Repetitive DNA, molecular cytogenetics and genome organization in the King scallop (Pecten maximus)

Maria A. Biscotti, Adriana Canapa, Ettore Olmo, Marco Barucca, Chee H. Teo, Trude Schwarzacher, Sven Dennerlein, Rica Richter, J.S. (Pat) Heslop-Harrison

Università Politecnica delle Marche, Italy
Department of Biology, University of Leicester, UK

Received 20 December 2006; received in revised form 15 June 2007; accepted 25 June 2007
Available online 19 July 2007

Abstract

We studied the structure, organization and relationship of repetitive DNA sequences in the genome of the scallop, Pecten maximus, a bivalve that is important both commercially and in marine ecology. Recombinant DNA libraries were constructed after partial digestion of genomic DNA from scallop with PstI and ApaI restriction enzymes. Clones containing repetitive DNA were selected by hybridisation to labelled DNA from scallop, oyster and mussel; colonies showing strong hybridisation only to scallop were selected for analysis and sequencing. Six non-homologous tandemly repeated sequences were identified in the sequences, and Southern hybridisation with all repeat families to genomic DNA digests showed characteristic ladders of hybridised bands. Three families had monomer lengths around 40 bp while three had repeats characteristic of the length wrapping around one (170 bp), or two (326 bp) nucleosomes. In situ hybridisation to interphase nuclei showed each family had characteristic numbers of clusters indicating contrasting arrangements. Two of the repeats had unusual repetitions of bases within their sequence, which may relate to the nature of microsatellites reported in bivalves. The study of these rapidly evolving sequences is valuable to understand an important source of genomic diversity, has the potential to provide useful markers for population studies and gives a route to identify mechanisms of DNA sequence evolution.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Satellite DNA; Cytogenetics; Tandem repeats; Fiber-FISH; Bivalvia; Mollusca

1. Introduction

Bivalves are of interest both because of their use as a high-value, commercial species of importance in marine ecology which can be farmed, or at least seeded and managed and because of their evolutionary position within the Mollusca and Lophotrochozoa, a sister group to the Ecdysozoa (including insects) and Deuterostomia (including higher animals). In the bivalves of the order Ostreoida, the scallop family, Pectinidae, includes some 350 living species with several major commercial species (such as the King or Great scallop, Pecten maximus, and Bay scallop, Argopecten irradians). Sister families include mussels and oysters. Although the group as a whole and most individual species are well circumscribed, their attribution to subfamilies and genera is sometimes equivocal, and there is limited information about phylogeny and relationships between the species, not least because most work has been based on adult morphology (Barucca et al., 2004). As has been pointed out, e.g. by Wang and Guo (2004), multiple major changes in chromosome number and structure have occurred during the evolution of Pectinidae, and Bouilly et al. (2004, 2005) have shown that bivalves can tolerate chromosome loss and aneuploidy. Clearly the evolutionarily extensive reorganization of the genomes has allowed the reproductive isolation and differentiation of the large number of species.

Abbreviations: bp, base pairs; DAPI, 4′,6-diamidino-2-phenylindole; EDTA, ethylenediamine tetraacetic acid; EMBL, European Molecular Biology Laboratory; Mbp, mega (million) base pairs; PCR, polymerase chain reaction; rDNA, ribosomal DNA; SDS, sodium dodecyl sulphate; SSC, salt sodium citrate, 150 mM sodium chloride and 15 mM trisodium citrate.

* Corresponding author. Tel.: +39 712204619; fax +39 71 2204609.
E-mail address: a.canapa@univpm.it (A. Canapa).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.06.027
The nucleus of the King scallop has \(2n=38\) chromosomes and the DNA content of the nucleus (1C) is 1.42 pg (Gregory, 2007), giving a genome size of 1400 Mbp, near the middle of the range of measured genome sizes in bivalves (890 Mbp in Crassostrea gigas, the edible oyster, through 1560 Mbp in Mytilus edulis, mussel, to more than 3000 Mbp in species in the orders Unionoida and Nuculoida). As with other species with genomes much larger than 100–200 Mbp, there is the expectation that repetitive DNA sequences represent a large proportion of the genome and are important for understanding species relationships, genome evolution and nuclear organization. While the chromosomes of many species of scallop have been studied and recent work has shown the chromosomal locations of the 45S and 5S rDNA loci in scallop (Wang and Guo, 2004; Insua et al., 2006), there has been only limited work at DNA level (Saavedra and Bachère, 2006), despite the commercial importance. Increasingly substantial databases of expressed sequences (ESTs, Johnston, 2006), and genomic or EST-derived microsatellites (Zhan et al., 2006a) from \(P\).

