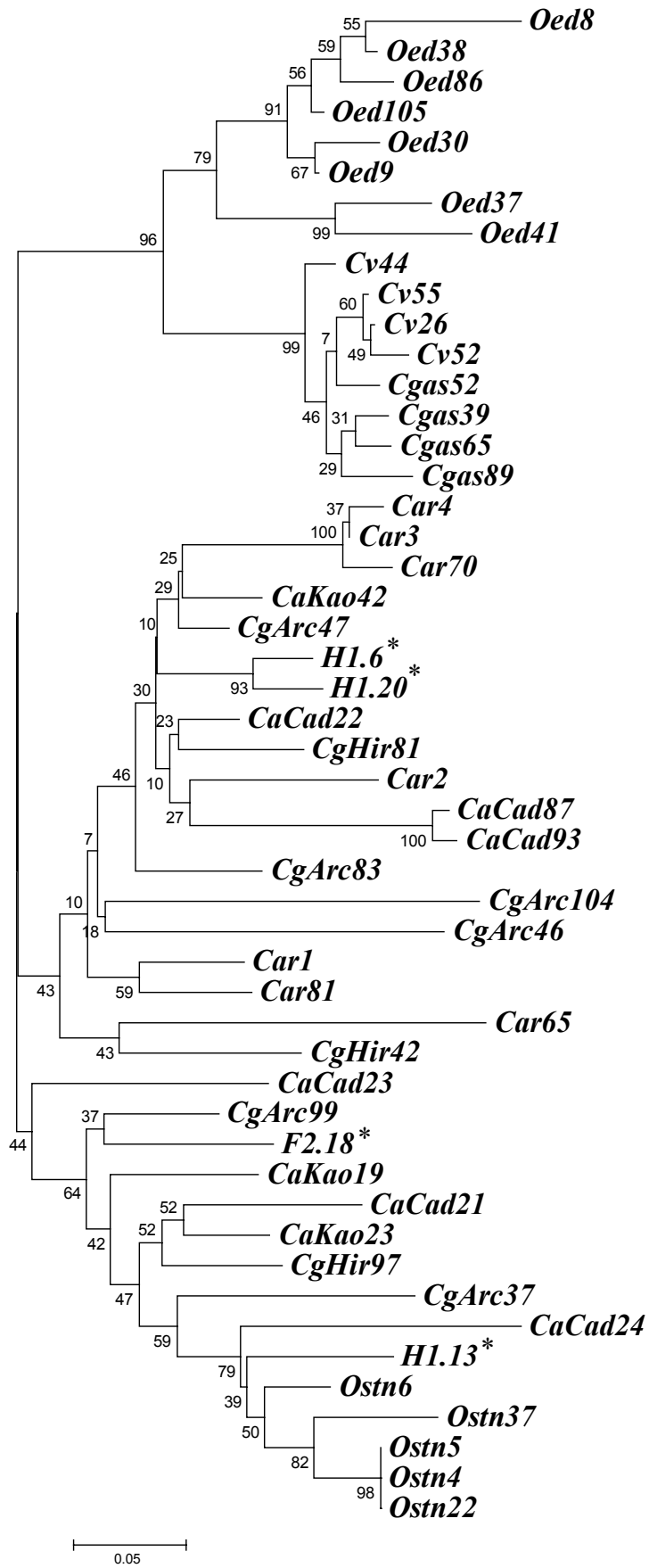


Fig. 2. Neighbour-joining tree of the 46 *Hind*III satellite DNA monomers using the Kimura-2 parameters method. Bootstrap values supporting each node are presented to the left of the tree branches. The name of the specimens corresponds to first letter of genus and species. In the case of *C. angulata* and *C. gigas*, the name from the population is also cited. The final number identifies the clone. Sequences obtained by Clabby et al., (1996) (marked with *) have been included. Our analysis reveals that the high divergence found by these authors was because these clones belong to different subfamilies: H1.6 and H1.20 belong to subfamily 1 and H1.13 and F2.18 to subfamily 2.

