Evidence for 5S rDNA Horizontal Transfer in the toadfish *Halobatrachus didactylus* (Schneider, 1801) based on the analysis of three multigene families

*Merlo et al.*
Evidence for 5S rDNA Horizontal Transfer in the toadfish *Halobatrachus didactylus* (Schneider, 1801) based on the analysis of three multigene families

Manuel A Merlo1, Ismael Cross1, José L Palazón2,3, María Úbeda-Manzanaro1,2, Carmen Sarasquete2 and Laureana Rebordinos1*

**Abstract**

**Background:** The Batrachoididae family is a group of marine teleosts that includes several species with more complicated physiological characteristics, such as their excretory, reproductive, cardiovascular and respiratory systems. Previous studies of the 5S rDNA gene family carried out in four species from the Western Atlantic showed two types of this gene in two species but only one in the other two, under processes of concerted evolution and birth-and-death evolution with purifying selection. Here we present results of the 5S rDNA and another two gene families in *Halobatrachus didactylus*, an Eastern Atlantic species, and draw evolutionary inferences regarding the gene families. In addition we have also mapped the genes on the chromosomes by two-colour fluorescence in situ hybridization (FISH).

**Results:** Two types of 5S rDNA were observed, named type α and type β. Molecular analysis of the 5S rDNA indicates that *H. didactylus* does not share the non-transcribed spacer (NTS) sequences with four other species of the family; therefore, it must have evolved in isolation. Amplification with the type β specific primers amplified a specific band in 9 specimens of *H. didactylus* and two of *Sparus aurata*. Both types showed regulatory regions and a secondary structure which mark them as functional genes. However, the U2 snRNA gene and the ITS-1 sequence showed one electrophoretic band and with one type of sequence. The U2 snRNA sequence was the most variable of the three multigene families studied. Results from two-colour FISH showed no co-localization of the gene coding from three multigene families and provided the first map of the chromosomes of the species.

**Conclusions:** A highly significant finding was observed in the analysis of the 5S rDNA, since two such distant species as *H. didactylus* and *Sparus aurata* share a 5S rDNA type. This 5S rDNA type has been detected in other species belonging to the Batrachoidiformes and Perciformes orders, but not in the Pleuronectiformes and Clupeiformes orders. Two hypotheses have been outlined: one is the possible vertical permanence of the shared type in some fish lineages, and the other is the possibility of a horizontal transference event between ancient species of the Perciformes and Batrachoidiformes orders. This finding opens a new perspective in fish evolution and in the knowledge of the dynamism of the 5S rDNA. Cytogenetic analysis allowed some evolutionary trends to be roughed out, such as the progressive change in the U2 snDNA and the organization of (GATA)n repeats, from dispersed to localized in one locus. The accumulation of (GATA)n repeats in one chromosome pair could be implicated in the evolution of a pair of proto-sex chromosomes. This possibility could situate *H. didactylus* as the most highly evolved of the Batrachoididae family in terms of sex chromosome biology.

* Correspondence: laureana.rebordinos@uca.es
1Laboratorio Genética, Facultad de Ciencias del Mar y Ambientales, CACYT MAR, Universidad de Cádiz, Puerto Real (Cádiz) 11510, Spain
2Full list of author information is available at the end of the article

© 2012 Merlo et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Background

The Lusitanian toadfish (*Halobatrachus didactylus*), a member of the Batrachoididae family, is found in natural habitats of soft sand and rocky sea beds, and reefs. The Batrachoididae family comprises 71 species, belonging to three subfamilies [1] that are widely distributed across the Atlantic, Pacific, and Indian Oceans. *H. didactylus* is the only species of this family found in the Iberian Peninsula [2] and populations are distributed along the coasts from the Bay of Biscay (Spain) to Ghana as well as the western Mediterranean [3].

On the coasts of Portugal and southwestern Spain, this species is of moderate commercial importance. However, the real importance of this species is for research purposes rather than its commercial value. It has traditionally been used as a model animal in toxicology experiments [4], and in hematology, reproduction and histophysiology studies [5-7]. Genetic studies in the Batrachoididae family have been focused mainly on cytogenetics (reviewed by [8]); there is only one published paper in which this fish family has been studied from a molecular perspective [9]. For *H. didactylus* in particular, the karyotype and the localization of various repetitive sequences have been described [10].

Multigene families are comprised of multiple genes that all descend from a common ancestral gene, have similar sequences and are functionally related [11]. Traditionally, it has been presumed that the evolution of multigene families is in accordance with a model known as the “Concerted Evolution Model,” in which the members of the family are homogenized by several different mechanisms such as unequal crossing-over and gene conversion [12]. This feature makes the units show more similarity within than between related species [13]. In the last two decades, a new model has been proposed to explain controversial situations found in some multigene families: the “Birth-and-Death Evolution Model”. Under this model new genes arise by duplications during the evolution of a group of organisms, and these new genes are either fixed as functional genes or become pseudogenes [14]. The histone and the immunoglobulin gene families are typical examples of birth-and-death evolution [11].