While the chromosomes of many species of scallop have been studied and recent work has shown the chromosomal locations of the 45S and 5S rDNA loci in scallop (Wang and Guo, 2004; Insua et al., 2006), there has been only limited work at DNA level (Saavedra and Bachère, 2006), despite the commercial importance. Increasingly substantial databases of expressed sequences (ESTs, Johnston, 2006), and genomic or EST-derived microsatellites (Zhan et al., 2006a) from oysters and mussels. We then examined the similarities in developmental programmes with other taxonomic groups. Examination of histone H3 gene sites in four scallop species (two Chlamys, Patinopecten yessoensis and Argopecten irradians) allowed Zhang et al. (2007) to see a highly variable distribution pattern of the gene, interpreted as indicating that gene duplication/diminution and chromosome rearrangements play an important role in genome evolution in the group. Many mitochondrial sequences have been isolated from pectinids (e.g. Barucca et al., 2004; Matsumoto and Hayami, 2000) and those with rDNA sequences and chromosomal distribution (Insua et al., 2006; López-Piñón et al., 2005; Odierna et al., 2006; Wang and Guo, 2004) have allowed some evolutionary relationships to be inferred.

As in other species, the study of the rapidly evolving genomic component of tandemly repeated satellite DNA – its primary sequence, abundance and genomic distribution – in molluscs is likely to be important for understanding species relationships, evolution and the generation of new diversity. In this study, we aimed to identify major repetitive DNA sequence families in the scallop, which diverged from sequences present in oysters and mussels. We then examined features of the sequences including their genomic organization in the context of the genome by Southern, nuclear and fibre in situ hybridisation.

2. Materials and methods

2.1. DNA isolation, cloning and Southern hybridisation

Scallops, oysters and mussels (Pecten maximus, Crassostrea gigas, Mytilus edulis) were purchased from local markets. The scallops were diver-caught from commercial fisheries off the North-East coast of Scotland between January and May. DNA was isolated using standard methods following Canapa et al. (2000).

To identify major repetitive DNA sequences, genomic libraries were constructed after partial digestion of DNA from scallop with \(Pst\)I and \(Apa\)I restriction enzymes. Fragments were excised from a 1.5% agarose gel and ligated in pUC19. After transformation, colonies with inserts were arrayed on 90 mm plates, DNA was transferred to charged membranes and the inserts screened by hybridisation following standard techniques (Sambrook et al., 2000). Parallel sets of membranes were probed with genomic DNA from scallops, oysters and mussels labelled with digoxigenin by random-primer labelling (Invitrogen). Colonies showing strong hybridisation to scallop, but not mussel or oyster, DNA were selected, grown and the inserts in the plasmids sequenced commercially. These were named in the form of \(PmPst8\) (\(Pm=\) Pecten maximus; \(Pst=\) first letters of restriction enzyme; number=serial number of clone) and submitted to the EMBL database (accession nos. AM279151, AM279152, AM279153, AM279154, AM279155, AM263449, AM282563, AM279461, and AM279462). Tandem repeats were analysed using dotplots (Sonnhammer and Durbin, 1995).

Genomic DNA (5 μg) from scallops was digested with \(Apa\)I, \(Pst\)I, \(Hae\)III, \(Eco\)RI, \(Hind\)III, and \(Msp\)I and genomic DNA from mussels and oysters was digested with \(Pst\)I and \(Hae\)III. Digests were electrophoresed on a 1% agarose gel and transferred onto positively charged membrane (Roche). A non-radioactive hybridisation method was used with digoxigenin-labelled genomic DNA as probe (Schwarzacher and Heslop-Harrison, 2000).

