J Mol Evol (2010) 70:413–426
DOI 10.1007/s00239-010-9341-3

Evolutionary Dynamics of the 5S rDNA Gene Family in the Mussel *Mytilus*: Mixed Effects of Birth-and-Death and Concerted Evolution

Ruth Freire · Alberto Arias · Ana M. Ínsua · Josefina Méndez · José M. Eirín-López

Received: 7 January 2010 / Accepted: 30 March 2010 / Published online: 13 April 2010

Abstract In higher eukaryotes, the gene family encoding the 5S ribosomal RNA (5S rRNA) has been used (together with histones) to showcase the archetypal example of a gene family subject to concerted evolution. However, recent studies have revealed conspicuous features challenging the predictions of this model, including heterogeneity of repeat units, the presence of functional 5S gene variants as well as the existence of 5S rDNA divergent pseudogenes lacking traces of homogenization. In the present work, we have broadened the scope in the evolutionary study of ribosomal gene families by studying the 5S rRNA family in mussels, a model organism which stands out among other animals due to the heterogeneity it displays regarding sequence and organization. To this end, 48 previously unknown 5S rDNA units (coding and spacer regions) were sequenced in five mussel species, leading to the characterization of two new types of units (referred to here as small-β 5S rDNA and γ-5S rDNA) coexisting in the genome with α and β rDNA units. The intense genetic dynamics of this family is further supported by the first description of an association between γ-5S rDNA units and tRNA genes. Molecular evolutionary and phylogenetic analyses revealed an extensive lack of homology among spacer sequences belonging to different rDNA types, suggesting the presence of independent evolutionary pathways leading to their differentiation. Overall, our results suggest that the long-term evolution of the 5S rRNA gene family in mussels is most likely mediated by a mixed mechanism involving the generation of genetic diversity through birth-and-death, followed by a process of local homogenization resulting from concerted evolution in order to maintain the genetic identities of the different 5S units, probably after their transposition to independent chromosomal locations.

Keywords Ribosomal RNA · 5S subunit · *Mytilus* · Molecular evolution · Birth-and-death · Concerted evolution

Electronic supplementary material The online version of this article (doi:10.1007/s00239-010-9341-3) contains supplementary material, which is available to authorized users.

R. Freire · A. Arias · A. M. Ínsua · J. Méndez · J. M. Eirín-López (✉)
XENOMAR Group, Departamento de Biología Celular y Molecular, Facultade de Ciencias, Universidade da Coruña, Campus de A Zapateira s/n, 15071 A Coruña, Spain
e-mail: jeirin@udc.es

Introduction

In higher eukaryotes, ribosomal DNA (rDNA) comprises two different gene families (Long and Dawid 1980), including a major family encoding 18S, 5.8S, and 28S rRNA, and a minor family encoding 5S rRNA. The genomic organization of 5S rDNA is characteristic by showing multiple repeats, tandemly arranged at one or several chromosomal locations, with each repeat including a highly conserved coding region and a variable nontranscribed spacer. However, alternative arrangements have been reported for 5S rDNA genes, including the presence of additional scattered 5S copies (Little and Braaten 1989), the linkage of 5S genes with other families such as 18–28S rDNA (Gerbi 1985), snRNAs (Manchado et al. 2006), trans-spliced leader and histone genes (Drouin and Moniz de Sá 1995; Eirín-López et al. 2004), as well as with microsatellite sequences (Cross and Rebordinos 2005).

Given the apparent homogeneity observed among the different copies, 5S genes (together with histones) have been used to showcase the archetypal example of a gene

family subject to concerted evolution. Under this model, repeated gene family members evolve together as a block displaying a high degree of homogeneity, as they diverge respect to repeats from other species (Arnheim 1983; Dover 1982). However, the theoretical expectations made by the concerted evolution model are challenged by three major molecular evolutionary features displayed by the 5S rDNA family: First, several 5S gene variants have been found in some fish and amphibian species, constituting a dual system in which oocytes display a 5S rRNA type different from that found in somatic cells (Alves-Costa et al. 2006; Komiya et al. 1986). Nevertheless, gene variants are not only circumscribed to this dual system (Cronn et al. 1996; Lazar et al. 1983). Second, 5S rDNA divergent pseudogenes (lacking traces of homogenization) have been found in unrelated taxa such as humans (Nielsen et al. 1993), fishes (Martins et al. 2002), flies (Sharp and Garcia 1988), and filamentous fungi (Rooney 2003). Third, the existence of different types of repeat units has been also corroborated based on the study of spacers. In this sense, it has been proposed that the presence of such variants could respond to a hybrid origin of the species, such as for instance in the case of cereals (Baum et al. 2009; Kellogg and Appels 1995). In the light of these data, different authors have proposed that the variation observed among 5S rDNA members best fits to a birth-and-death model of long-term evolution promoting genetic diversity (Fujiwara et al. 2009; Rooney and Ward 2005). Accordingly, new 5S copies are created by recurrent gene duplication events and some copies are maintained in the genome for a long time whereas others are deleted or become nonfunctional through deleterious mutations (Nei and Rooney 2006).

Bivalve molluscs stand out among protostome animals for being one of the most widely studied group of organisms as it pertains to 5S rDNA, encompassing high levels of gene variation as well as a vast diversity of gene arrangements. The molecular organization of 5S rDNA has been studied in cockles (Freire et al. 2005; Insua et al. 1999), mussels (Insua et al. 2001), oysters (Cross and Rebordinos 2005), scallops (Lopez-Piñon et al. 2008), and razor-clams (Fernandez-Tajes and Mendez 2009; Vierna et al. 2009). Mussels are of special interest within this taxonomic group due to the heterogeneity they display in 5S rDNA organization, including two types of repeat units (referred to as α and β units) with divergent spacers characterized in *Mytilus galloprovincialis* and *M. edulis* (Insua et al. 2001). Furthermore, additional amplification products have been observed in both species, suggesting that additional 5S rDNA types could coexist with those already described in the genome of these organisms.

In the present work, we have broadened the scope in the study of protostome ribosomal RNA gene families by characterizing 48 new 5S rDNA units (coding and spacer regions) in five mussel species. Our results reveal the presence of high levels of gene variation and diversity of gene arrangements, including the description of two types of units previously unknown, referred to here as small-β 5S rDNA (possibly representing a pseudogenized 5S type) and γ-5S rDNA, the latter one coexisting in the genome of mussels with the previously described α and β units. The intense genetic dynamics of this family is further supported by the discovery (for the first time) of an association between γ-5S rDNA units and tRNA genes. Molecular evolutionary analyses unveil an extensive lack of homology among spacer sequences belonging to different rDNA types (α, β, and γ), suggesting the presence of independent evolutionary pathways leading to their differentiation. Overall, our results suggest that the long-term evolution of the 5S rDNA in mussels is most likely mediated by a mixed mechanism involving the generation of genetic diversity through birth-and-death followed by a process of homogenization acting at a local level in different 5S arrays.

Materials and Methods

Sample Collection and DNA Extraction

Mytilus trossulus specimens were collected in the Gulf of Gdansk, Baltic Sea (Poland, European sample) and in Esquimalt Lagoon, Vancouver Island (Canada, American sample). *M. californianus* specimens were collected in Point no Point at Vancouver Island (Canada), and *M. coruscus* were collected in Otsuchi Bay (Japan). *M. galloprovincialis* and *M. edulis* individuals were collected in Galicia (Spain) and Zeeland (Holland), respectively. Genomic DNA was extracted after two 15-min washes in phosphate-buffered saline and sterile deionized water (Winnepenninckx et al. 1993).

PCR Amplification, Cloning, and Sequencing

PCR reactions were performed in a final volume of 25 µl using the primers A–B and C–D (Insua et al. 2001) from a single individual in each of the species studied. The resulting products were electrophoresed in 2% agarose gels and subsequently purified using the Geneclean kit (Qbiogene). The resulting DNA fragments were cloned in the pGEM-T Easy Vector System II (Promega) and plasmid DNA was purified using the alkaline lysis method (Birnboim and Doly 1979). Sequencing reactions were performed using an ALFexpress automatic sequencer (GE Healthcare). The obtained nucleotide sequences have been deposited in the EMBL nucleotide database with accession numbers FN561814-61 (see Supplementary Table 1 for details).