The ribosomal genes are divided in two clusters: the major (45S rDNA) and the minor (5S rDNA) clusters. The 45S rDNA has a transcribed unit which comprises two external transcribed spacers (5′ ETS and 3′ ETS), the coding for 18S, 5.8S and 28S rRNAs, and two internal transcribed spacers, which separate the 18S rRNA from the 5.8S rRNA (ITS-1) and the 5.8S rRNA from the 28S rRNA (ITS-2); the transcribed units are separated by intergenic spacers (IGS). Meanwhile, the 5S rDNA comprises a conserved coding region of 120 bp and a non-transcribed spacer (NTS) which is variable among species in length and sequence. The two ribosomal clusters represent well-known examples of concerted evolution.

The small nuclear RNAs (snRNA) are components of the small nuclear ribonucleoprotein particles (snRNP), a complex which participate in the splicing process of the mRNA precursors [12]. Most of the snRNA are transcribed by RNA polymerase II, except for U6 snRNA which is transcribed by RNA polymerase III [15]. Within the group of fishes, there have been very few studies made of the snRNA gene family at molecular and cytogenetic level. It has been presumed that the snRNA gene families undergo a concerted evolution [11].

In this work a molecular characterization has been made of three multigene families: 5S rDNA, ITS-1 (from 45S rDNA) and the U2 snRNA gene of *H. didactylus*. In addition, double-FISH technique was used to ascertain the U2 snRNA gene localization and the possible co-localization of this probe with the 5S rDNA, 18S rRNA gene and (GATA)n probes. The co-localization of (GATA)n and the 18S rRNA gene was also tested. This technique not only provides various different cytogenetic markers, but also represents a starting point for the production of a genetic map in *H. didactylus*. The nucleotide variation has been estimated to assess the homogenizing forces which support the concerted evolution in the three multigene families. Finally, the results have been compared with those obtained from other members of the Batrachoididae family, and the evolutionary difference between those and *H. didactylus* has been established.

Results

Sequence analysis

After PCR amplification, two electrophoretic bands of 5S rDNA sequence could be observed, named type α and type β, and whose sizes were 420 and 200 bp respectively. The amplification with the type β specific primer gave a 190 bp band in 9 specimens of *H. didactylus* (data not shown). The sequencing results demonstrate that the 5S rDNA sizes varied from 413 to 424 bp in type α, and from 197 to 207 bp in type β 5S rDNA sequences. In the former the variability was due a variable poly-A region at position ~71, and in the latter by a 10 bp insertion-deletion (indel) localized immediately after the poly-T termination signal. In two clones an electrophoretic band of 840 bp was obtained; this corresponded to a dimeric form of 5S rDNA type α. From these findings, it would therefore be possible to ascertain the complete sequence for 5S rDNA of type α.

Both types of 5S rDNA sequence conserved all promoter and regulatory regions necessary for transcription by RNA polymerase III (pol III), such as the TATA-like promoter element around the position ~30 upstream
from the transcription starting point, the poly-T terminator region situated at the 3’ end of the coding region and the Internal Control Regions (Box A, Intermediate Element and Box C) located inside coding regions between the positions +50 to +65, +68 to +73, and +81 to +98 respectively. The RNAstructure 5.2 program showed a stable secondary structure of 5S rRNA in both cases (type α and type β) (Figure 1), with similar free energy values, and were in accordance with secondary structures predicted in previously described models [16].

Meanwhile, the ITS-1 and U2 snDNA sequences showed one electrophoretic band of 650 and 800 bp respectively. The sequencing results demonstrated that the ITS-1 sequence presented an exact size of 664 bp, and 820 bp in the U2 snDNA sequence. Furthermore, three putative promoter and regulator sequences in the U2 snRNA gene have been found (Figure 2). These include the Proximal Sequence Element (PSE), the Distal Sequence Element (DSE) and the 3’ box within the spacer region. The PSE and 3’ box were located by alignment with consensus sequences obtained from the Moronidae species, *Dicentrarchus labrax* and *Dicentrarchus punctatus*. From the alignment of 126 bp upstream from the coding region two conserved regions were detected (Figure 2a). The more proximal of these is between position −60 and −23, while the more distal region comprises a 17 bp box, the sequence of which has not previously been described as a promoter or regulator region. The 100 bp downstream from the coding region was also aligned and a more conserved box was found at 8 nucleotides downstream from the 3’ end of the coding region (Figure 2b). A putative DSE octameric sequence was localized at −311 bp from the transcription starting point (Figure 2d), and it was located in direct orientation with respect to the consensus octameric sequence [17].

**Variability analysis**

The three multigene families studied showed low values of nucleotide variability, in both coding regions and spacers. As expected, the spacer regions had higher values than coding regions, except in the U2 snRNA gene (Table 1). On the whole, the most variable multigene family was the U2 snDNA. The 5S rDNA type α also showed high nucleotide variability. The less variable genes were the 5S rDNA type β and the 18S rRNA-ITS-1-5.8S rRNA sequence. Furthermore, divergence values between the two types of 5S rDNA were different depending on the fragment considered in the study. Thus, the coding region showed a low value of divergence ($d = 0.087 \pm 0.028$), whereas divergence in the NTS was higher ($d = 0.751 \pm 0.218$).

The gene families under study were subjected to BLASTN search at the NCBI database; all coding regions (5S rRNA, 18S rRNA, 5.8S rRNA and U2 snRNA) showed high degrees of identity with others obtained from a wide variety of fish species. Spacer regions did not present any significant homologies, except for the 5S rRNA gene spacer (NTS) type β which, surprisingly, presented a high degree of homology with the NTS from *Sparus aurata* described by [18] (Figure 3). The analysis with the type β demonstrated that this type was present in all specimens of *H. didactylus* studied, as well as in all the species included for this experiment (see Methods section), except in *Solea senegalensis* and *Engraulis encrasicholus* (Figure 4). However, these two species showed positive bands in the parallel experiment with ITS-1 primers.