Two of the repetitive elements, \(PmPst3\) and \(PmApa22\), with 178 and 156 bp long monomers were selected for analysis by real-time PCR with the Chromo4 System (Bio-Rad). Primers used were \(Pm\)3...30F: 5′-GGGAGATGTGATCAACAGG-3′ and \(Pm\)3...283R: 5′-CGCGTTACGTGGTTAATTG-3′, giving a 76 bp monomer fragment product and 254 bp dimer product, and \(Pm\)22...13F: 5′-CCCTTGGGTCAGAGTC-3′ and \(Pm\)22...231R: 5′-GAAACATCCTTGGTGAAGG-3′ giving a 219 bp product. The SYBR Green JumpStart Taq ReadyMix (Sigma) was used for quantitative PCR following the manufacturer’s instructions, with a 20 μl reaction volume including 0.5 μM final concentration of each primer, the ReadyMix buffer and six or seven concentrations over a 500 or 1000-fold range of genomic DNA from scallops, oysters, mussels and the isolated plasmids containing cloned repeats. The PCR program followed denaturation at 94 °C for 1 min, followed by 40 cycles of denaturation at 94 °C (10 s), annealing and extension at 57 °C (17 s) and a 5 s pause at 80 °C during which SYBR green fluorescence was measured. Primers and conditions were adjusted to give single major products corresponding to 219 bp (\(PmApa22\)) and 254 bp (\(PmPst3\)). A short (17 s) low temperature (57 °C) step was found to be optimum for annealing and amplifying only a single product from the tandem repeats; higher temperatures would have increased any selectivity of the primers and made the \(Taq\) polymerase extension faster. The primers which gave optimum products also gave some primer-related PCR products which had a low denaturation temperature; hence the quantitative fluorescence readings were made at 80 °C to denature these products and only measure the
monomers of the tandem repeat. Real time PCR showed exponential amplification of both repeat families 1 and 5 from the clones and genomic DNA from scallops and mussels, with correlation coefficients ($r^2$) greater than 0.99 over the appropriate range of template DNA concentrations. Copy numbers were calculated from the difference in threshold cycle number ($C_t$) between standards and genomic DNA products, based on the average of three or four independent reactions on different days. No quantitative amplification of products closely related to the repeat families was seen from oyster genomic DNA with these primers.

2.2. In situ hybridisation

Scallops between 5 and 8 cm in diameter were obtained from Millport Research Station, Isle of Cumbrae, Scotland for in situ hybridisation to extended DNA fibres. A small tissue portion was excised from the adductor muscle of *Pecten maximus* and homogenised in 500 μl of Nuclei Isolation Buffer (NIB; Schwarzacher and Heslop-Harrison, 2000) containing 0.1% mercaptoethanol. The suspension culture was filtered through a 64 μm filter to remove the large tissue clumps before centrifugation at 2000 × g for 10 min at 4 °C and resuspension in NIB containing 10% Triton-X and mercaptoethanol. Chromosome fibres were stretched on poly-L-lysine coated slides as described by Schwarzacher and Heslop-Harrison (2000).

For interphase chromosome preparation, scallops were treated with 0.005% colchicine in seawater for 16 h before fixation of gill tissue in fresh 3:1 ethanol:acetic acid. Preparations were made by maceration of tissue in 45% acetic acid and dropping onto chromic-acid cleaned slides. Fluorescent in situ hybridisation followed standard procedures (Schwarzacher and Heslop-Harrison, 2000). 5S rDNA was amplified by PCR from *P. maximus* genomic DNA with the primers 5′-AGCCCGGTAGTAGTACTTGG-3′ and 5′-CCGACGTTGCTTAACTTCG-3′ under usual conditions (López-Piñón et al., 2005) and was labelled by PCR with digoxigenin-11-dUTP (Roche) as a control for hybridisation. Pm probes were labelled with biotin-16-dUTP by random primer labelling (Invitrogen Bioprime kit). The probe mixture consisted of 50% formamide, 2 × SSC, 10% dextran sulfate, 2 μg salmon sperm DNA, 125 μM EDTA, 0.125% SDS and 10–15 ng of the labelled probes (5S rDNA or Pm series probe) in 50 μl. The probe mixture was denatured at 80 °C for 10 min then cooled on ice, placed on the slides, covered with a plastic coverslip and the probe and preparation was then denatured together at 75 °C for 5 min before cooling slowly to 37 °C for overnight hybridisation. The probe mixture and weakly hybridised probe were washed off with the most stringent wash being carried out in 20% formamide and 0.2 × SSC at 42 °C, allowing probe sequences with more than 80% similarity to the target chromosomes to remain stably hybridised. Biotin and digoxigenin probes were detected with streptavidin-Alexa594 (Molecular Probes) and anti-digoxigenin-FITC (Roche) respectively using standard methods. Preparations were counterstained with DAPI before mounting in Citifluor antifade mountant (Schwarzacher and Heslop-Harrison, 2000). Images were taken using a Zeiss Axioplan fluorescent microscope with a CCD camera and