Molecular Evolutionary Analyses

A total of 62 5S rDNA repeat units (encompassing coding and spacer sequences) belonging to five different *Mytilus* species were analyzed. Among them, 48 units are described for the first time in the present work (see Supplementary Table 1 for details). The identity of the sequences obtained was corroborated by using the BLAST program (Altschul et al. 1990). Multiple sequence alignments were conducted and edited for potential errors using the CLUSTAL_X (Thompson et al. 1997), BIOEDIT (Hall 1999), and GENEDOC (Nicholas et al. 1997) programs. The extent of the nucleotide variation was estimated by means of different methods depending on the nature of the data as follows: In one hand, within species nucleotide polymorphism (π) was estimated as the average number of nucleotide differences per site between two sequences; and between species nucleotide divergence (K) was estimated as the average number of nucleotide substitutions per site between species (using in both cases the Jukes and Cantor correction), using the DnaSP version 5 program (Librado and Rozas 2009). On the other hand, the extent of the nucleotide variation across species both within and between types of units was estimated by means of the Kimura 2-parameter evolutionary distances (d) using the complete deletion option in all cases, using the MEGA version 4 program (Tamura et al. 2007). Standard errors of the estimations were calculated by using the bootstrap method (1,000 replicates).

The neighbor-joining tree-building method (Saitou and Nei 1987) was used to reconstruct the phylogenetic trees in this work. In order to assess that our results are not dependent on this choice, phylogenetic analyses were completed by reconstructing maximum parsimony trees. The reliability of the resulting topologies was tested by using both the bootstrap (Felsenstein 1985) and the interior branch-test (Sitnikova 1996) methods, producing the bootstrap (BP) and confidence probability (CP) values, respectively, for each interior node in the trees after 1,000 replicates. Given the known conservative nature of the bootstrap method, BP $>$ 80% was interpreted as high statistical support for groups, whereas CP $\geq$ 95% was considered statistically significant (Sitnikova et al. 1995).

Nucleotide Variation Across 5S Coding Regions and Reconstruction of Ancestral Sequences

The analysis of the nucleotide variation across 5S coding regions was performed using a sliding-window approach as implemented in the program DnaSP version 5 (Librado and Rozas 2009) by estimating the nucleotide polymorphism within species (π) and the nucleotide divergence between species (K), as well as the nucleotide variation across species both within and between types of units (d), calculated as detailed above. A window length of 10 bp and a step size of 5 bp were used. Ancestral sequences corresponding to the internal nodes of the phylogenies of α, β, and γ-5S spacer regions were reconstructed by maximum likelihood using the baseml program within the PAML ver. 4.1 package (Yang 2007), allowing the estimation of the nucleotide substitutions involved in the differentiation among *Mytilus* 5S spacer sequences. The statistical selection of best-fit models of nucleotide substitution for spacer regions was performed by using the program jModelTest version 0.1.1 (Posada 2008) using the Akaike Information Criterion (AIC) for model selection. The following models were defined as the two most fitted to the data (only those indicated in first position were used for the reconstruction of ancestral sequences): K80/TPM3uf (Kimura 1980) with 5 free parameters and unequal base frequencies ($-\ln L = 259.2374$, AIC $= 580.4747$), and HKY (Hasegawa et al. 1985) ($-\ln L = 261.6628$, AIC $= 583.3527$) for α-5S spacers; HKY + G (Hasegawa et al. 1985) including rate variation among sites ($-\ln L = 2880.8982$, AIC $= 5875.7963$), and K80 + G/TPM3uf + G (Kimura 1980) with 5 free parameters and unequal base frequencies including rate variation among sites ($-\ln L = 2880.6340$, AIC $= 5877.2680$) for β-5S spacers; HKY + I (Hasegawa et al. 1985) including a proportion of invariable sites ($-\ln L = 2461.6980$, AIC $= 4989.3959$), and HKY + G (Hasegawa et al. 1985) including rate variation among sites ($-\ln L = 2461.7249$, AIC $= 4989.4497$) for γ-5S spacers.

Results

Molecular Features of 5S rDNA Units in *Mytilus*

Amplifications of 5S rDNA units were performed using two different sets of primers (see "Materials and Methods" section), yielding three main fragments of approximately 250, 770, and 1000 bp in *M. californianus* and *M. trossulus* from the Canadian coast, representing the α, β, and γ-5S rDNA units, respectively. Amplifications from European *M. trossulus* representatives resulted in two main bands of approximately 250 and 770 bp, although some individuals also presented the 1000 bp band. In *M. coruscus* individuals 5S rDNA amplifications consisted of only one product of approximately 300 bp. The identity of the bands obtained was corroborated as 5S rDNA in all cases. Furthermore, two γ-5S rDNA specific primers were designed (F: 5′-AAA ATT GTC TAC CTC CGA CGG G-3′; and R: 5′-ACT AGC CTC ACA ACA TAG TCT GC-3′) and tested in *M. galloprovincialis* and *M. edulis* individuals, resulting in the 1000 bp fragment and evidencing that γ-5S rDNA units also occur in these mussel species.

Among the 62 5S rDNA repeat units from mussels used in the present study, 48 units are described for the first time in the present work (see Supplementary Table 1 for details). The coding region was identified in the units by using the 5S rRNA form the mussel *M. edulis* as a reference (Fang et al. 1982), starting with GTC at 5′ and ending with ACA at 3′ (119 bp in length for all sequences), agreeing with the consensus defined for the predicted 5S rRNA secondary structure in all cases except for the small-β type (Fig. 1). BLAST analyses of α, β as well as the newly described γ-5S rDNA units revealed scarce or total absence of homology among their spacer regions in *M. galloprovincialis* and *M. edulis*. Homology was also absent in comparisons with *M. coruscus* spacers. Surprisingly, the 250 bp fragment obtained in amplifications from *M. californianus* (previously referred to as α-5S rDNA due to its similarity in length to these type of sequences) displayed homology to the 3′ end of β spacer sequences, lacking homology with α sequences. These sequences were subsequently reassigned to a new 5S rDNA type defined to here as small-β type.

The length and base composition of the 5S rDNA units in *Mytilus* representatives is summarized in Table 1. The three types of 5S rDNA units sequenced (α, β, and γ) consisted of fragments falling within size ranges of 260–264, 751–794, and 995–1000 bp, respectively, in the case of *M. trossulus*; 240, 723–734, and 981–986 bp in the case of *M. californianus*; and a single size class of 311–314 bp in the case of *M. coruscus*. In the case of *M. galloprovincialis* and *M. edulis*, the newly described γ-5S rDNA unit consisted of 994–1000 and 1011–1014 bp, respectively. While the length of the spacer region was heterogeneous across different rDNA types and species, the GC content was always smaller in spacers when compared with coding regions. γ-5S coding regions display the highest average GC content among the 5S types (54.70%); on the other hand, the small-β type (which lacks the typical secondary 5S rRNA structure) displays the lowest GC average content (51.27%). In all cases, α spacers and γ spacers displayed the smallest and greatest average GC content, respectively.

Identification of a tRNA-Arg Gene Linked to γ-5S rDNA Units

The presence of a tRNA-Arg gene in the spacer regions of γ-5S rDNA units was revealed through BLAST analyses and confirmed with the program tRNAscan-SE version 1.21 (Lowe and Eddy 1997), occupying an opposite direction to the 5S rDNA gene in all cases. Table 2 details the exact location of such tRNA in all clones investigated. The secondary structure of the tRNA-Arg gene detected in *Mytilus* 5S was subsequently predicted (Fig. 2), and the alignment of the tRNA-Arg sequences showed five polymorphic sites (three substitutions and two indels). All sequences displayed the A and B boxes involved in transcription by RNA polymerase III (Paule and White 2000), with consensus sequences detailed in Table 3. Two differences (boldfaced in Table 3) were found in *Mytilus* consensus sequences with respect to the consensus sequence described for yeast tRNA (Marck et al. 2006). One of these differences is represented by an indel in one clone of *M. galloprovincialis* (Mg-γ3) which affects the secondary structure of tRNA. All sequences showed at least four T residues at 3′ end of coding region and no TATA elements were found upstream the coding region.