**Phylogenetic analysis**

The Maximum Likelihood (ML) tree obtained (Figure 5) was compared with the Neighbor-Joining (NJ) tree previously described by [9] for other species of the
Batrachoididae family (Western batrachoids). As in the NJ tree, the ML tree divided the Western batrachoids species in two clades, characterized by different NTS types. In the clade 2 (following the nomenclature devised by [9]) of these species, the types $\alpha$ and $\beta$ of $H$. didactylus were grouped on the same branch, but within that branch, the type $\beta$ clade formed a monophyletic group with the clade 2 of the Western batrachoids. The two sequences of $S$. aurata mentioned earlier were clustered within the $H$. didactylus type $\beta$ clade with a high branch support. Therefore, $S$. aurata did not group with the two other sparid species, which formed a separate cluster from that of the Batrachoididae. Within clade 1 of the Western batrachoids, the clones of the species $Ampichthys cryptocentrus$ were clustered in the same group, whereas in the NJ tree previously described by [9], these clones were divided into two groups.

Cytogenetic analysis

Four double-FISH experiments were carried out with the following: the U2 snRNA gene with 5S rDNA (Figure 6a); the 18S rRNA gene with U2 snRNA gene (Figure 6b); the (GATA)$_n$ sequence with U2 snRNA gene (Figure 6c); and the (GATA)$_n$ sequence with 18S rRNA gene (Figure 6d). The 5S rDNA probe only includes the type $\alpha$. The double-FISH treatments with 5S-18S and 5S-GATA probes had been performed in previous work [10]. The U2 snDNA probe was localized in the middle of the long arm of a medium-size submeta/subtelocentric chromosome pair. No co-localization between the U2 snRNA gene probe and the remaining probes was detected. Double-FISH treatment with (GATA)$_n$, and 18S rRNA gene probes showed a non co-localized pattern.

Discussion

Two types of 5S rDNA have been found in $H$. didactylus, named here as type $\alpha$ and type $\beta$. The NTS sequences between the two types do not share significant homologies. Furthermore, the degree of divergence of the coding region between the two types is more than 5 times higher than the nucleotide variability within the coding region of type $\alpha$, and more than 28 times higher than such variability in type.

### Table 1 Polymorphism by gene region

| Region     | Length (bp) | %GC | S  | h  | $\pi$       |
|------------|-------------|-----|----|----|-------------|
| 5S rRNA ($\alpha$) | 120         | 52.50-55.83 | 12  | 10  | 0.018±0.003 |
| NTS ($\alpha$)    | 289-304     | 36.75-39.79 | 41  | 14  | 0.02±0.006  |
| 5S rRNA ($\beta$) | 120         | 56.67-57.50 | 3   | 4   | 0.004±0.002 |
| NTS ($\beta$)      | 77-87       | 38.96-43.02 | 4   | 5   | 0.006±0.003 |
| 18S rRNA           | 182         | 48.90-49.45 | 3   | 4   | 0.004±0.001 |
| ITS-1               | 388         | 62.89-63.40 | 9   | 5   | 0.008±0.002 |
| S.85 rRNA           | 93-94       | 50.54-51.06 | 1   | 2   | 0.002±0.002 |
| U2 snRNA            | 189         | 35.45-38.10 | 12  | 7   | 0.024±0.003 |
| U2 spacer           | 623-631     | 36.61-37.88 | 31  | 8   | 0.017±0.003 |

%GC: GC content; S: number of polymorphic sites; h: number of haplotypes; $\pi$: nucleotide variability.
β (Table 1). Both types presented all the regulatory elements which make them functional genes. The existence of two types of 5S rDNA appears to be common in fish species [9,19,20], and this situation has been commonly explained as a dual expression system, where one type is expressed in both somatic and oocyte cells and the other type only in oocyte cells [21].

In a previous study [9] with four species of the Batrachoididae family from the Western Atlantic, it has been found that two species also shared two types of 5S rDNA, and the two other species have only one type. Therefore, it cannot be concluded that a dual system of 5S rDNA is generally established in the Batrachoididae family. In the four Western Atlantic species the NTS sequences are each similar to the others (see Figure 3 of [9]), whereas in H. didactylus these sequences are distantly related to those of the four species. This feature could indicate that H. didactylus has evolved independently from the Western Atlantic species, due to the geographic isolation between them, which could have contributed to the greater divergence found between Eastern and Western Atlantic species.

The sequences of both coding regions and the NTSs are quite conserved in H. didactylus, and the secondary structure agrees with previous described models [16] (Figure 1). Therefore, concerted evolution is the more feasible model which is acting in each locus of this multigene family. Under this model, the different units of the gene family tend to homogenize their sequences by means of mechanisms such as unequal cross-over and gene conversion [12], thus spreading mutation events along the multigene family. Concerted evolution makes the sequence more similar within the species than it is among related species [13]. Traditionally it has been established that the 5S rDNA gene family is a paradigm for the concerted evolution model; however, although

Figure 3 5S rDNA type β, common to Sparus aurata and Halobatrachus didactylus. Alignment of the two clones of Sparus aurata (GenBank Acc. No. AY330701.1 and AY330702.1) and the type β clones of Halobatrachus didactylus. A schematic representation of 5S rDNA is included, highlighting the A box, the Intermediate Element (IE), the C box, the poly-T terminator region, and the TATA-like box.