Fig. 1. Restriction enzyme digestions of genomic DNA and Southern hybridisation of tandemly repeated DNA probes from scallop. Tracks in each panel, left to right: Bioline Hyperladder I; Scallop *ApaI, PstI, HaeIII, EcoRI, HindIII, MspI*; Mussel *PstI, HaeIII*; Oyster *PstI, HaeIII* genomic DNA digests; Hyperladder I. A) Ethidium bromide stained gel; B) PmPst15 (family 1); C) PmPst3 (family 1); D) PmPst8 (family 2); E) PmPst49 (family 3); F) PmPst50 (family 4); G) PmApa22 (family 5).
processed in Photoshop using only functions which affect the whole image equally (except for filling small areas around cropped pictures in Fig. 5C).

3. Results

3.1. Isolation of repetitive DNA

Digests of genomic DNA from scallops with a total of eight restriction enzymes showed no clear bands representing restriction satellites (Fig. 1). DNA was cloned from the Apal and PstI digests (Fig. 1A, tracks 1 and 2) and the resultant colonies were hybridised with labelled genomic DNA from scallops, oysters and mussels. Out of about 200 colonies screened, 13 showed strong hybridisation with scallop but not mussel or oyster DNA, and these clones were selected for growing and sequencing. Three clones had short inserts without any sub-repeat structure and were not further analysed. Eight clones between 127 bp and 891 bp long included multiple copies of six non-homologous tandem repeats. Southern hybridisation with all repeat families to genomic DNA digests showed ladders of hybridised bands, characteristic of tandemly repeated sequences. All clones hybridised strongly to scallop DNA, with evidence for higher order structure in the sequence arrays by the presence of particular fragments and ladders up to relatively high molecular weight. There was almost no detectable hybridisation to the oyster DNA digests (right hand two tracks in each gel) from any clone, while some showed weak hybridisation to mussel DNA.

3.2. Analysis of repetitive DNA

Clones PmPst15 (including five complete units) and PmPst3 (including one complete unit) are members of the tandemly repeated DNA family named family 1; these units measure ca. 178 bp (range 177 to 180, average similarity between all 6 copies 89.9%). There was no significant similarity with other sequences in the Genbank/EMBL database. Southern hybridisation (Fig. 1B and C) showed a 178 bp ladder characteristic of a tandemly repeated unit which was found in the Apal, PstI, HaeIII and MspI digests; short fragments (< 100 bp) hybridising to Mbol and Sau3a digests were consistent with the presence of multiple restriction sites for these enzymes in the sequence (data not shown). The sequences PmPst15 and PmPst3 show internal PstI sites (consistent with their presence in the partial digest used for cloning), although, despite showing a ladder by hybridisation, no Apal, HaeIII or MspI sites were present in the sequences, indicating that the clone was a variant of the repeat without these sites. The clone showed some similarity with mussel but not oyster DNA, and the PCR primers designed for scallops did not amplify an appropriate product from oyster DNA. Real-time PCR showed that the 178 bp repeat was present in 10,190 copies in the scallop genome, representing 0.13% of the DNA. The copy number in mussel was calculated as 113 per genome, consistent with the Southern hybridisation results.

The second repeat family 2, represented by PmPst8, consisted of 11 units of a short repetitive unit of 38 bp (Fig. 2, the final G is missing in the last unit) with an average of 90% similarity. In genomic DNA digests, a c. 200 bp ladder was revealed in some digests indicating a higher-order structure. The PstI site (CTGTAG), used to construct the library, flanked the PmPst8 clone (seen in the last bases CTGTA/G, where/indicates the cutting site), defining a 418 bp unit consistent with the dimer band seen in the Southern hybridisation. Families 3 (PmPst49) and 4 (PmPst50) were also short, with a 44 bp repeat. Family 4 included a 22 bp sub-repeat and showed minimal hybridisation to high molecular weight DNA from mussels, in contrast with the other five families which showed weak hybridisation.

