5S rDNA and U2 snRNA are linked in the genome of Crassostrea angulata and Cra...
I Cross; L Rebordinos
Genome; Dec 2005; 48, 6; Health & Medical Complete
pg. 1116

NOTE / NOTE

5S rDNA and U2 snRNA are linked in the genome of *Crassostrea angulata* and *Crassostrea gigas* oysters: does the $(CT)_n \cdot (GA)_n$ microsatellite stabilize this novel linkage of large tandem arrays?

I. Cross and L. Rebordinos

Abstract: The 5S rRNA genes from 2 species of the Ostreidae family, *Crassostrea angulata* and *Crassostrea gigas*, were molecularly characterized. The genes were amplified, cloned, and sequenced. The results revealed a 5S rDNA tandem array with a nucleotide sequence in an inverted position within the nontranscribed spacer region that corresponded to the U2 small nuclear RNA (snRNA) gene. The sequence analysis indicated that both genes could be functionally active. The presence of the microsatellite $(CT)_n \cdot (GA)_n$ at the 3′ end of both genes and the possible involvement of concerted evolution are discussed.

Key words: Crassostrea angulata, Crassostrea gigas, 5S rDNA, U2 snRNA, microsatellite, concerted evolution.

Résumé : Les gènes codant pour les ARNr 5S chez 2 espèces de la famille des ostréidés, *Crassostrea angulata* et *Crassostrea gigas*, ont été caractérisés au niveau moléculaire. Les gènes ont été amplifiés, clonés, et séquencés. Les résultats ont révélé l'existence d'une suite en tandem de gènes d'ADNr 5S et la présence d'une séquence nucléotidique correspondant au gène du petit ARN nucléaire (snRNA) U2 en position inversée au sein de l'espaceur non-transcrit (NTS). Une analyse de séquence a indiqué que les deux gènes pourraient s'avérer fonctionnels. La présence du microsatellite $(CT)_n \cdot (GA)_n$ à l'extrémité 3′ des 2 gènes et la possible implication d'un mécanisme d'évolution concertée sont discutées.

Mots clés : Crassostrea angulata, C. gigas, ADNr 5S, snRNA U2, microsatellite, évolution concertée.

[Traduit par la Rédaction]

Introduction

Crassostrea angulata and *Crassostrea gigas* are 2 species of the family *Ostreidae* with similar morphology. They can only be distinguished genetically by RFLP analysis of the cytochrome oxidase subunit I mitochondrial gene (Boudry et al. 1998). The 5S rDNA in bivalve mollusks have only been studied in the mussels *Mytilus edulis* and *Mytilus galloprovincialis* (Insua et al. 2001), and in *Cerastoderma edule* (Insua et al. 1999).

Recently in our laboratory, we used fluorescence in situ hybridization to locate the 5S rDNA gene on 2 medium-size chromosomal pairs in *C. angulata* ($2n = 20$) (Cross et al. 2005) that are different from those carrying the 18S–5.8S–28S rDNA genes (Cross et al. 2003).

In some species, 2 or more variants of the 5S rRNA gene have been described (Cloix et al. 2000; Insua et al. 2001). Each 5S rRNA gene contains a coding region of 120 nucleotides and an intergenic, nontranscribed spacer (NTS) of variable length and sequence (Suzuki et al. 1996). The genes in mammals range from 1.6 to 2.2 kb (Suzuki et al. 1994), while nonmammalian vertebrates, plants, and invertebrates contain much shorter DNA sequences (175–880 bp). In Atlantic salmon, the two 5S rDNA variants differ in the insertion of a 270 bp DNA segment in the NTS (Pendás et al. 1994). In *Xenopus laevis*, two 5S rDNA variants are associated with either the oocytes or somatic cells (Miller et al. 1978). Both types are active during oogenesis, but the oocyte

Received 7 March 2005. Accepted 22 June 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 9 December 2005.

Corresponding Editor: W. Traut.