Figure 4 Agarose gel electrophoresis of 5S rDNA type β species-specific PCR. Positive products of 190 bp are marked by an arrow. Hdi: Halobatrachus didactylus; Ppl: Ponerichthys plecotodon; Tma: Thalassophryne maculosa; Acr: Amphiichthys cryptocentrus; Bma: Batrachoides manglae; Sau: Sparus aurata (positive control); Ppa: Pagonius pagonius; Dsa: Diplodus sargus; Dia: Dicentrarchus labrax; Are: Argyrosomus regius; Pme: Plectorhinchus mediterraneus; Sse: Solea senegalensis; Een: Engraulis encrasicolus; C+: negative control.
this is the most common model observed for this gene, in more and more studies that are published, it is concluded that the 5S rDNA gene family can follow a different evolution model. Some authors have described 5S rDNA gene families which fit better with a birth-and-death evolution model than with the concerted model [9,14,22], or with a mixed model with elements of both birth-and-death and concerted evolution [23]. In the *H. didactylus* case, although the concerted model is acting at locus level, at genome level, the presence of two types and the between-species clustering of the multigene family indicate a birth-and-death evolution model.

As mentioned above, surprisingly the 5S rDNA type β presented homology in the NTS sequence with those obtained in *S. aurata* [18] (Figure 3). This species and *H. didactylus* belong to different superorders (Acanthopterygii and Paracanthopterygii respectively), which diverged 55 My ago [24]. Both species could have inherited the type β from an ancestral species common to the two superorders, and could have maintained in some lineages and lost in others along the fish evolution. Using the type β specific primer, we were able to conclude that this sequence is found in all specimens studied of *H. didactylus*, as well as in *S. aurata* and the remaining species belonging to Batrachoideiformes (*Porichthys plectrodon*, *Thalassophryne maculosa*, *Amphichthys cryptocentrus* and *Batrachoides manglae*) and Perciformes order (*Pagrus pagrus*, *Diplodus sargus*, *Dicentrarchus labrax*, *Argyrosomus regius* and *Plectorhinchus mediterraneus*) (Figure 4). However, the β type is not present in either the Pleuronectiformes...
(S. senegalensis) or Clupeiformes orders (E. encrasicholus). Nevertheless, taking into account the highly dynamism to which the NTS are subjected, and the period elapsed since the two superorders began to diverge, another theory could be applied: a Horizontal Transference (HT) between ancient species of Perciformes and Batrachoidiformes groups. Therefore, the new type found in H. didactylus was not generated by a duplication within the own genome, but rather by an “extragenomic duplication”.

The HT phenomenon has been documented extensively in prokaryotes, and is a well-known mechanism of gene exchange [25]; it has been assumed that between 88% and 98% of the expansion of protein families in prokaryotes is due to HT [26]. However, the occurrence of HT between eukaryote organisms remains obscure [27]. It has been postulated that HT is a very important mechanism, the source for much biological innovation; it has been assumed that between two species phylogenetically distant; (ii) the rate of nucleotide variability of a transferred gene might be lower than that of the other orthologous genes; and (iii) in a phylogenetic analysis there should be incongruence between the taxonomic situation and the NTS-based phylogeny.

Another alternative way of horizontal transfer should not be discounted: the sperm-mediated gene transfer (SMGT). It is well-known that sperm cells are able to capture exogenous DNA and to transfer it to the oocyte at fertilization [43]. This capacity has led to sperm cells being used as vectors for transferring DNA in transgenic experiments [44-46]. The frequency of sperm-exogene DNA uptake depends on several factors, such as the presence of inhibitory glycoproteins and the maintenance of optimum conditions for the DNA uptake in terms of quantity, length and primary structure of the exogene DNA [47]. It has been established that the greater the quantity of exogenous DNA, the higher the possibility of take-up by sperm cells [47,48] reported that the DNA content in sediments is 3 to 4 orders of magnitude higher than in the water column. Taking this into account and the benthic behaviour of the majority of species of the Batrachoidiformes order, the possibility of a SMGT is plausible.
It is commonly considered that the U2 snRNA gene also evolves in a concerted fashion [11]. However, there are few studies with this gene to conclude unequivocally that the concerted evolution is the general model followed by the U2 snRNA gene. The organization of U2 snRNA has only been investigated in humans [49], primates [50] and moronid fishes [51], and the same characteristics are shared, i.e., U2 snRNA are clustered in tandem arrays and undergo a concerted evolution model. The U2 sequence homology observed in many different specimens and units has led us to confirm a similar situation with the U2 snRNA of H. didactylus. In all the studies mentioned above (including the present work) the U2 snRNA gene is localized in one chromosome, although the position inside the chromosome is not conserved. A dispersed gene arrangement facilitates a birth-and-death evolution model [37]; therefore a clustered arrangement could be propitious for a concerted evolution model. Moreover, taking into account the results obtained by [8] with other species from the Batrachoideidae family, it is possible that there has been a transition from a birth-and-death to a concerted evolution model within this fish family, since three distinct patterns of hybridization have been seen in only five species analyzed: dispersed, clustered, and both patterns together.