Phylogenetic Analysis of *Mytilus* 5S rDNA Units

Overall, the 5S coding region comprised 119 bp in all species (regardless of the type of unit) displaying reduced levels of nucleotide variation as expected. On the other hand, the high rate of evolution displayed by the spacer regions does not allow for their combined analysis from a single alignment and thus, independent alignments for coding regions (Supplementary Fig. 1), complete units (Supplementary Figs. 2, 3, and 4) and spacer regions (Supplementary Figs. 5, 6, and 7) belonging to α, β, and γ-5S types were performed. The alignment of complete α-5S rDNA units (coding + spacer) encompassed 265 positions, with the 5S coding region being virtually identical for all species. Indeed, the highest nucleotide diversity was found in the case of European *M. trossulus* representatives, being as low as 0.005 ± 0.002 differences per site. The alignment of β-5S rDNA units (863 positions) displayed 394 variable sites in the spacer region and only 2 in the coding region, leading to the highest nucleotide diversity levels in European and American *M. trossulus* representatives (0.158 ± 0.040 and 0.152 ± 0.043 differences per site, respectively). Finally, the alignment of γ-5S rDNA comprised 1092 positions, with nucleotide diversity values ranging between 0.001 ± 0.001 differences per site (in American *M. trossulus*) and 0.019 ± 0.006 differences per site (in *M. galloprovincialis*).

The phylogeny reconstructed based on 5S coding regions is shown in Fig. 3, depicting a high degree of conservation among sequences. The resulting topology reveals a clustering pattern based on the type of unit (instead of by species) in the case of γ-5S, in which the three European species (*M. galloprovincialis*, *M. edulis*, and European *M. trossulus*) are grouped in a single monophyletic group. β type sequences also follow a clustering pattern based on the type of unit in the case of *M. californianus*, representing a divergent group from the remainder of the β-type sequences, which intermingle extensively with α-5S sequences.

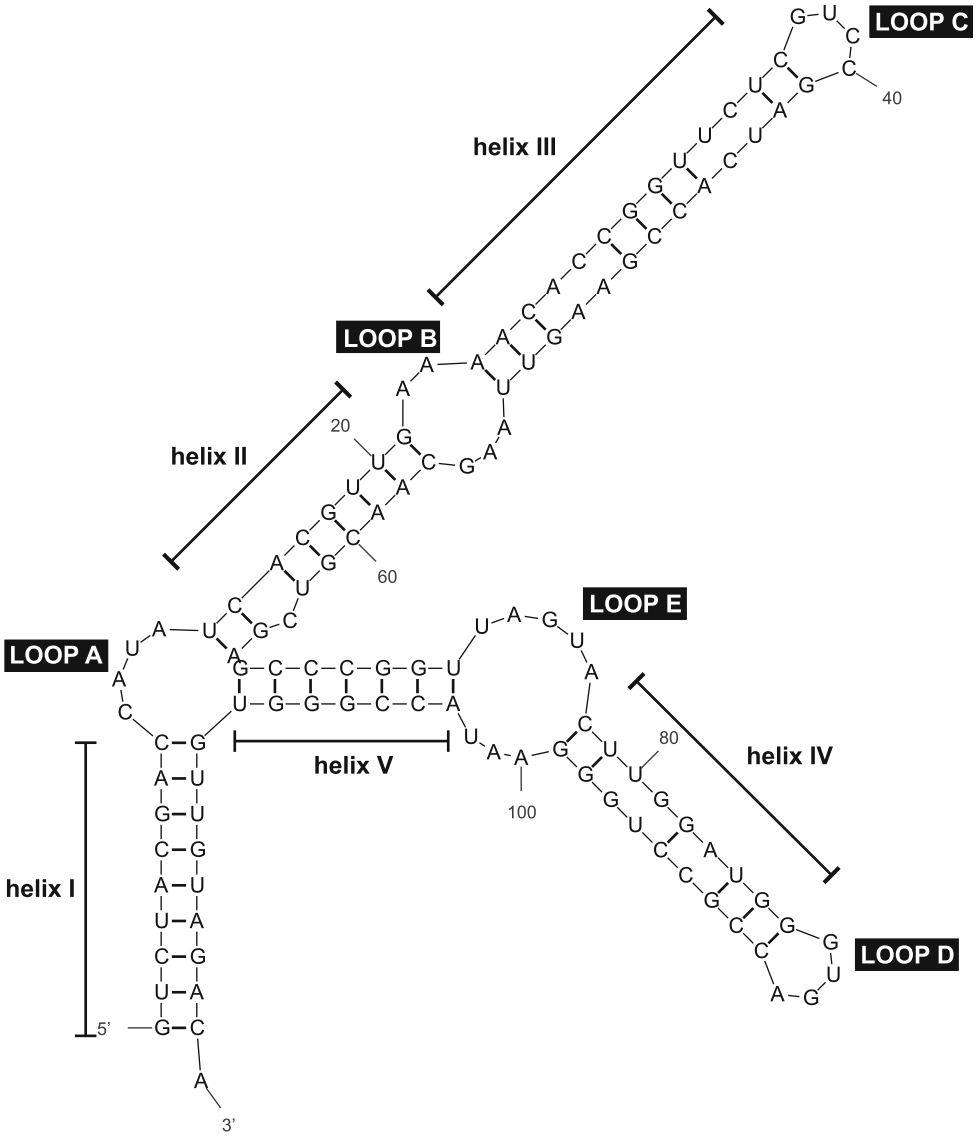

Fig. 1 Consensus secondary structure predicted for the 5S rDNA gene in *Mytilus*. The positions corresponding to different regions including helices I–V and loops A–E are indicated

The phylogenies reconstructed for the three different spacer regions reflect, as expected, high evolutionary rates when compared with 5S coding segments and the groups defined by the topologies are strongly supported by bootstrap and confidence probabilities. In the case of α-5S sequences (Fig. 4a), a species-based clustering pattern is evident, defining the α spacer from *M. trossulus* as the most divergent respect to the remaining European species. On the contrary, the tree topology based on *β*-5S sequences encompasses more complexity, highlighting the divergent nature of *M. californianus* *β* and small-*β* sequences (Fig. 4b), as well as revealing the absence of a species-specific clustering pattern, which is concomitant with the increased levels of intraspecific and interspecific variation. Furthermore, this topology discriminates between two subtypes of *β* type sequences with monophyletic origins, encompassing *M. galloprovincialis*, *M. edulis*, and *M. trossulus* on group 1, and *M. edulis and M. trossulus* on group 2 (Fig. 4b). Finally, phylogenetic inference based on *γ* type sequences unveils a topology depicting a species-specific clustering pattern very similar to that obtained in the case of α sequences, with *M. caliornianus* representing the most divergent species in this case (Fig. 4c).

Nucleotide Variation in Coding and Spacer Regions

The mechanisms underlying 5S rDNA evolution in *Mytilus* were further investigated by analyzing the nucleotide variation responsible for the differentiation of the 5S coding region into different species as well as into different types of units. Table 4 shows nucleotide variation estimates obtained within and between *Mytilus* species revealing that, with the exception of *M. coruscus*, polymorphism and divergence acquire approximately similar magnitudes in coding regions. The apparent absence of a clear interspecies differentiation process is further supported by the lack

Table 1 Size and GC content of 5S rDNA units in *Mytillus* representatives

Species/5S rDNA type	No. of clones	Total		Gene		Spacer	
		Length	GC (%)	Length	GC (%)	Length	GC (%)
M. galloprovincialis							
α-5S rDNA[a]	3	258	36.43	119	52.94	139	22.30
β-5S rDNA[a]	3	775	37.81	119	52.94	656	35.06
γ-5S rDNA	3	994–1000	39.74	119	54.9	875–881	37.68
M. edulis							
α-5S rDNA[a]	3	257	36.45	119	52.66	138	22.46
β-5S rDNA[a]	4	749–776	38.41	119	52.94	630–657	35.72
γ-5S rDNA	3	1011–1014	39.68	119	54.62	892–895	37.7
M. trossulus							
European coast							
α-5S rDNA	3	262	35.11	119	52.94	143	20.28
β-5S rDNA	3	751–776	38.08	119	52.94	632–657	35.34
γ-5S rDNA	3	995–1000	39.16	119	55.46	876–881	36.95
American coast							
α-5S rDNA	5	264	34.93	119	53.11	145	20
β-5S rDNA	4	761–794	37.3	119	53.15	642–675	35.06
γ-5S rDNA	3	996–1000	38.71	119	54.70	877–881	36.58
M. californianus							
Small-β-5S rDNA	7	240	38.02	119	51.27	121	25.03
β-5S rDNA	6	723–734	38.80	119	52.66	604–615	36.09
γ-5S rDNA	3	981–986	37.76	119	53.80	862–867	35.56
M. coruscus							
5S rDNA	3	311–314	39.01	119	54.11	192–195	29.83

[a] Data from (Insua et al. 2001)

of fixed differences between species in the 5S coding region. Indeed, the representation of polymorphism and divergence across the 5S coding region reveals a strong correlation between both estimates, reaching high values in the regions corresponding to the loops A, B, and C and the helices II, III, and V, which include the Box A and the Internal Element (Fig. 5a). Coding regions were further analyzed based on the type of unit to which they are appended (Table 5), unveiling low levels of nucleotide variation within each one of the three types in comparison with the variation determined between types. Again, the representation of the nucleotide variation across the 5S coding region unveils the highest values in the regions corresponding to the loops A, B, and C and the helices II, III, and V, which include the Box A and the Internal Element (Fig. 5b). Overall these results indicate that the evolutionary process leading to the diversification and differentiation of 5S coding regions seem to be predominantly based on the type of unit rather than on the concrete species to which they belong.