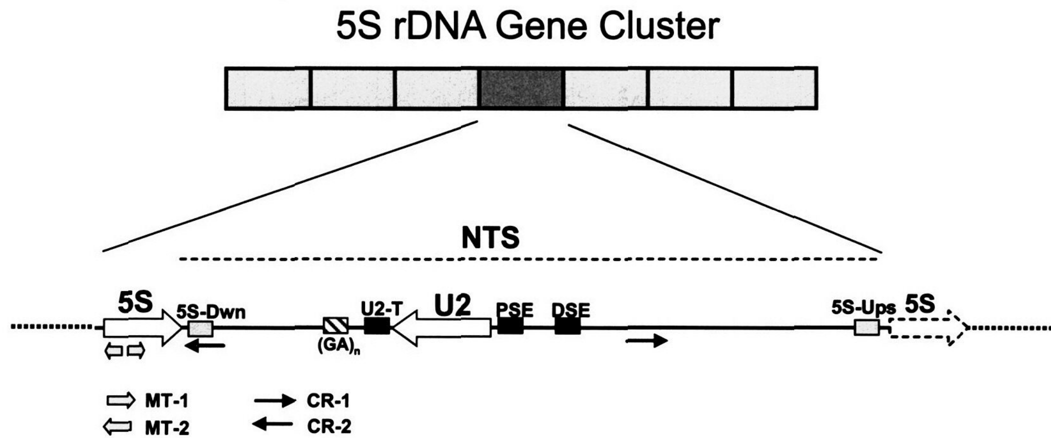
I. Cross and L. Rebordinos.[1] Laboratorio de Genética, Facultad de Ciencias del Mar y Ambientales, Universidad de Cádiz, Polígono del Río S. Pedro s/n. 11510, Puerto Real, Cádiz, Spain.

[1]Corresponding author (e-mail: laureana.rebordinos@uca.es).

 doi: 10.1139/G05-075 © 2005 NRC Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fig. 1. Schematic representation of the 5S rDNA gene cluster in *Crassostrea angulata* and *Crassostrea gigas*. 5S rDNA of 10 individuals of *C. angulata* and 10 individuals of *C. gigas* was amplified by PCR using primers MT1 and MT2. The product lengths from both species were virtually identical, at approximately 1100 bp. Six clones from *C. angulata* (2 clones of 3 individuals) and 5 clones from *C. gigas* (1 or 2 clones of 3 individuals) were sequenced. Sequences are available in the EMBL database under accession Nos. AY765358–63 (*C. angulata*) and AY765364–68 (*C. gigas*). Position and orientation of upstream regulatory sequences (5S-Ups) and downstream regulatory sequences (5S-Dwn) are shown. PSE regulatory regions are the proximal element, DSE regulatory regions are the distal element, and U2-T represents the terminating sequences in the U2 snRNA gene. $(GA)_n$ is the microsatellite sequence and MT-1, MT-2, CR-1, and CR-2, indicate the positions and primers used for the complete molecular characterization of the tandem array.



type is inactivated progressively during embryonic development, giving rise to a selective expression of the somatic type during later larval stages.

In the present study, the full versions of the 5S gene tandemly repeated cluster from *Crassostrea angulata* and *Crassostrea gigas* were cloned using PCR products. Clones of both species were sequenced and analyzed. The results were compared with other species, and the possible involvement of concerted evolution is discussed.

Results and discussion

This is the first time that the molecular characterization of both a U snRNA gene and a complete 5S rDNA gene has been carried out in a species of oyster. The results showed that the U2 snRNA and 5S rDNA genes were closely linked in a gene cluster in the genomes of *C. angulata* and *C. gigas* (Fig. 1). Using the specially designed primers in PCR, the sequence analysis of the clones indicated that the cluster was tandemly repeated. The 5S coding region of *Crassostrea* was 120 bp long (Fig. 2), with an average NTS of 980 bp (970–997 bp). The length of analyzed clones differed mainly in the varying number of $(CT)_n$·$(GA)_n$ microsatellite repeat unit.

Only 1 type of 5S rDNA tandem repeat was found in *Crassostrea*; however, in other species, several repeated patterns that vary in both spacer length and sequence have been observed (Insua et al. 2001; Pendás et al. 1994). With respect to 5S rDNA transcription, the sequence analysis indicated that the 5S rDNA gene is possibly a possible functional gene that contains a TATA box upstream of the transcription start site, as well as internal control regions within the 5S coding sequence. Certain DNA sequences at the 3′ end associated with the transcription termination were also seen. Furthermore, the abundance of thymine (>50%) at 70 bp downstream of the coding region could indicate functional terminators (Bogenhagen and Brown 1981) (Table 1).