Despite the scarcity of studies describing the molecular characterization of the U2 snRNA in fish species, the only four species characterized to date, Solea senegalensis [12], Dicentrarchus labrax and D. puntatus [51], and A. regius [52] all showed a close linkage with another member of the U snRNA family. This linkage has been applied as a useful tool for the identification of different species of sole [53]. However, in H. didactylus, the U2 was not found in linkage with any other U snRNA; this absence is, so far, the only instance found within a fish group. There are insufficient data to conclude whether the U2 linkage in fishes is a general rule or is an exceptional case of only some fish families. In either case, the analysis of this gene family could be an interesting tool for evolutionary and phylogenetic purposes.

The search for promoter elements in the U2 spacer regions yielded four putative regions (Figure 2). The nearest to the 5’ coding region was located at −60 bp, and, as described by other authors [15,54], the PSE element is located between positions −60 and −50; therefore, the box found at position −60 could be the PSE regulatory element. A putative 3’ box was also localized at 8 nt downstream from the coding region, which is close to the localization described by [54] for this box (9 nt downstream). The DSE enhancer element is independent of the position and orientation, so these characteristics make it difficult to find a possible DSE by aligning with other sequences. Despite this, it has been reported that the DSE element is usually found around 200–250 bp upstream from the PSE [55]. A putative DSE element has been observed in H. didactylus at −244 bp upstream from the PSE, so that could be the enhancer element. The presence of all these regulator elements favours the supposition of a transcriptionally active gene.

The ITS-1 region without 18S and 5.8S coding regions comprises 388 bp, and the GC content is between 62.89% and 63.40%. Although this GC content is similar to the average for the Osteichthyes group (68.0% ± 4.2), the ITS-1 size is one of the smallest within this group, which varies between 318 and 1518 bp [56]. When the ITS-1 sequence was subjected to BLASTN search, an ITS-1 fragment aligned with another from many fish species, reflecting the moderate nucleotide conservation of this sequence compared with other spacer regions (IGS, NTS, etc.). This region undergoes some selective pressure, because their secondary structures are necessary for the rRNA maturation process [57]. These features make this sequence suitable for phylogenetic analysis at the species level.

The relatively low value of the nucleotide variability in the spacer regions of the three sequences treated here (Table 1) leads us to accept that their evolution has been according to the concerted model. Unequal cross-over and gene conversion are the common mechanisms that the multigene families undergo with a concerted evolution model. However, coding regions undergo another additional mechanism, such as the purifying selection, which makes them less variable. The differences in the variability found in the coding region in each gene could be due to opposing differences between homogenizing and diversifying forces. Therefore, the dynamic of the concerted evolution process depends on two conflicting forces: the homogenizing forces and the forces which generate variability [58]. It has been postulated that the unequal cross-over and gene conversion mechanisms occur more frequently in chromosome regions adjacent to the telomeres [11], thus increasing the effectiveness of concerted evolution. On this point, cytogenetics data support the conclusions already obtained with nucleotide diversity, i.e., the concerted mechanisms are more frequent in the multigene family which is situated in the telomeric position, as is the 18S rRNA, resulting in less nucleotide variation. Conversely, the two most variable multigene families (5S rDNA and U2 snRNA gene) are localized in internal positions.

None of the four probes used to hybridize in H. didactylus chromosomes were co-localized (Figure 6; Table 2); therefore they represent four different chromosomal markers, which is useful in a preliminary genome mapping. Comparisons with similar results obtained from other species of Batrachoideidae family offer some evolutionary trends (Table 2; see also [8] and [10]). The five implicated species show a variable fundamental num-
Table 2 Localization summary of the five probes used in *Halobatrachus didactylus* and other Batrachoididae species

| Species                        | 2n | FN   | (GATA)ₙ | SS rDNA | 18S rRNA | U2 snRNA | Gene co-localization |
|-------------------------------|----|------|---------|---------|----------|----------|----------------------|
| *Halobatrachus didactylus* (present work) | 46 | 72   | 1 SMT (qSc) | 1 SMT (qSc) | 1 SMT (pTl) | 1 SMT (qI) | None |
| *Batrachoides manglae* [8]    | 46 | 48   | Dispersed (+++) | 2 A (1 qStl and 1 qSc) | 1 STL (pTl) | Dispersed (+++) | 1 A (qStl) | None |
| *Porichthys plectrodon* [8]   | 46 | 62   | Dispersed (++) | 1 MT (qI) | 1 SMT (pTl) | Dispersed (+) | None |
| *Thalassophryne maculosa* [8] | 46 | 64   | Dispersed (+++) | 1 STL (qTl) | 1 SMT (pTl) | Dispersed (+) | 1 STL (qI) | 18S rRNA and U2 snRNA |
| *Amphichthys cryptogenus* [8] | 46 | 52   | Dispersed (++) | 1 STL (qSc) | 1 SMT (pTl) | Dispersed (+) | 1 SMT (qI) | 18S rRNA and 5S rDNA |

+: low abundance; ++: medium abundance; +++: high abundance; A: acrocentric; STL: subtelocentric; SMT: submetacentric; MT: metacentric; q: long arm; p: short arm; Sc: subcentromeric; I: internal; Stl; subtelomeric; Tl: telomeric.