Given the limited variation detected in coding regions, the study of 5S rDNA evolution in *Mytilus* was complemented with the analysis of spacer regions. Table 6 shows polymorphism and divergence values in the three spacer regions of the different *Mytilus* species analyzed in the present work, indicating close intraspecies relationships. The high numbers of fixed differences between spacer sequences from different *Mytilus* species (which contrasts with the absence of such differences in coding regions as indicated in Table 4) reinforces the notion pointing towards a long-term evolutionary process promoting the differentiation of α, β, and γ rDNA types. Such a behavior in the dynamics of nucleotide substitutions is observed in comparisons between interspecies nucleotide divergence versus the numbers of fixed differences among *Mytilus* species (Fig. 6), revealing a significant positive correlation in the three types of units.

Discussion

General Overview of 5S rDNA Organization in *Mytilus*

Amplifications of 5S rDNA in *Mytilus* demonstrated that these mussel species exhibit, at least in part, the conventional tandem arrangement organization of 5S rDNA units.

Table 2 Location of tRNA in γ-5S rDNA sequences

Species/clone	Start	End	AA	Anti-codon
M. galloprovincialis				
Mg γ1	690	618	Arg	CCT
Mg γ2	691	619	Arg	CCT
Mg γ3	690	620	Arg	CCT
M. edulis				
Me γ1	709	636	Arg	CCT
Me γ2	706	635	Arg	CCT
Me γ3	709	637	Arg	CCT
M. trossulus				
Mt(E) γ1	685	613	Arg	CCT
Mt(E) γ2	683	611	Arg	CCT
Mt(E) γ3	680	608	Arg	CCT
Mt(A) γ1	680	608	Arg	CCT
Mt(A) γ2	683	611	Arg	CCT
Mt(A) γ3	683	611	Arg	CCT
M. californianus				
Mca γ1	664	592	Arg	CCT
Mca γ2	670	598	Arg	CCT
Mca γ3	664	592	Arg	CCT

Note: Mt(E) and Mt(A) denote European and American *M. trossulus*, respectively. AA stands for the amino acid encoded by the tRNA

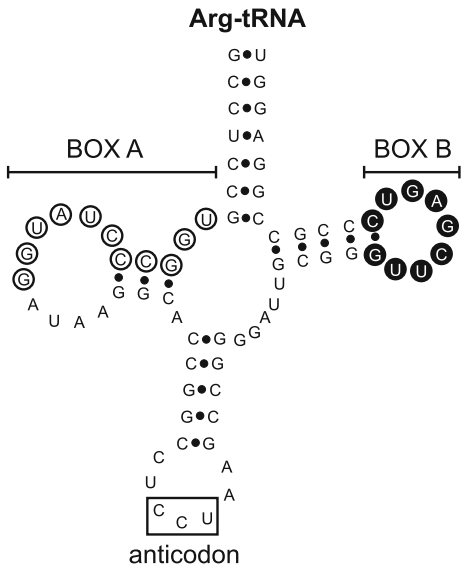

Fig. 2 Consensus secondary structure predicted for the tRNA-Arg linked to γ-5S rDNA sequences from *Mytilus*. The regions corresponding to the anticodon as well as the regions corresponding to boxes A and B involved in transcription by RNA polymerase III are indicated in the figure

According to PCR amplification patterns, *M. trossulus* and *M. californianus* contain multiple 5S rDNA units with different sizes, in agreement with previous reports for *M. galloprovincialis* and *M. edulis* (Insua et al. 2001). In contrast, only one amplification product was obtained in the case of *M. coruscus*. Such results seem to suggest the presence of a heterogeneous evolutionary process guiding 5S rDNA evolution in *Mytilus*, since duplications arose in some of the lineages but not in others. Three different 5S rDNA units were characterized in *M. trossulus* and *M. californianus* based on spacer regions. Two of these units described in *M. trossulus* showed high homology with α and β sequences described in *M. edulis* and *M. galloprovincialis*, based on spacer regions (Insua et al. 2001). Sequences encompassing homologies to β-5S rDNA were characterized in *M. californianus* as well, however, homologous sequences to α-5S rDNA were lacking in this species. Surprisingly, the putative α-5S rDNA detected in *M. californianus* turned out to be highly similar to the 3′ end of the β-5S spacer. Subsequently, this 5S type was referred to as small-β 5S rDNA. Given the absence of α-5S rDNA sequences in *M. californianus*, it is thus probable that the differentiation of this 5S rDNA type arose at the moment of the radiation leading to the formation of the *M. edulis* complex.

Table 3 Promoter elements of tRNA-Arg in *Mytilus*

	A-Box	B-Box	Reference
Consensus	TRGYnnAnnnG	GWTCRAnnC	Marck et al. (2006)
Mytilus	TGGCCC**W**ATGG	GTT**y**GAGTC	This study

Note: Differences respect to the consensus sequence described for yeast tRNA (Marck et al. 2006) are indicated in boldface

Furthermore, an additional 5S rDNA type was identified in *M. trossulus* and *M. californianus* (referred to as γ-5S rDNA), as well as in *M. edulis* and *M. galloprovincialis*, encompassing approximately 1000 bp in length and lacking homology neither to α nor β-5S rDNA types in the spacer region. Moreover, γ-5S rDNA sequences were found linked to tRNA-Arg genes oriented in opposite direction inside the spacer region. Although the presence of multiple bands in 5S rDNA amplification patterns was observed, in agreement with previous reports in *M. edulis* and *M. galloprovincialis* (Insua et al. 2001), our results represent the first description of additional 5S rDNA types (γ-5S rDNA and small-β rDNA) in these species.

Evolutionary Implications of the Linkage Between 5S rDNA and tRNA-Arg Genes

The analysis of γ-5S rDNA units unveiled the existence of linkage with tRNA-Arg genes (Fig. 2), representing the first report of such an arrangement in a higher eukaryote. Indeed, the only other report of such an association was described for the hemiascomycete *Yarrowia lipolytica* (Acker et al. 2008). Since most of the sequence information required for the expression and regulation of 5S rDNA

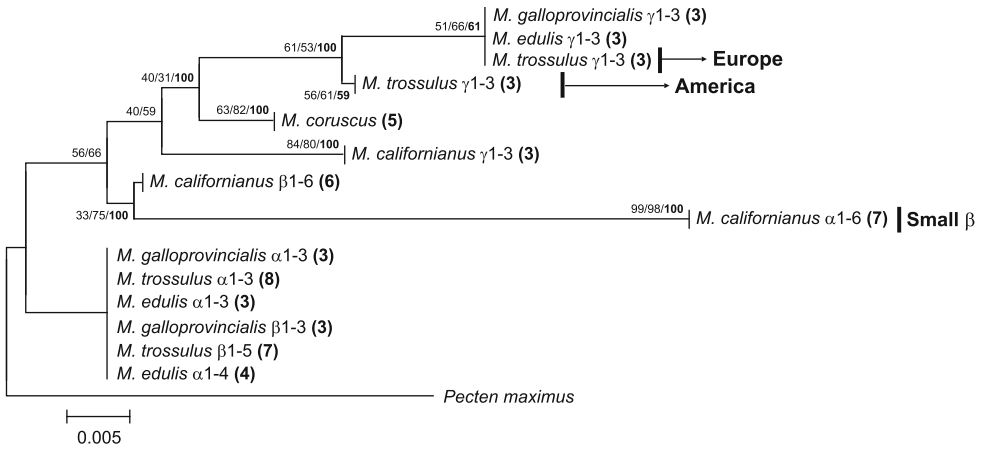

Fig. 3 Phylogenetic relationships among 5S rDNA coding regions in *Mytilus* species based on Kimura 2-parameter evolutionary distances. Numbers for interior nodes represent bootstrap and confidence probabilities based on 1000 replicates, followed by the BP corresponding to the maximum parsimony tree topology (only shown when greater than 50%). The topology was rooted with the 5S coding sequence from the bivalve mollusc *Pecten maximus* (GenBank accession AM691825). Coding sequences are referred to as α, β, and γ depending on the type of unit to which they belong, followed by the number of the clone. *Numbers* in *boldface* near species names indicate the number of sequences analyzed