Repeat family 5, represented by PmApa22 (Fig. 3), consisted of two 156 bp monomer repeat units. The sequence was noteworthy for its enrichment in runs of single nucleotides; more than 75% of the 82 G residues were in the context of two or more Gs, with four occurrences of AGGG in each monomer. GG, AA, TT and CC all occurred significantly more often than expected given
the base composition of the sequence (Chi-squared = 19, 7.2, 4.9, 14 respectively, statistic for 1 dof, 3.84 for $P = 0.05$ and 10.83 for $P = 0.001$). One monomer included three closely positioned HaeIII sites. The other repeat family analyses showed no overall significant runs of single nucleotides. Real-time PCR indicated a copy number of 272 with homology to the primers.

The repeat in family 6 (PmPst9) had a monomer size of 326 bp, and each monomer included two subrepeats of 163 bp with lower (75%) similarity at DNA level (Fig. 4). Comparison with the EMBL sequence database revealed an expressed mRNA sequence clone from the Bay scallop (*Argopecten irradians*), CN783316 (Roberts and Goetz, 2003) which included five copies of a 143 bp repeat, homologous to repeat family 6 in the 5′ end of the submitted sequence. Although the sequence was from a spat EST library, no open reading frame was evident in the sequence. The family 6 sequence also

---

Fig. 4. Alignment of repeat family 6 PmPst9 with 3.5 monomers, Pm9-a, b, c and -d, divided into two more degenerate units, 1 and 2 with 75% similarity. Part of an EST sequence from *Argopecten*, Ap-CM783316 (reverse complement) aligns well with the first part of the Pecten family 6 sequence.

---

Fig. 5. Interphase and fibre in situ hybridisation of selected repetitive DNA motifs (labelled red) and the 5S rDNA probe (labelled green) in nuclei counterstained blue with DAPI. The green 5S rDNA site shows four (two pairs) of major sites and dispersed hybridisation arising from decondensation and expression. A) Family 1 (PmPst15) showed discrete hybridisation sites varying in size and excluded from nucleoli, indicating its presence at multiple chromosomal locations, and very long, continuously labelled DNA fibres were present indicating that the tandem repeats occurred in long, uninterrupted arrays. In one nucleus from a probably dying cell, the sites of family 1 had aggregated (right of picture). B) Family 2 (PmPst8), showing two major sites of the repeat ($x$ indicates site of green precipitate). C) A composite image showing family 5 (PmApa22) located at some six major sites with diffuse hybridisation indicating location at multiple smaller sites in the genome. D) In situ hybridisation of PmPst15 (family 1; red) and the 5S rDNA sequence to extended chromatin fibres showing non-interspersed arrangements; E) Family 5 (PmApa22) hybridised to extended chromatin fibres showing abundance and some shorter fibres. A, B, C scale bar = 10 μm; D, E scale bar = 5 kb.
showed limited similarity (70% over 80 bp) with a satellite DNA sequence from the bivalve Donax trunculus with a monomer length of 155 bp (Plohl and Cornudella, 1996). Two further abundant sequences, PmPst5 and PmPst6, did not include a clear repetitive unit, but PmPst10 showed significant similarity (76% over 259 bp) with the database sequence isolated from Chlamys farreri (EMBL accession no. AY682116; Zhan et al., 2006b) in the Pectinidae.

Interphase and fibre in situ hybridisation of selected repetitive DNA motifs showed that they had different genomic distributions (Fig. 5). Family 2 (PmPst8) was concentrated on a single pair of chromosomes, showing two distinct sites in each nucleus. Family 1 (PmPst15) showed many discrete hybridisation sites varying in size, indicating its presence at multiple chromosomal locations, and very long, continuously labelled DNA fibres were present indicating that the tandem repeats occurred in long, uninterrupted arrays. In contrast, family 5 (PmApa22) shows six to eight major sites with some more diffuse hybridisation, in agreement with its lower genomic abundance. The presence of shorter labelled fibres suggested the presence of dispersed repetitive elements. With all probes and the 5S rDNA (seen hybridising to uninterrupted arrays: not shown) sequences, no hybridisation was regularly interspersed. This result provides a control of the specificity of hybridisation and shows that the elements were not interspersed with each other.