Conclusions

For 5S rDNA, a mixed evolutionary model has been proposed, in which the presence of two different types that group in a between-species pattern reveals a birth-and-death model, but the high homogeneity of the sequences in each type indicates a homogenization derived by the concerted model. Moreover, the existence of the birth-and-death process could not be originated by a duplication of a pre-existing 5S rDNA; instead, an ancient HT event between Batrachoidiformes and Perciformes could have introduced the new type. The 5S rDNA provides the requirements for a HT because of its recognized ability to transpose to other loci [66]. Therefore future research on this type of transfer should also be directed towards ascertaining the specific mechanism by which the transfer took place.

The U2 snRNA gene was not co-located with the other repetitive sequence, so four different chromosomal markers have been obtained in *H. didactylus*, which is important for producing a preliminary physical map of the species. Moreover, cyogenetic data have shown that the U2 snRNA gene have experienced a transition from a dispersed organization to a clustered organization during the Batrachoididae evolution. A similar case has been detected for (GATA)ₙ repeats. All these characteristics make the Batrachoididae family an interesting group for studying the multigene families from the evolutionary perspective.

Methods

Sampling, PCR amplification, cloning and sequencing

Up to six specimens (3 males and 3 females) of *H. didactylus* were collected from natural populations in the Bay of Cadiz (SW Spain) for molecular purposes. Meanwhile, cyogenetic analysis was made from three specimens. Sampling procedures comply with the ARRIVE guidelines and with the UCACG052009 Regulation of the Bioethical Committee from the University of Cádiz.

Genomic DNA was isolated from muscle tissue of the six specimens using FastDNA kit® (Q-Biogene). Extraction quality was validated by electrophoresis in agarose gel (1.5%) containing 0.5 μg ml⁻¹ ethidium bromide. Three individuals were used to amplify the ITS-1 and U2 snRNA multigene families, while four were amplified for 5S rDNA (Table 3). These PCR amplifications were made using the following primers: A and B described by [67] for 5S rDNA amplification; Sp18 and Sp58 described by [68] for 18S rRNA-ITS-1-5.8S rRNA amplification; and U2ang-Fwd and U2ang-Rev described by [69] for U2 snDNA amplification. An internal reverse primer specific to the NTS of the 5S rDNA type β (SSBETAI NT: 5’-CGAGGTGCAGGAGACAAACAC-3’) was designed to corroborate the amplification in the following fish species: *H. didactylus*, *S. aurata* (positive
controls), four species of the Batrachoididae family (Porichthys pleciodon, Thalassophryne maculosa, Amp-hichthys cryptocentrus and Batrachoides manglae), two species of the Sparidae family (Pagrus pagrus and Diploplus sargus), Dicentrarchus labrax (Moronidae family). Argyrosomus regius (Sciaenidae family), Plectorhinchus mediterraneus (Haemulidae family), Solea senegalensis (Soleidae family) and Engraulis encrasicholus (Engraulidae family). Parallel to this experiment another PCR was carried out with the ITS-1 primers in the same species, as positive control for genomic DNA. Reactions were carried out in a final volume of 50 μl containing 60–80 ng of genomic DNA, 3 mM Cl2Mg, 300 μM dNTP, 0.2 pmol of the forward and reverse primers and 3 U of Taq polymerase (Euroclone). The PCR conditions were according to [70] and were performed in a Gene Amp® PCR System 2700 (Applied Biosystems) thermal cycler.

The PCR products were purified using the NucleoSpin® Extract II kit (Macherey-Nagel), and cloned into pGEM-T Easy Vector (Promega). Plasmid DNA was extracted using NucleoSpin® Plasmid (Macherey-Nagel). DNA sequencing was performed with fluorescence-labeled terminator (BigDye Terminator 3.1 Cycle Sequencing Kit; Applied Biosystems) in an ABI3100 Genetic Analyzer.

**Sequence analysis**

From 7 to 9 clones per specimen were sequenced in 5S rDNA, and 3 clones per specimen in both ITS-1 and U2 snRNA gene (Table 3). Sequence data have been deposited with the GenBank Data Library under Accession Numbers JN406320 to JN406370.

The sequences were aligned in MAFFT [71] using the slow and iterative refinement method strategy (FFT-NS-i). Consensus sequences were obtained with the Bioedit software [72]. A sequence-similarity search of coding and spacers sequences from the different multigene families was performed in BLASTN [73] to determine the similarities of the sequences obtained with other sequences from the GenBank database. Promoter and regulator regions of the 5S gene were detected by comparison with models previously described [20]. Regarding the U2 snRNA gene, the PSE and 3’ box elements were detected by aligning, respectively, the 5’ end and the 3’ end (respect to the coding region) with those obtained in the Moronidae fish family [51]. The putative DSE region was detected also comparing with those of the Moronidae family and with the consensus octameric sequence described for this region (ATGCAAT) [17]. DnaSP version 5 program [74] was used to obtain the nucleotide variability (n), the number of polymorphic sites (s) and the number of haplotypes (h) in both coding and spacer regions of the three gene families. On the other hand, evolutionary distance between 5S rDNA types was calculated, using MEGAS5 program [75], by the number of base substitutions per site from averaging over all sequence pairs, using the Kimura-2-parameter [76] in relation to evolutionary distances (d). Standard error estimates were obtained by a bootstrap procedure (1000 replicates), and complete deletion option was also applied.