resides within its coding region, a mechanism based on transposition through a RNA intermediate has been proposed as the responsible for its linkage to other genes (Drouin and Moniz de Sá 1995), being this also valid for the case of tRNA genes as depicted by the identification of promoter elements in their coding regions (Table 3). Both 5S rRNA and tRNA are transcribed by RNA polymerase III and their co-transcription was described in the case of *Y. lipolytica*. However, the notion of a cooperative transcription can be hardly reconciled with the fact that both 5S and tRNA genes are arranged in opposite directions in *Mytilus*. Different kinds of linkage have been reported for 5S rDNA units, including associations with genes encoding major rRNA, snRNAs and trans-splicer leader (Drouin and Moniz de Sá 1995; Manchado et al. 2006), as well as linkage with histone repetitive units in *M. galloprovincialis* (Eirín-López et al. 2004). The existence of two types of linkage in *Mytilus* (histones and tRNA) clearly emphasizes the intense dynamics of the 5S rDNA.

Molecular Variation Across 5S rDNA Types

The 5S coding regions described in the present work are characteristic by displaying a high degree of conservation among species and types of units as depicted by the phylogeny shown in Fig. 3. The coding regions of both α and β units showed similar sequences, in contrast to the case of γ-5S rDNA in which higher levels of variability were detected, including up to four nucleotide substitutions respect to the archetypal 5S rRNA sequence from the mussel *Mytilus edulis* (Fang et al. 1982). In functional terms, all sequences fit the canonical secondary structure predicted for 5S rRNA (Fig. 1) with the exception of the small-β type, which displays nucleotide replacements in the conserved internal control regions described for *Xenopus laevis* (Pieler et al. 1987) and *Drosophila melanogaster* 5S genes (Sharp and Garcia 1988). The presence of mutations in conserved functional elements could imply that small-β 5S rDNA sequences represent pseudogene copies, a notion further supported by the increased branch lengths in the phylogeny shown in Fig. 3. The presence of 5S rDNA pseudogenes has already been described in different species including humans (Nielsen et al. 1993), rats (Frederiksen et al. 1997), *Drosophila* (Sharp et al. 1984), fishes (Martins et al. 2002), and filamentous fungi (Rooney 2003), as a consequence of recurrent gene duplication events.

The phylogenetic analyses of *Mytilus* 5S spacer sequences depict *M. californianus* as the most divergent mussel species (excluding the controversial *M. coruscus*) in agreement with previous studies employing allozyme loci (Chichvarkhin et al. 2000) and satellite DNA (Martinez-Lage et al. 2005). On the other hand, the clustering of *M. galloprovincialis*, *M. edulis*, and *M. trossulus* 5S sequences in the phylogenies depends on the type of spacer analyzed. Thus, while a species-specific clustering pattern is observed in the independent analyses of α and γ-5S spacer sequences (with *M. galloprovincialis* and *M. edulis* being the closest species), analyses of β-5S spacer sequences reveal an intermingled clustering pattern defining two different subtypes (subtype 1 and subtype 2, Fig. 4b) as well as the presence of a highly divergent small-β type in *M. californianus*. The absence of subtype 2 β-5S spacer in *M. galloprovincialis* could be due to the later appearance of this species compared to *M. edulis* and *M. trossulus*, probably during the Pleistocene

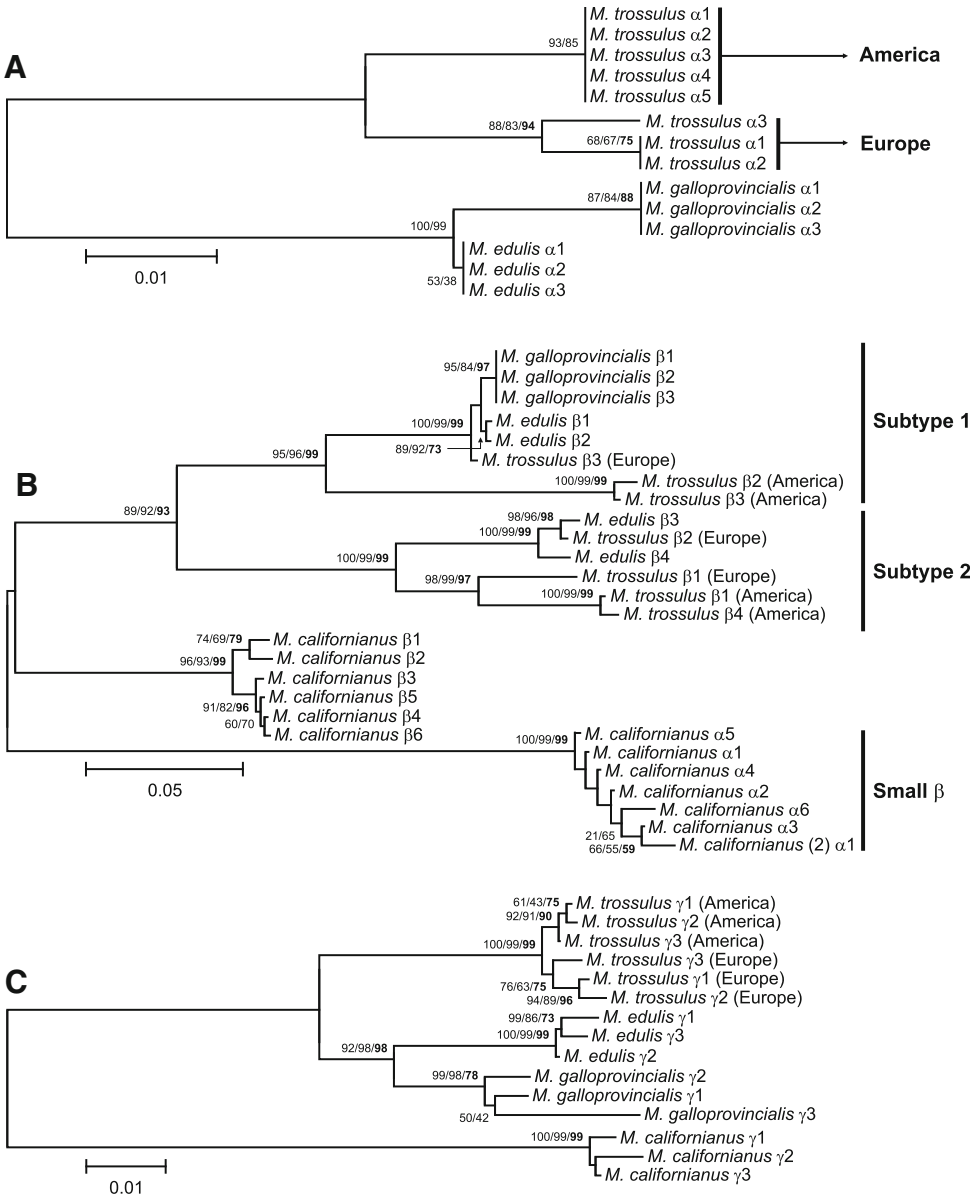

Fig. 4 Phylogenetic neighbor-joining trees of type α (**a**), type β (**b**), and type γ (**c**) 5S rDNA from *Mytilus* based on Kimura 2-parameter evolutionary distances. Types of units, clone numbering and BP and CP probabilities are indicated as in Fig. 3

Table 4 Polymorphism (π, bold values), divergence (K, lower diagonal), and fixed differences (upper diagonal) in 5S coding regions among *Mytilus* species

	M. galloprovincialis	*M. edulis*	*M. trossulus*	*M. californianus*	*M. coruscus*
M. galloprovincialis	**0.021 ± 0.006**	0	0	0	0
M. edulis	0.018 ± 0.008	**0.189 ± 0.004**	0	0	0
M. trossulus	0.018 ± 0.007	0.017 ± 0.006	**0.017 ± 0.004**	0	0
M. californianus	0.039 ± 0.012	0.038 ± 0.011	0.038 ± 0.010	**0.031 ± 0.003**	2
M. coruscus	0.021 ± 0.008	0.020 ± 0.007	0.022 ± 0.007	0.049 ± 0.016	**0.003 ± 0.002**

concomitantly with the formation of the Mediterranean Sea (Barsotti and Meluzzi 1968).