Additionally, the sequence analysis of the U2 snRNA gene, including both the coding and noncoding regions, revealed that it is a putative active and functional gene. It contains all of the regulatory elements upstream and downstream of the coding region, as well as the protein-binding sites within the coding sequences that are necessary for mRNA splicing (Table 1 and Fig. 3). Among the analyzed clones, polymorphisms were found in certain repeats. Repeats were seen at 8–13 copies (Fig. 4). In all of the clones except for Cang1A, 1B, and 3B, T was substituted by C after the second repeat. Additionally, clone Cang3B showed an inverse substitution (C substituted by T) after the sixth repeat.

Our recent discovery of the 5S rRNA gene on 2 chromosomal pairs (Cross et al. 2005), together with the linkage of the 5S and U2 snRNA genes, leads to the possibility of concerted evolution. The presence of the microsatellite $(CT)_n$·$(GA)_n$ at the 3′ end of both gene clones could play a very important role in concerted evolution mechanisms. Several genes in humans and other organisms, including U1 snRNA, U2 snRNA, and 5S and 45S rRNA, are organized in tandemly repeated arrays and contain microsatellite sequences (CT in snRNA and GT in rRNA) downstream of the transcription unit (van Arsdell and Weiner 1984; Htun et al. 1985; Little and Braaten 1989; Sorensen and Frederiksen 1991). These specific repeating units found in all analyzed gene families suggest the function of tandem array maintenance (Liao and Weiner 1995) and concerted evolution of loci. A human study has provided evidence that a CT microsatellite stabilizes a long tandem array integration of transcriptionally active U2 snRNA genes (Bailey et al. 1998). It has been shown that CT repeating units can assume a triplex structure, removing the strain rich in CT and allowing recombination within the CT microsatellite (Mirkin et al. 1987). Hence, the

© 2005 NRC Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fig. 2. Nucleotide sequences of the 5S rDNA coding regions of *Crassostrea angulata* and *Crassostrea gigas* aligned to *Mytilus edulis*, *Mytilus galloprovincialis* (Insua et al. 2001) and *Cerastoderma edule* (Insua et al. 1999). Dots represent identical residues. Base substitutions are indicated by the respective bases. Internal control regions are underlined.

```
Cang/gig    GTCTACGGCCATATCACGTTGAAAGCACCGGTTCTCGTCCGATCACCGAAGTTAAGCAAC  60
Mytilus     .......A...............A....................................  60
C.edule     .......................A....................................  60

C.ang/gig   GTAGAGCTTGGTTAGTACTTGGATGGGTGACCGCCTGGGAATACCAGGTGTCGTAGACTT  120
Mytilus     ..C....CC...................................G.....T......A.  120
C.edule     ..C....CC...................................G....CT........  120
```

Table 1. Regulatory elements of the 5S rDNA gene and U2 snRNA gene in *Crassostrea angulata* and *Crassostrea gigas*.

Regulatory element	Description
5S rDNA	
TATA Box	–30 TATATTACTTGTTTCTAACTC-CTAGTTGTT –1
Terminator	Residues poli-T (+121 to +183)
U2 snRNA	
PSE	–50 TTATTTGAAAATGATTTTATT –30
DSE	–150 TTGCAAAT –142
Terminator	+194 TTTATAACAAAATAATAG +211

Fig. 3. RNA secondary structure of the U2 snRNA gene of *Crassostrea angulata* and *Crassostrea gigas* based on a currently accepted model (Yu et al. 1999). Interaction sites are as follows: U2–U6 Helix I, nucleotides 20–22 and 25–28; U2–U6 Helix II, nucleotides 1–8; U2–U6 Helix III, nucleotides 40–50; ChnRNP binding site, nucleotides 15–28; ChnRNP + U2–U6 Helix II overlap, nucleotides 20–22 and 25–28; branchpoint interaction site (underlined), nucleotides 34–39; Sm binding site (underlined), nucleotides 101–107; and U2A′–U2B′ binding site, nucleotides 154–183.