The dimers obtained from 5S rDNA type α were used to extract the complete sequence of the coding region, and this sequence was subjected to the RNAstructure 5.2 program [77]. A consensus sequence from 5S rDNA type β was also obtained to be folded in RNAstructure 5.2 program. The secondary structures obtained were compared with those predicted in previous described models [16].

**Phylogenetic analysis**

A phylogenetic analysis was carried out with the NTS sequences of *H. didactylus* and those of other four species from the Batrachoididae family, which were extracted from GenBank database (from GU645582 to GU645701). Six sequences of two different species of the Sparidae family were also included in the analysis, two from *Pagrus pagrus* (Acc. No. HM004380 and HM004384), two from *Pomadasy auriga* (Acc. No. HM004389 and HM004390) and two from *Sparus aurata* (Acc. No. AY330701 and AY330702). In addition, two sequences of *Dicentrarchus labrax* were rooted as outgroup (Acc. No. HM014364 and HM014367). A Neighbor-Joining (NJ) [77] tree was firstly inferred in MEGA5 [75]. The Branch support was tested by the bootstrap method [78]. The evolutionary distances were computed using the p-distance
method [79]. This NJ tree was used thereafter as the starting tree for the inference of a Maximum-Likelihood tree in the PhyML 3.0 program [80]. Previously, a statistical selection of best-fit models of nucleotide substitution had been performed by using the jModelTest program, version 0.1.1 [81], using the Akaike Information Criterion (AIC) for model selection. The GTR model [82] was chosen as the best for the PhyML 3.0 program (−lnL= 12156.5732, AIC= 25252.6211). The branch support of the ML tree was measured by applying the non-parametric version of the approximate likelihood ratio test (SH-aLRT) [83], and the option SPR was also chosen for tree improvement.

**Cyto genetic techniques**

Chromosome preparations were made from cephalic kidney as described by [2]. Slides were pretreated with RNase, pepsin and formaldehyde according to [84]. Finally, the samples were dehydrated in successive steps using ethanol and stored at −80°C up to the moment of hybridization.

5S rDNA and U2 snRNA probes were labeled by the PCR-based method using PCR primers as previously mentioned. Similarly, 18S rDNA probe was labeled by PCR-based method using PCR primers as described by [86], using the primers (GATA)7 and (TATC)7. The labelling of GATA probe was performed by Nick Translation procedure according to the manufacturer’s instructions.

Double-FISH techniques were performed according to [10]. Finally, better images were obtained with an epifluorescence microscope (Axioskop 2 Plus, Zeiss), equipped with a cooled camera (CoolSnap, Photometrics® Inc.).

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

MAM carried out the molecular genetic studies, the sequence alignment, the phylogenetic analysis, the cytogenetic techniques, and drafted the manuscript. IC participated in the sequence and phylogenetic analysis and helped to draft the manuscript. JLP, MUM and CS carried out the sampling, tissue extraction, and chromosome preparations. LR conceived and coordinated the study, participated in its design, discussed the results and corrected the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

This work was supported by grants from the Junta de Andalucía (Spain) to the PAI BIO-219 group (LR), CACyTMAR and the “Proyecto de excelencia” RNM-03074 (CS). MUM holds a CSIC JAE fellowship.

**Author details**

1Laboratorio Genética, Facultad de Ciencias del Mar y Ambientales, CACyTMAR, Universidad de Cádiz, Puerto Real (Cádiz) 11510, Spain. 2Instituto de Ciencias Marinas de Andalucía – CSIC, Polígono Río San Pedro, Puerto Real (Cádiz) 11510, Spain. 3Instituto de Investigaciones Científicas, Universidad de Oriente, Isla de Margarita, Venezuela.

Acknowledgements

This work was supported by grants from the Junta de Andalucía (Spain) to the PAI BIO-219 group (LR), CACyTMAR and the “Proyecto de excelencia” RNM-03074 (CS). MUM holds a CSIC JAE fellowship.

Received: 29 March 2012 Accepted: 3 October 2012

**References**

1. Greenfield DW: Allenobatrachus, a new genus of Indo-Pacific toadfish (Batrachoididae), Pac Sci 1997, 51:306–313.

2. Palazón JL, Núñez M, Sarasquete C: Conventional karyotype and nucleolar organizer regions of the toadfish Holobatrachus didactylus (Schneider, 1801) (Pisces: Batrachoididae), Sci 2003, 67:445–449.

3. Roux C: Batrachoididae. In Fishes of North-eastern Atlantic and Mediterranean. 3rd edition. Edited by Whitehead PJP, Bauchot ML, Hureau JC, Nielsen JL, Torteron E. Paris: UNESCO, 1986:1360–1361.

4. Sarasquete C, Gutiérrez M, Establier R: Efecto del mercurio inorgánico (HgCl2) sobre la sangre y riñón del pez sapo marino, Holobatrachus didactylus (Schneider, 1801). Inv Pesq 1982, 46:323–330.

5. Desantis S, Cirillo F, Deftorio M, Megalofonou P, Palazón JL, Sarasquete C, de Merlio G: Histochemical study of glycoconjugates in the toadfish, Holobatrachus didactylus oesophagus epithelium, Histo Histopathol 2007, 22:23–35.

6. Palazón JL, Arias AM, Sarasquete C: Aspects of the reproductive biology of the toadfish, Holobatrachus didactylus Schneider, 1801) (Pisces: Batrachoididae), Sci 2001, 65:131–138.