Quite on the contrary, α-5S spacer sequences are characteristic by displaying reduced levels of variation compared with β and γ-5S spacers. Such differences were further investigated by the study of the nucleotide substitution patterns leading to the different spacer sequences across *Mytilus* species by reconstructing the ancestral

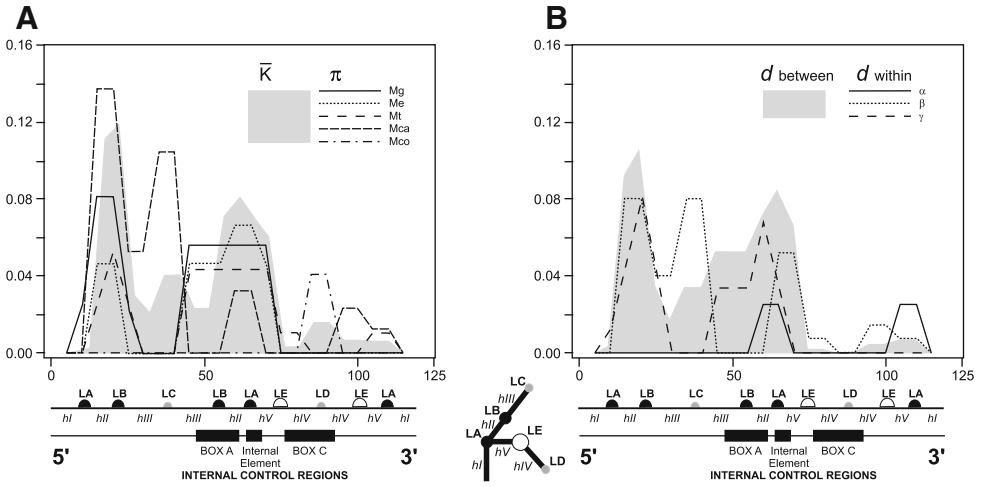

Fig. 5 Polymorphism (π) and average divergence (K) within and between *Mytilus* species, respectively (**a**) and nucleotide variation (d) within and between types of units (**b**) across the 5S coding region. The diversity values were estimated using a sliding-window approach with a window length of 10 bp and a step size of 5 bp. Loops A to E are indicated as LA–LE and helices I to V are indicated as *hI–hV*. Species are referred to as follows: *Mg M. galloprovincialis*, *Me M. edulis*, *Mt M. trossulus*, *Mca M. californianus*, *Mco M. coruscus*. The corresponding secondary rRNA structure (indicating the different loops and helices) as well as the internal control regions (indicated by *black boxes* on the 5S coding region represented by a *thin line*) are indicated below the graphs

Table 5 Evolutionary distances (Kimura 2-parameters) among *Mytilus* 5S coding regions based on different types of units. Distances within spacer types (bold values), between spacer types (lower diagonal), and fixed differences between spacer types (upper diagonal) are indicated

	Type α	*Type β*	*Type small-β*	*Type γ*
Type α	**0.002 ± 0.002**	0	0	1
Type β	0.004 ± 0.003	**0.005 ± 0.004**	0	0
Type small-β	0.058 ± 0.022	0.054 ± 0.021	**0.006 ± 0.005**	0
Type γ	0.035 ± 0.015	0.032 ± 0.014	0.074 ± 0.024	**0.017 ± 0.007**

Table 6 Polymorphism (π, bold values), divergence (K, lower diagonal), and fixed differences (upper diagonal) in α, β and γ 5S rDNA spacer regions from *Mytilus*

	M. galloprovincialis	*M. edulis*	*M. trossulus*	*M. californianus*
M. galloprovincialis	**(α) 0.052 ± 0.004**	(α) 2	(α) 10	(α) –
	(β) 0.000 ± 0.000	(β) 3	(β) 3	(β) 16
	(γ) 0.020 ± 0.007	(γ) 23	(γ) 36	(γ) 93
M. edulis	(α) 0.014 ± 0.010	**(α) 0.000 ± 0.000**	(α) 8	(α) –
	(β) 0.115 ± 0.093	**(β) 0.150 ± 0.038**	(β) 0	(β) 9
	(γ) 0.044 ± 0.017	**(γ) 0.006 ± 0.002**	45	(γ) 110
M. trossulus	(α) 0.093 ± 0.043	(α) 0.080 ± 0.037	**(α) 0.020 ± 0.005**	(α) –
	(β) 0.164 ± 0.073	(β) 0.160 ± 0.048	**(β) 0.184 ± 0.022**	(β) 8
	(γ) 0.062 ± 0.018	(γ) 0.063 ± 0.018	**(γ) 0.008 ± 0.001**	(γ) 107
M. californianus	(α) –	(α) –	(α) –	**(α) –**
	(β) 0.302 ± 0.105	(β) 0.318 ± 0.082	(β) 0.326 ± 0.068	**(β) 0.145 ± 0.013**
	(γ) 0.140 ± 0.054	(γ) 0.154 ± 0.060	0.149 ± 0.043	**(γ) 0.006 ± 0.002**

Note: Data for different spacer types are indicated near spacer type in parentheses

sequences for the internal nodes in the topologies shown in Fig. 4. Overall, a total of 71 and 93 nucleotide substitutions were involved in the differentiation of mussel β and γ-5S spacer sequences, respectively, quite in contrast with the case of α-5S spacer sequences which require only 4 nucleotide substitutions in order to accomplish the observed diversity in *Mytilus* species (Fig. 7). Under the assumption that 5S spacer regions are not subject to any

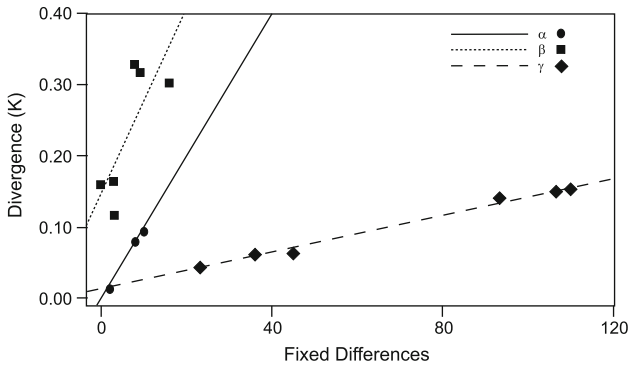

Fig. 6 Graphic representation of the nucleotide divergence against the numbers of fixed nucleotide differences among *Mytilus* species in 5S spacer regions

selective constraints, similar mutation rates would be expected for the different types of units. Far from this notion, our results strongly support a recent evolutionary origin for the α-5S rDNA type, based on its lesser degree of variation. Different hypothesis could account for this observation: (a) not enough time has elapsed since the origin of the α spacer in order to accumulate the same levels of variation presented by the diverse *β* and *γ*-5S rDNA types; and (b) *β* and *γ*-5S rDNA types encompass a higher level of specialization. However, and given the absence of α-5S spacer sequences in *M. californianus* and the unlikelihood of selection acting differentially across 5S spacer sequences, our results seem to strengthen the hypothesis pointing towards a recent evolutionary origin of the α-5S rDNA type.