CT microsatellite could be considered to be a spacer between U2 repeats, rather than a microsatellite fitted within a repeating unit. Concerted evolution could explain array maintenance and homogenization, as well as the CT microsatellite polymorphisms observed in *Crassostrea* in this study.

The linkage between the 5S and U2 snRNA genes in the *C. angulata* and *C. gigas* genomes could be explained by 5S rDNA retrotransposition mechanisms via an RNA intermediate, as evidenced in rats and mice (Drouin 2000). 5S rRNA genes linked to other tandemly repeated gene families have been observed in diverse genomes (Drouin and Moniz de Sá 1995; Pelliccia et al. 2001; Eirín-López et al. 2004). Although the mechanisms of transposition have not been completely resolved, the mechanisms by which 5S is subsequently spread to all members of a tandemly repeated multigenic family after its transposition to a single repeating unit are well known. Further studies are needed to investigate the possible transposition of the 5S rDNA gene to other

© 2005 NRC Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fig. 4. Alignment of CT microsatellites from 11 clones (GenBank accession Nos. AY765358–68). Polymorphisms are shown. Repeating units are highlighted and substitutions are in bold and underlined.

```
5S_Cang1A   T-CCCGCCTCTCTCTCTCTCTCTCTCTCTCTTCAATCTT
5S_Cang1B   T-CCCGCCTCTCTCTCTCTCTCTCTCTCTCTTCAATCTT
5S_Cang2A   T-CCCGCCTCTCCCTCTCCCTCTCTCTCT----TTCAATCTT
5S_Cang2B   TTCCCGCCTCTCCCTCTCCCTCTCTCTCTCT--TTCAATCTT
5S_Cang3A   -TCCCGCCTCTCCCTCTCTCTCTCTCT------TTCAATCTT
5S_Cang3B   TTCCCGGCTCTCTCTCTCTTTCTCTCTCTCTCTTTCAATCTT
5S_Cgig1A   T-CCCGCCTCTCCCTCTCTCTCTCTCTCT----TTCAATCTT
5S_Cgig1B   T-CCCGCCTCTCCCTCTCTCTCTCTCT------TTCAATCTT
5S_Cgig2A   T-CCCGCCTCTCCCTCTCTCTCT----------TTCAATCTT
5S_Cgig3A   T-CCCGCCTCTCCCTCTCTCTCTCTCTCT----TTCAATCTT
5S_Cgig3B   T-CC-GCCTCTCCCTCTCTCTCTCTCT------TTCAATCTT
```

loci, in addition to the multigenic families described up to date.

Acknowledgements

This work has been supported by projects INTERREG IIIA OPAM from the EEC and CVI219 from the Junta de Andalucia (Spain). The authors would like to thank Dr. A. Martínez-Lage (University of La Coruña), for the primer sequences, Rosa Vázquez (Laboratory of Marine Culture, University of Cádiz) for the microalgae cultures used in the feeding of *C. angulata*, and S.L. Amalthea for the samples of *C. gigas*.

References

Bailey, A.D., Pavelitz, T., and Weiner, A.M. 1998. The microsatellite sequence $(CT)_n$•$(GA)_n$ promotes stable chromosomal integration of large tandem arrays of functional human U2 small nuclear RNA genes. Mol. Cell. Biol. **18**: 2262–2271.

Bogenhagen, D.F., and Brown, D.D. 1981. Nucleotide sequences in *Xenopus* 5S DNA required for transcription termination. Cell, **14**: 261–270.

Boudry, P., Heurtebise, S., Collet, B., Conette, F., and Gérard, A. 1998. Differentiation between populations of the Portuguese oyster *Crassostrea angulata* (Lamark) and the Pacific oyster, *Crassostrea gigas* (Thunberg), revealed by mtDNA RFLP analysis. J. Exp. Mar. Biol. Ecol. **226**: 279–291.

Cloix, C., Tutois, S., and Mathieu, O. 2000. Analysis of 5S rDNA arrays in *Arabidopsis thaliana*: physical mapping and chromosome-specific polymorphisms. Genome Res. **10**: 679–690.

Cross, I., Vega, L., and Rebordinos, L. 2003. Nucleolar organizing regions in *Crassostrea angulata*: chomosomal location and polymorphism. Genetica, **119**: 65–74.

