J Mol Evol (1999) 49:250–259

JOURNAL OF MOLECULAR EVOLUTION

Origin and Evolution of Paralogous rRNA Gene Clusters Within the Flatworm Family Dugesiidae (Platyhelminthes, Tricladida)

Salvador Carranza,* Jaume Baguñà, Marta Riutort

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avinguda Diagonal 645, 08071 Barcelona, Spain

Received: 24 March 1998 / Accepted: 17 March 1999

Abstract. Analysis of the 18S rDNA sequences of five species of the family Dugesiidae (phylum Platyhelminthes, suborder Tricladida, infraorder Paludicola) and eight species belonging to families Dendrocoelidae and Planaridae and to the infraorder Maricola showed that members of the family Dugesiidae have two types of 18S rDNA genes, while the rest of the species have only one. The duplication event also affected the ITS-1, 5.8S, ITS-2 region and probably the 28S gene. The mean divergence value between the type I and the type II sequences is 9% and type II 18S rDNA genes are evolving 2.3 times more rapidly than type I. The evolutionary rates of type I and type II genes were calibrated from biogeographical data, and an approximate date for the duplication event of 80–120 million years ago was calculated. The type II gene was shown, by RT-PCR, to be transcribed in adult individuals of *Schmidtea polychroa,* though at very low levels. This result, together with the fact that most of the functionally important positions for small-subunit rRNA in prokaryotes have been conserved, indicates that the type II gene is probably functional.

Key words: 18S rDNA — Polymorphism — Dugesiidae — Platyhelminthes — Variability — Multigene family — Metazoan — Phylogeny

Introduction

The ribosomal DNA (rDNA) array in eukaryotes consists of the 18S, 5.8S, and 28S coding regions (or their homologues), separated by two internal transcribed spacers (ITS-1 and ITS-2). An external transcribed spacer (ETS) is located upstream of the 18S gene. Adjacent copies of the rDNA repeat unit are separated by a nontranscribed spacer (NTS), also called an intergenic spacer. Once nuclear rDNA genes had been studied in detail, it became clear that the multiple copies had not evolved independently (Dover and Coen 1981; Krystal et al. 1981; Arnheim 1983). The process that is believed to maintain similarities among the different copies of the repeat is referred to as concerted evolution (Dover 1982; Hillis and Dixon 1991) and two molecular mechanisms are assumed to account for it: gene conversion (Lassner and Dvorak 1986; Hillis et al. 1991) and unequal crossing-over (Arnheim 1983; Coen and Dover 1983). Differences between arrays are mostly length variations within the NTS (e.g., Arnheim et al. 1982; Spencer et al. 1987), although intragenomic variation also occurs among the multiple copies of the ribosomal coding regions within an individual (Carranza et al. 1996; Buckler et al. 1997; Stothard et al. 1997; Telford and Holland 1997; VanCampenhout et al. 1997).

In a previous paper (Carranza et al. 1996), we reported the first example of 18S rDNA polymorphism within the genome of a metazoan species, *Schmidtea mediterranea,* a free-living platyhelminth of the order Tricladida. the phylogenetic analysis (Carranza et al. 1996) suggested that the 18S rDNA duplication is common to most species belonging to the family Dugesiidae,

* *Present address:* Department of Zoology, The Natural History Museum, Cromwell Road, SW7 5BD London, UK
Correspondence to: M. Riutort; *e-mail:* martar@porthos.bio.ub.es

Table 1. Taxonomic classification and species name, type of 18S sequence found, references, and accession numbers for the 18 sequences sampled in this study

Phylum Platyhelminthes[a]	Type of 18S sequence	Reference(s)	Accession No.
O Seriata			
SO Tricladida			
IO Paludicola			
F Dugesiidae			
Schmidtea mediterranea	I and II	Carranza et al. (1996)	U31084 (type I)/U31085 (type II)
Schmidtea polychroa	I and II	This study	AF013152 (type I)/AF013154 (type II)
Dugesia subtentaculata	I and II	This study	M58343 (type I)/AF013155 (type II)
Dugesia japonica	I and II	This study/Katayama *et al.* (1995)	AF01353 (type I)/D83382 (type II)
Girardia tigrina	I and II	This study	AF013157 (type I)/AF013156 (type II)
F Planariidae			
Phagocata ullala	Only one type	This study	AF013149
Phagocata sp.	Only one type	This study	AF013150
Crenobia alpina	Only one type	This study	M58345
Polycelis nigra	Only one type	This study	AF013151
F Dendrocoelidae			
Dendrocoelum lacteum	Only one type	This study	M58346
Dendrocoelopsis sp.	Only one type	Katayama et al. (1996)	D85087
IO Maricola			
F Uteriporidae			
Ectoplana sp.	Only one type	Katayama et al. (1996)	D85088
Uteriporus sp.	Only one type	This study	AF013148