	*		*						
<i>C. angulata</i> 1	TTGAATTTT	ACATAGGAAT	ATATAGAGTA	AATCTTTAAA	AATCTT-CTT	CTCAGAAACC			60
<i>C. gigas</i> 1	.C...A....-			
<i>C. ariakensis</i>	.C.....-			
<i>O. edulis</i>	.CA..G....CA..G..A.-AG....			
<i>C. virginica</i>	.CA..G....	.T..T.A..	G.....GAA.-AG....			
<i>C. gasar</i>	.CA..G...AT.A..	G.....GAA.-AG....			
<i>C. angulata</i> 2	.C-..G....A.-A...T			
<i>C. gigas</i> 2	.C-..G....A.-A...T			
<i>O. stentina</i>	.C-..G....A.A...TA...T			
	* * *	*	*	* * *		* * *			
<i>C. angulata</i> 1	AATCAGCCAG	GAAAGCTGAA	ACTTGTGTGG	AAGCATCCTC	AGGTAGTGTA	GATTCAAAGT			120
<i>C. gigas</i> 1	.T.....			
<i>C. ariakensis</i>	..T.....A.....			
<i>O. edulis</i>	.C.G.....	A.....G	.T..A.A..A	...T.T.T	.TA.....C.	...T..A.			
<i>C. virginica</i>	.C.G.A....	A.G..A....	.T..A.....	...GT..T.T	..AA.....	...T..A.			
<i>C. gasar</i>	.C.GGA....	A.G..A....	.T..A.....	...GT..T.T	..A.....	...T..A.			
<i>C. angulata</i> 2	.T.TG....	A.....C..	.T..AAAA..GT.			
<i>C. gigas</i> 2	.T.AG....	A.....C..	.T..AAAA..A.....GT.			
<i>O. stentina</i>	.TGAC....C..	.T..AAAA..A...A.	T...TTGT.			
	* *	* *							
<i>C. angulata</i> 1	TGTGAAAATC	ATGACCCCCA	GGGGTA-GGG	TGGGGCCACA	ATGGGGGG-				169
<i>C. gigas</i> 1G--				
<i>C. ariakensis</i>G-	..A.....	..T.....-				
<i>O. edulis</i>	..T.....	..G.....GC-..AA-...A				
<i>C. virginica</i>	..TC.....G	...-..A..-A...A				
<i>C. gasar</i>	..TC.....G	...-..A..-A...A				
<i>C. angulata</i> 2	..TC.....	..GT...G-A				
<i>C. gigas</i> 2	..TC.....	..GT...G-A				
<i>O. stentina</i>	..TC.....	..GT...G-	..A.....	G.A...A				

Fig. 3. Aligned *Hind*III satellite DNA consensus sequences of the seven oyster species studied. The complete *Hind*III sequence of *C. angulata* subfamily 1 is shown. Dots denote nucleotides that are the same as the first sequence and dashes denote gaps. Below the sequences (*) marks the 16 positions for subfamily 1 and 2 defined in *C. angulata* and *C. gigas* and present in the remaining species.

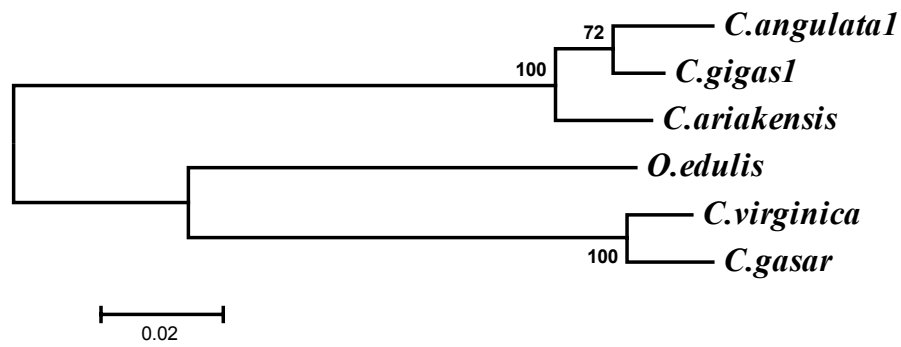


Figure 4. Phylogenetic tree of the oyster species analysed as deduced from their *Hind*III subfamily 1 monomeric sequences using the Neighbour-joining phylogenetic-inference method and the Kimura-2 parameters model. Bootstrap values supporting each node are presented to the left of the tree branches