Cross, I., Díaz, E., Sánchez, I., and Rebordinos, L. 2005. Molecular and cytogenetic characterization of *Crassostrea angulata* chromosomes. Aquaculture, **247**: 135–144.

Drouin, G. 2000. Expressed retrotransposed 5S rRNA genes in the mouse and rat genomes. Genome, **43**: 213–215.

Drouin, G., and Moniz de Sá, M. 1995. The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. Mol. Biol. Evol. **12**: 481–493.

Eirín-López, M., Fernanda Ruiz, M., González-Tizón, A., Martinez, A., Sánchez, L., and Méndez, J. 2004. Molecular evolutionary characterization of the Mussel *Mytilus* histone multigene family: first record of a tandemly repeated unit of five histone genes containing an H1 subtype with "orphon" features. J. Mol. Evol. **58**: 131–144.

Htun, H., Lund, E., Westin, G., Petterson, U., and Dahlberg, J.E. 1985. Nuclease S1-sensitive sites in multigene families, human U2 small nuclear RNA genes. EMBO J. **4**: 1839–1845.

Insua, A., Freire, R., and Méndez, J. 1999. The 5S rDNA of the bivalve *Cerastoderma edule*: nucleotide sequence of the repeat unit and chromosomal location relative to 18S-28S rDNA. Genet. Sel. Evol. **31**: 509–518.

Insua, A., Freire, R., Ríos, J., and Méndez, J. 2001. The 5S rDNA of mussels *Mytillus galloprovincialis* and *M. edulis*: sequence variation and chromosomal location. Chromosome Res. **9**: 495–505.

Liao, D., and Weiner, A.M. 1995. Concerted evolution of the tandemly repeated genes encoding primate small nuclear RNA (the *RNU2* Locus) does not prevent rapid diversification of the $(CT)_n$•$(GA)_n$ microsatellite embedded within the U2 repeat unit. Genomics, **30**: 583–593.

Little, R.D., and Braaten, D.C. 1989. Genomic organization of human 5S rDNA and sequence of one tandem repeat. Genomics, **4**: 376–383.

Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V., and Brown, D.D. 1978. The nucleotide sequence of oocyte 5S DNA in *Xenopus laevis*. II. The GC-rich region. Cell, **13**: 717–725.

Mirkin, S.M., Lyamichev, V.I., Drushlyak, K.N., Dobrynin, V.N., Filippov, S.A., and Frank-Kamenetskii, M.D. 1987. DNA H form requires a homopurine–homopyrimidine mirror. Nature (London), **330**: 495–497.

Pelliccia, F., Barzotti, R., Bucciarelli, E., and Rocchi, A. 2001. 5S ribosomal and U1 small nuclear RNA genes: a new linkage type in the genome of a crustacean that has 3 different tandemly repeated units containing 5S ribosomal sequences. Genome, **44**: 331–335.

Pendás, A.M., Morán, P., Freije, J.P., and García-Vásquez, E. 1994. Chromosomal location and nucleotide sequence of two tandem repeats of the Atlantic salmon 5S rDNA. Cytogenet. Cell Genet. **67**: 31–36.

Sorensen, P.D., and Frederiksen, S. 1991. Characterization of human 5S rRNA genes. Nucleic Acids Res. **19**: 4147–4151.

Suzuki, H., Moriwaki, K., and Sakurai, S. 1994. Sequences and evolutionary analysis of mouse 5S rDNAs. Mol. Biol. Evol. **11**: 704–710.

Suzuki, H., Sakurai, S., and Matsuda, Y. 1996. Rat rDNA spacer sequences and chromosomal assignment of the genes to the extreme terminal region of chromosome 19. Cytogenet. Cell Genet. **72**: 1–4.

van Arsdell, S.W., and Weiner, A.M. 1984. Human genes for U2 small nuclear RNA are tandemly repeated. Mol. Cell. Biol. **4**: 492–499.

Yu, Y.-T., Schart, E.C., Smith, C.M., and Seitz, J.A. 1999. The growing world of small nuclear ribonucleoproteins. *In* The RNA world: the nature of modern RNA suggests a prebiotic RNA. *Edited by* R.F. Gesteland, T. Cech, and J.F. Atkins. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. pp. 487–524.

© 2005 NRC Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.