The presence of identity between α and *β*-5S spacers at 3′ terminal regions, previously reported for molluscs (Insua et al. 2001), has been also demonstrated in the present work, finding homologies between *M. trossulus* α and *β*-5S spacer sequences. However, the greatest level of identity between spacer types was found in the case of *M. californianus* which displays a group of α-5S spacers (referred to as small-*β*) more closely related to *β*-5S spacers. Finally, none of the three types of 5S rDNA units described in the present study displayed sequence similarity in the spacer region respect to *M. coruscus*. The phylogenetic and taxonomic status of this species is complex and its relationship with the remainder of the *Mytilus* species uncertain (Chichvarkhin et al. 2000; Gosling 1992; Martinez-Lage et al. 2005). Although the high temperatures used in PCR amplifications could represent a bias in detecting additional 5S rDNA types in *M. coruscus*, the validity of the primers used in the present study is supported by the successful amplification of 5S rDNAs in other mollusc species above the Family level, including cockles (Freire et al. 2005; Insua et al. 1999) and scallops (Lopez-Piñon et al. 2008), revealing the presence of two 5S rDNA types in some of the species (i.e., *Cerastoderma glaucum*). In our opinion, the aforementioned experimental results (together with the conservation of 5S rDNA sequences in *Mytilus*) rule out the possibility of losing amplification accuracy resulting from increased astringency due to high temperatures. Thus, the existence of a single 5S rDNA type lacking similarity with any of the other types of 5S rDNA identified in the present work strongly support the divergent nature of *M. coruscus*.

Long-Term Evolution of 5S rDNA Sequences in *Mytilus*

Although concerted evolution has been recently discarded (in favor of a birth-and-death mechanism) as the major model guiding the long-term evolution of several multigene families such as for instance histones (Eirín-López et al. 2009; Nei and Rooney 2006), the case of ribosomal DNA seems to be otherwise more complex, probably involving a combined effect of concerted evolution and birth-and-death evolution (Fujiwara et al. 2009). The results presented in this work unveil an apparent absence of

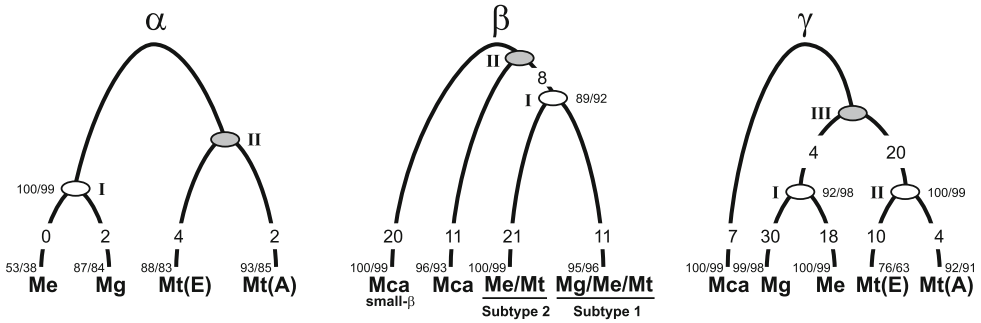

Fig. 7 Evolutionary pathways leading to the differentiation of 5S spacer sequences in mussels (species are referred to as in Fig. 5). The numbers of nucleotide changes from ancestral sequences reconstructed for nodes I–III are indicated in the corresponding internal branches. Confidence levels for the internal nodes are indicated as in Fig. 3. Ancestral sequences were reconstructed by maximum likelihood using the K80/TPM3uf model (Kimura 1980) for α-5S spacers, the HKY + G model (Hasegawa et al. 1985) for *β*-5S spacers, and the HKY + I model (Hasegawa et al. 1985) in the case of *γ*-5S spacers (see "Materials and Methods" section)

a clear interspecies differentiation across 5S coding regions, a notion reinforced by: (a) the lack of fixed differences between species (Table 4), as indicated by the strong correlation observed between the polymorphism within species and the divergence between species (Fig. 5); and (b) the low levels of nucleotide variation within 5S coding regions in comparisons between different types of units (Table 5). An extensive lack of homology was also detected among spacer sequences belonging to different rDNA units (α, β, and γ), suggesting the presence of independent evolutionary pathways leading to their differentiation. Oppositely to the case of coding regions, comparisons among spacers from different units (Table 6) reveal the presence of close intraspecies relationships encompassing high numbers of fixed differences among different *Mytilus* species. Both observations are illustrated by the comparison between the nucleotide divergence and the numbers of fixed differences among *Mytilus* species (Fig. 7), revealing a significant positive correlation in the three types of spacers. In addition, a total absence of stretches of homogenized sequences by gene conversion events in spacer sequences belonging to each of the three types of units was confirmed by using the GENECONV program (Sawyer 1999). Overall these results seem to indicate that a process involving an extensive genetic homogenization within species does not represent the major mechanism guiding the long-term evolution of 5S rDNA units.

Although the results previously discussed do not fit the predictions made by the concerted evolution model, they can be still reconciled with a critical role for this evolutionary model in 5S rDNA evolution. Different studies have put forward a hypothesis in which the homogenization of rDNA units would occur locally within arrays, implying that selective mechanisms operate in the coding region eliminating mutations without affecting spacer regions (Cronn et al. 1996; Kellogg and Appels 1995). In this scenario, concerted evolution would eliminate variants or spread them throughout individual arrays, but not between different arrays. These observations are specially relevant in the case of *Mytilus*, in which 5S rDNA units might be located at physically different chromosomal loci, such as for the case of *M. edulis* and *M. galloprovincialis* (Insua et al. 2001). Indeed, fluorescent in situ hybridization (FISH) experiments using specific probes for α and β spacers reveal that they occupy different chromosomal locations in *M. galloprovincialis* (manuscript in preparation).

It is thus possible that a first stage of 5S gene evolution would had involved the generation of genetic diversity through recurrent gene duplications (birth-and-death) in the genome of these organisms, followed by the transposition of several units to different chromosomal locations leading to the their subsequent independent evolution. Indeed, a recent study has suggested that the increase in RNA gene copy number in yeast most likely represents an evolutionary strategy focused towards the reduction of rDNA transcriptional activity, which interferes with cohesion between rDNA loci of sister chromatids (Ide et al. 2010). However, while the experimental basis for this study is focused on the disruption of RNA polymerase I (transcribing 35S rRNA genes in yeast), further studies will be needed in order to extend this conclusion to the case of 5S rRNA genes (transcribed by RNA polymerase III). Given that 5S rRNA genes are represented in multiple copies in the genome, variations from the consensus sequence (including pseudogenes), are expected to have little effect in the fitness of an organism. At this point, the local effect of concerted evolution could represent the second step in 5S rRNA evolution, following the genetic diversification generated by birth-and-death.

Homogenization would be relevant at a local level within arrays at different loci, allowing for the accumulation of neutral or nearly neutral mutations in spacers, and eliminating those arrays in which mutations in coding regions have reached a threshold that results in a decrease of the organism's fitness (Cronn et al. 1996; Kellogg and Appels 1995). However, although this hypothesis is consistent with observations made on *Drosophila* (Samson and Wegnez 1988) and *Saccharomyces* (Szostak and Wu 1980), it is important to bear in mind that a recent and rapid process of gene duplication followed by selection could also result in a pattern of homogeneity (without invoking concerted evolution) similar to that observed in the present work. This latter scenario raises an important question as it pertains to whether the presence of heterogeneous functional constraints associated with different rDNA units are in fact responsible for the differentiation of their spacer regions. In this sense, previous reports reveal a lack of significant nucleotide differences among rDNA types in regulatory segments upstream spacer regions (Insua et al. 2001), suggesting that these regions do not represent primary targets for selection. On the other hand, it has been suggested that the 5S array as a whole (i.e., numbers of functional copies) could be selectively constrained (Kellogg and Appels 1995).

Concluding Remarks

Birth-and-death has been proposed as a very important mechanism in guiding the long-term evolution of the 5S rDNA family in different organisms (Fujiwara et al. 2009; Rooney and Ward 2005). The results in the present work reveal that *Mytilus* 5S rDNA genes do not represent an exception to this rule. Our results suggest that the

long-term evolution of mussel 5S rDNA is most likely mediated by a mixed mechanism in which the generation of genetic diversity is achieved through birth-and-death (recurrent gene duplication) followed by the local homogenization of the different units through concerted evolution (probably after their physical transposition to independent chromosomal locations). However, even though the observed patterns of 5S rDNA evolution could also result from a process of gene duplication and selection without invoking homogenization, a substantial effect of concerted evolution could not be ruled out until the presence of heterogeneous selective constraints acting on different 5S types is demonstrated.

Acknowledgments This work was supported by a grant within the Ramon y Cajal Subprogramme (MICINN, Spanish Government) awarded to JME-L. RF was supported by an Isabel Barreto contract from the Xunta de Galicia (SPAIN).