7. Sarasquete C: Estudios de las series eritrocítica, leucocítica y trombocítica de la sangre periférica del pez sapo marino, Holobatrachus didactylus. Inv Pesq 1983, 47:410–412.

8. Ubéda-Manzanares M, Merlo MA, Palazón JL, Cross I, Sarasquete C, Rebordinos L: Chromosomal mapping of the major and minor ribosomal genes, (GATA)n and U2 snRNA gene by double-colour FISH in species of the Batrachoididae family. Genetica 2010, 138:787–914.

9. Ubéda-Manzanares M, Merlo MA, Palazón JL, Sarasquete C, Rebordinos L: Sequence characterization and phylogenetic analysis of the 5S ribosomal DNA in species of the family Batrachoididae. Genomre 2010, 53:723–730.

10. Merlo A, Cross I, Palazón JL, Sarasquete C, Rebordinos L: Chromosomal mapping of the major and minor ribosomal genes, (GATA)n and (TTAGGG)n by one-color and double-color FISH reveals three chromosomal markers in the toadfish Holobatrachus didactylus (Teleostei: Batrachoididae). Genetica 2007, 131:195–200.

11. Nei M, Rooney AP: Concerted Birth-and-Death evolution of multigene families. Annu Rev Genet 2005, 39:121–152.

12. Manchado M, Zasiti E, Cross I, Merlo A, Infante C, Rebordinos L: 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 2006, 49:79–86.

13. Pinhal D, Araki CS, Gadig OB, Martins C: Molecular organization of 5S rDNA in sharks of the genus Rhizoprionodon: insights into the evolutionary dynamics of 5S rDNA in vertebrate genomes. Genet Res, Camb 2009, 91:61–72.

14. Verna J, González-Tojal AM, Martínez-Lage A: 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 Genom 2009, 47:635–644.

15. McNamara-Schroeder KJ, Hennessy RF, Harding GA, Jensen RC, Stumph WE: The Drosophila U1 and U6 gene proximal sequence elements act as important determinants of the RNA polymerase specificity of small nuclear RNA gene promoters in vitro and in vivo. J Biol Chem 2001, 276:31786–31792.

16. Szymanski M, Barciszewka MZ, Erdmann VA, Barciszewski J: 5S ribosomal RNA database. Nucleic Acids Res 2002, 30:176–178.

17. Tebø G, Mattaj IW: The Xenopus laevis U2 gene Distal Sequence Element (enhancer) is composed of four subdomains that can act independently and are partly functionally redundant. Mol Cel Biol 1989, 9:1682–1690.

18. Solà L, Rossi AR, Annesi F, Gonnord E: Cytogenetic studies in Sparus auratus (Pisces, Perciformes): molecular organization of SS rDNA and chromosomal mapping of SS and 4SS ribosomal genes and of telomeric repeats. Hereditas 2003, 139:232–236.

19. Zhu HP, Lu MX, Gao FY, Zhang ZH, Yang LP, Gui JF: Chromosomal localization of rDNA genes and genomic organization of SS rDNA in Oreochromis mossambicus, O. urolepis hornorum and their hybrid. J Genet 2010, 89:163–171.
66. Drouin G, Moriz de Sá M: The concerted evolution of SS ribosomal genes linked to the repeat units of other multigene families. Mol Biol Evol 1995, 12:481–493.

67. Pendas AM, Morán P, Freije JP, García-Vázquez E: Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon SS rDNA. Cytogenet Cell Genet 1994, 67:31–36.

68. Kurtiwa K, Hanzawa N, Yoshito T, Kimura S, Nishida M: Phylogenetic relationships and natural hybridization in rabbitfishes (Teleostei: Siganidae) inferred from mitochondrial and nuclear DNA analyses. Mol Phylogenet Evol 2007, 45:69–80.

69. Cross I: Caracterización poblacional y citogenética de Crassostrea angulata. Departamento de Biomedicina, Biotecnología y Public Health, 2005.

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

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

72. Kimura M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987, 4:406–425.

73. Kimura M: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Biol 1980, 162:111–120.

74. Saitou N, Nei M: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011, 28:2731–2739.

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

76. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O: New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol 2010, 59:307–321.

77. Posada D: jModelTest: phylogenetic model averaging. Mol Biol Evol 2008, 23:1253–1256.

78. Lanave C, Preparata G, Saccone C, Serio G: A new method for calculating evolutionary substitution rates. J Mol Biol 1984, 180:86–93.

79. Sattou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1991, 8:501–511.

80. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O: New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol 2010, 59:307–321.

81. Posada D: jModelTest: phylogenetic model averaging. Mol Biol Evol 2008, 23:1253–1256.

82. Lanave C, Preparata G, Saccone C, Serio G: A new method for calculating evolutionary substitution rates. J Mol Biol 1984, 180:86–93.

83. Anisimova M, Gili M, Dufayard J-F, Desimoz C, Gascuel O: Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst Biol 2011, 60:685–699.

84. Jansen G, Devaere S, Weeks PHH, Adriaans D: Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences. Mol Phylogenet Evol 2006, 38:655–76.

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

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

Cite this article as: Merlo et al.: Evidence for SS rDNA Horizontal Transfer in the toadfish Holobatrachus didactylus (Schneider, 1801) based on the analysis of three multigene families. BMC Evolutionary Biology 2012 12:201.