4. Discussion

4.1. Scallop repetitive DNA family structures

Six families of moderate/low copy, non-homologous, tandemly repeated DNA sequences were isolated from scallop (Pecten maximus). They differ in monomer length, presence of similar sequences in other bivalves and in the variation in sequence between copies of the repeats. Three families have short monomer lengths of 38, 44 (with a 22 bp subrepeat) and 45 bp, while the remaining families have repeats of 156, 178 and 326 bp, characteristic of the length folding around one or two (326 bp) nucleosomes. The 156 bp sequence of family 5 was unusual in having multiple runs of the same base although no substructure could be regarded as a minisatellite or microsatellite. Consistent with their diversity and moderate copy number, no major restriction bands were seen in genomic DNA digests. Petrović and Plohl (2005) were able to identify two non-homologous tandemly repeated sequences, both 169 bp long, by electrophoresis of HindIII digests of genomic DNA from the Coquina clam, Donax trunculus. They quantified the abundance of the family named DTF1 and found it represented 0.1% of the genomic DNA. Even in Pecten maximus we found 0.13% abundance of the repeats of the PmPst3 family 1 quantified here by real-time PCR, and consistent with the hybridisation results shown. The Thiriot-Quévauvrez (2003) review detailed morphological analysis and differential staining methods (including silver-staining, C- and G-banding) applied to molluscan chromosomes, thereby allowing elucidation of cytotaxonomic relationships within the class. However, as pointed out by Wang and Guo (2004), at least in the Queen scallop the chromosomes show limited band differentiation (Insua et al., 1998), and bands alone are not distinctive or reliable enough for practical routine chromosome identification (Wang and Guo, 2004). Restriction enzyme banding of chromosomes has been shown to allow many chromosomes to be identified in the oyster (Leitão et al., 2002, 2004), perhaps because this leads to removal of particular sequences rather than requiring differential staining of large blocks of repetitive DNA. Both methods complement the analysis of repetitive DNA probes, and suggest that the scallop genome does not have extremely abundant repetitive DNA sequence families.

4.2. Repetitive DNA evolution

An alternative approach to analysis of repetitive DNA has been taken in the Cupped oyster, Crassostrea virginica by Gaffney et al. (2003) who identified several abundant sequence families by analysis of 730 kb of arbitrarily cloned genomic survey sequence. One sequence family, CvA, was present in 9.7% of the ≈1 kb clones analysed and was related to transposable elements, including inverted repeats and degenerate microsatellite motifs, along with a tandemly repeated core region. Three other less abundant elements with some similarities in structure to CvA were also found. CvB, for example, included a unique region followed by tandem repeats of a 168 bp element, and a number of imperfect microsatellites (named ‘protomicrosatellites’ by the authors), along with other nucleotide repetitions. It is interesting to note that some of the repetitive DNA sequence families isolated here from scallops had some short and non-random nucleotide repetitions, although none showed any conspicuous inverted repeats, secondary structures or other transposon-related characteristics. Various authors have reported microsatellite markers from bivalves. However, the length of the repeats and number of repetitions is notably variable and relatively low compared with mammals. Reece et al. (2004), for example, reported dinucleotide, trinucleotide and tetranucleotide microsatellites of only 8, 9 or 12 bases, and showed that these have useful variability in allele size between accessions. Cruz et al. (2005) reported that microsatellite motifs showed a low genomic density compared with other eukaryotes, even considering only 5 dinucleotide repeats as a microsatellite. In most organisms microsatellite arrays would be longer, and perhaps this relates to the protomicrosatellite sequences seen in the larger satellite DNA families analysed here, and may relate, like satellite DNA evolution, to slippage replication.