References

Acker J, Ozanne C, Kachouri-Lafond R, Gaillardin C, Neuveglise C, Marck C (2008) Dicistronic tRNA-5S rRNA genes in *Yarrowia lipolytica*: an alternative TFIIIA-independent way for expression of 5S rRNA genes. Nucl Acids Res 36:5832–5844

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Alves-Costa FA, Wasko AP, Oliveira C, Foresti F, Martins C (2006) Genomic organization and evolution of the 5S ribosomal DNA in Tilapiini fishes. Genetica 127:243–252

Arnheim N (1983) Concerted evolution of multigene families. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer Associates, Sunderland, MA, pp 38–61

Barsotti G, Meluzzi C (1968) Osservazioni su *Mytilus edulis* L. e *Mytilus galloprovincialis* Lamarck. Conchiglie 4:50–58

Baum BR, Edwards T, Johnson DA (2009) Phylogenetic relationships among diploid *Aegilops* species inferred from 5S rDNA units. Mol Phylogenet Evol 53:34–44

Birnboim HC, Doly J (1979) A rapid alkalyne procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Chichvarkhin AY, Kartavtsev YF, Kafanov AI (2000) Genetic relationships among some species of mytilidae (mollusca: Bivalvia) from the northern pacific ocean. Genetika 36:1206–1220

Cronn RC, Zhao X, Patterson AH, Wendel JF (1996) Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. J Mol Evol 42:685–705

Cross I, Rebordinos L (2005) 5S rDNA and U2 snRNA are linked in the genome of *Crassostrea angulata* and *Crassostrea gigas* oysters: does the (CT)n(GA)n microsatellite stabilize this novel linkage of large tandem arrays? Genome 48:1116–1119

Dover G (1982) Molecular drive: a cohesive mode of species evolution. Nature 299:111–117

Drouin G, Moniz de Sá MM (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 JM, Ruiz MF, González-Tizón AM, Martínez A, Sánchez L, 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

Eirín-López JM, González-Romero R, Dryhurst D, Méndez J, Ausió J (2009) Long-term evolution of histone families: old notions and new insights into their diversification mechanisms across eukaryotes. In: Pontarotti P (ed) Evolutionary biology: concept, modeling, and application. Springer-Verlag, Berlin Heidelberg, pp 139–162

Fang BL, De Baere R, Vandenberghe A, De Watcher R (1982) Sequences of three molluscan 5S ribosomal RNAs confirm the validity of a dynamic secondary structure model. Nucl Acids Res 10:4679–4685

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Fernandez-Tajes J, Mendez J (2009) Two different size classes of 5S rDNA units coexisting in the same tandem array in the razor clam *Ensis macha*: is this region suitable for phylogeographic studies? Biochem Genet 47:775–788

Frederiksen S, Cao H, Levan G, Hallenberg C (1997) The rat 5S rRNA bona fide gene repeat maps to chromosome19q12-qter and the pseudogene repeat maps to 12q12. Cytogenet Cell Genet 76:101–106

Freire R, Insua A, Méndez J (2005) *Cerastoderma glaucum* 5S ribosomal DNA: characterization of the repeat unit, divergence with respect to *Cerastoderma edule*, and PCR-RFLPs for the identification of both cockles. Genome 48:427–442

Fujiwara M, Inafuku J, Takeda A, Watanabe A, Fujiwara A, Kohno S, Kubota S (2009) Molecular organization of 5S rDNA in bitterlings (Cyprinidae). Genetica 135:355–365

Gerbi SA (1985) Evolution of ribosomal RNA. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum Press, New York, pp 419–517

Gosling EM (1992) Genetics of *Mytilus*. In: Gosling EM (ed) The mussel *Mytilus*: ecology, physiology, genetics and culture. Elsevier, Amsterdam, pp 309–382

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98

Hasegawa M, Kishino K, Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Ide S, Miyazaki T, Maki H, Kobayashi T (2010) Abundance of ribosomal RNA gene copies maintains genome integrity. Science 327:693–696

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

Kellogg EA, Appels R (1995) Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics 140:325–343

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

Komiya H, Hasegawa M, Takemura S (1986) Differentiation of oocyte- and somatic-type 5S rRNAs in animals. J Biochem (Tokyo) 100:369–374

Lazar E, Haendler B, Jacob M (1983) Two 5S genes are expressed in chicken somatic cells. Nucleic Acids Res 11:7735–7741

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

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

Long EO, Dawid IB (1980) Repeated genes in eukaryotes. Annu Rev Biochem 49:727–764

Lopez-Piñon MJ, Freire R, Insua A, Mendez J (2008) Sequence characterization and phylogenetic analysis of the 5S ribosomal DNA in some scallops (Bivalvia: Pectinidae). Hereditas 145:9–19

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 25:955–964

Manchado M, Zuasti E, Cross I, Merlo A, Infante C, Rebordinos L (2006) Molecular characterization and chromosomal mapping of the 5S rRNA gene in *Solea senegalensis*: a new linkage to the U1, U2, and U5 small nuclear RNA genes. Genome 49:79–86

Marck C, Kachouri-Lafond R, Lafontaine I, Westhof E, Dujon B, Grosjean H (2006) The RNA polymerase III-dependent family of genes in hemiascomycetes: comparative RNomics, decoding strategies, transcription and evolutionary implications. Nucl Acids Res 34:1816–1835

Martinez-Lage A, Rodriguez-Farina F, Gonzalez-Tizon A, Mendez J (2005) Origin and evolution of *Mytilus* mussel satellite DNAs. Genome 48:247–256

Martins C, Wasko AP, Oliveira C, Porto-Foresti F, Parise-Maltempi PP, Wright JM, Foresti F (2002) Dynamics of 5S rDNA in the tilapia (*Oreochromis niloticus*) genome: repeat units, inverted sequences, pseudogenes and chromosome loci. Cytogenet Genome Res 98:78–85

Nei M, Rooney AP (2006) Concerted and birth-and-death evolution in multigene families. Annu Rev Genet 39:121–152

Nicholas KB, Nicholas HB Jr, Deerfield DW II (1997) Gene Doc: Analysis and visualization of genetic variation. EMBNEW NEWS 4:14

Nielsen JN, Hallenberg C, Frederiksen S, Sorensen PD, Lomolt B (1993) Transcription of human 5S rRNA genes is influenced by an upstream DNA sequence. Nucl Acids Res 21:3631–3636

Paule MR, White RJ (2000) Transcription by RNA polymerases I and III. Nucl Acids Res 28:1283–1298

Pieler T, J. H, Roeder RG (1987) The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. Cell 48:91–100

Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256

Rooney AP (2003) Selection for highly biased amino acid frequency in the TolA cell envelope protein of proteobacteria. J Mol Evol 57:731–736

Rooney AP, Ward TJ (2005) Evolution of a large ribosomal RNA multigene family in filamentous fungi: birth and death of a concerted evolution paradigm. Proc Natl Acad Sci USA 102:5084–5089

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

Samson M-L, Wegnez M (1988) Bipartite structure of the 5S ribosomal gene family in a *Drosophila melanogaster* strain, and its evolutionary implications. Genetics 118:685–691

Sawyer SA (1999) GENECONV: a computer package for the statistical detection of gene conversion. Department of Mathematics, Washington University, St. Louis, MO

Sharp S, Garcia A (1988) Transcription of the *Drosophila melanogaster* 5S RNA gene requires an upstream promoter and four intragenic sequence elements. Mol Cell Biol 8:1266–1274

Sharp S, Garcia A, Cooley L, Saavedra C, Soll D (1984) Transcriptionally active and inactive gene repeats within the *D. melanogaster* 5S RNA gene cluster. Nucl Acids Res 12:7617–7632

Sitnikova T (1996) Bootstrap method of interior-branch test for phylogenetic trees. Mol Biol Evol 13:605–611

Sitnikova T, Rzhetsky A, Nei M (1995) Interior-branch and bootstrap tests of phylogenetic trees. Mol Biol Evol 12:319–333

Szostak JW, Wu R (1980) Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. Nature 284:426–430

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 25:4876–4882

Vierna J, Gonzalez-Tizon A, Martinez-Lage A (2009) Long-term evolution of 5S ribosomal DNA seems to be driven by birth-and-death processes and selection in *Ensis* razor shells (mollusca: Bivalvia). Biochem Genet 47:635–644

Winnepenninckx B, Backeljau T, De Watcher R (1993) Extraction of high molecular weight DNA from molluscs. Trends Genet 9:407

Yang Z (2007) PAML 4: a program package for phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591