Our method of isolation involving genomic libraries and screening for repetitive sequences with scallops which were not abundant in oysters or mussels should have selected abundant and evolutionarily conserved transposons present in all three genomes. The CvA and CvB families (Gaffney et al., 2003) had core repeats that were conserved among widely diverged molluscs ranging from oysters to blood arks, clams and even Drosophila. These authors point out that it would be interesting to investigate the relationship of their transposon-related sequences with satellite DNA. It is probable that we selected against transposon-like elements (both class I retrotransposons
and class II DNA elements) which would show sequence conservation among oysters, mussels and scallops.

Other tandemly repeated families, in particular those related to centromere organization and function, are found in many molluscs, and Canapa et al. (2000) have reported a sequence from the Antarctic scallop, *Adamussium colbecki*, which includes regions similar to the mammalian CENP-B box. A centromere-located satellite conserved across oyster species, that is approximately 168 bp long and represents 1% to 4% of the genome, has been isolated by Clabby et al. (1996) and independently by López-Flores et al. (2004), who also showed that it had transposon characteristics, which are substantially more abundant than the copy number of the genus-specific repeats analysed here. The tandem repeat families which we isolated from scallops (like the families isolated from mussels by Martinez-Lage et al., 2002, 2005) show specificity to the single genus, indicating that new or diverged sequences have appeared and amplified during speciation. This evolution of tandem repeats during speciation is a characteristic of many tandem array families (although by no means all, particularly not including transposon-related sequences and coding repeats) in both mammals (e.g. Chaves et al., 2005) and plants (Heslop-Harrison, 2000), and the Nijman and Lenstra model (2001) suggests that rapid amplification of homogeneous repeat units is followed sequentially by mutation and independent amplification of coexisting sequence variants.

Within both family 1 and family 2, we detected sequence variants (90% similarity between copies of family 1 and family 2 where 6 and 11 monomers were analysed, respectively). The lengths of the 11 family 2 monomers were conserved. These results are consistent with the careful analysis carried out by Petrović and Plohl (2005) who sequenced ten copies of DTF1, a 169 bp tandem repeat from clams. They also identified two regions that were conserved in nine of the sequences, and 169 bp tandem repeat from clams. They also identified two regions similar to the mammalian CENP-B box. A tandem repeat families which we isolated from scallops (like the families isolated from mussels by Martínez-Lage et al., 2002, 2005) show specificity to the single genus, indicating that new or diverged sequences have appeared and amplified during speciation. This evolution of tandem repeats during speciation is a characteristic of many tandem array families (although by no means all, particularly not including transposon-related sequences and coding repeats) in both mammals (e.g. Chaves et al., 2005) and plants (Heslop-Harrison, 2000), and the Nijman and Lenstra model (2001) suggests that rapid amplification of homogeneous repeat units is followed sequentially by mutation and independent amplification of coexisting sequence variants.

Both within family 1 and family 2, we detected sequence variants (90% similarity between copies of family 1 and family 2 where 6 and 11 monomers were analysed, respectively). The lengths of the 11 family 2 monomers were conserved. These results are consistent with the careful analysis carried out by Petrović and Plohl (2005) who sequenced ten copies of DTF1, a 169 bp tandem repeat from clams. They also identified two regions that were conserved in nine of the sequences, and 169 bp tandem repeat from clams. They also identified two regions similar to the mammalian CENP-B box. A tandem repeat families which we isolated from scallops (like the families isolated from mussels by Martínez-Lage et al., 2002, 2005) show specificity to the single genus, indicating that new or diverged sequences have appeared and amplified during speciation. This evolution of tandem repeats during speciation is a characteristic of many tandem array families (although by no means all, particularly not including transposon-related sequences and coding repeats) in both mammals (e.g. Chaves et al., 2005) and plants (Heslop-Harrison, 2000), and the Nijman and Lenstra model (2001) suggests that rapid amplification of homogeneous repeat units is followed sequentially by mutation and independent amplification of coexisting sequence variants.

4.3. Conclusions

The results here show that the scallop genome has a number of unrelated moderately abundant families of tandemly repeated DNA with contrasting characteristics. The families have limited similarity with those reported in other molluscs even within the Ostreida, and hence have shown rapid evolutionary divergence in sequence and/or copy number. The study of these rapidly evolving sequences is not only valuable for understanding an important source of genomic diversity, but also has the potential to provide useful markers for population studies, and gives a route for identifying mechanisms of DNA sequence evolution.

References