[a] *Schmidtea, Dugesia,* and *Girardia,* formerly within the genus *Dugesia,* are considered new genera of the family Dugesiidae, as suggested in the taxonomic revision of the family Dugesiidae (Ball 1974; DeVries and Sluys 1991). O, order; SO, suborder; IO, infraorder; F, family.

whereas it is probably lacking in the sister families Planariidae and Dendrocoelidae. Sequence analyses also showed that type I and type II 18S rDNA sequences were highly divergent. It was also shown that, although the distribution of changes along the type II sequences seemed to maintain the secondary structure, probably due to functional constraints, it was not possible to detect the expression of the type II gene by Northern blot experiments.

The aims of this paper are (1) to establish the presence or absence of the duplicated 18S genes in the genome of other dugesiid species as well as in planariids and dendrocoelids; (2) to attempt to determine which parts of the rDNA array were affected by the duplication event; (3) to compare the evolutionary rates of type I and type II genes and to date the duplication event; and, finally, (4) to determine whether there is any evidence to indicate that the type II 18S gene is functional.

Materials and Methods

Organisms

Table 1 lists the species studied, types of 18S sequence found, and references. EMBL accession numbers for the 13 18S sequences identified in this study plus 5 other turbellarian sequences taken from GenBank are also listed.

DNA Extraction

High molecular weight DNA was purified according to a modification (Garcia-Fernàndez et al. 1993) of the guanidine isothiocyanate method initially described for RNA (Chirgwin et al. 1979).

Amplification and Sequencing

The 18S rDNA genes were PCR-amplified in two halves and 10 internal primers were used to sequence both chains completely. To amplify the type I 18S rDNA gene of *D. japonica* two type I-specific primers were designed on the basis of the type I 18S sequences known from *G. tigrina, S. polychroa, S. mediterranea,* and *D. subtentaculata:* type IF (5′-CGC GAA GGT AAT TCA AAT CTC CTT-3′) and its complementary type IR. Together with two universal primers, these two type I-specific primers allowed us to amplify the type I 18S molecule of *D. japonica* in two parts (1F-typeIR and typeIF-9R). The 10 internal primers used to sequence all the 18S rDNA and the PCR conditions were as described by Carranza et al. (1996).

In *Schmidtea mediterranea* two sets of primers were used to amplify the fragment of the rDNA cluster spanning from the 18S rDNA to the 28S rDNA. The primers used were 18STIF 5′ TGTCGTGTCA GGAATAGTGG 3′ and 28SR 5′ CCTTGTTAGT TTCTTTTCCT C 3′ to amplify the type I cluster and 9F 5′ GTAGGTGAAC CTGCGGAAGG 3′ and 28SR (see above) to amplify the type II cluster.

The PCR products were purified with the GenecleanR II kit (BIO 101 Inc.) and directly sequenced using an automated ABI 377 sequencer or cloned into the pUC18 vector as described by Carranza et al. (1996). Cycle sequencing using dye-labeled terminators (PrismTM Ready Reaction DyeDeoxyTM Terminator Cycle Sequencing Kit) was performed in a Perkin–Elmer 480 DNA Thermal Cycler according to the manufacturer's instructions.

RNA Extraction

Total cellular RNA was isolated by the guanidine isothiocyanate method (Chirgwin et al. 1979) from live individuals of three species of the family Dugesiidae (*Dugesia subtentaculata, Schmidtea polychroa*, and *Girardia tigrina*), three species of the family Planariidae (*Phagocata ullala, Crenobia alpina*, and *Polycelis nigra*), and one species of the family Dendrocoelidae (*Dendrocoelum lacteum*).

RT-PCR

RNA extracted from each species was resuspended in 20 μl of H_2O DEPC, 15 U of DNAse (RNAse free) was added, and the mixture was incubated at 37°C for 20 min. The volume was then increased to 100 μl with H_2O DEPC and phenol extraction was performed to remove the DNAse. The RNA was precipitated with 0.1 vol of 3 *M* NaOAc, pH 5.2, and 2 vol of absolute ethanol; the solution was held at −70°C for 30 min and then centrifugated for 30 min at 12,500 rpm on a microfuge at RT. The pellet was washed in 70% ethanol and air-dried for 20 min at RT.

About 1 μg of RNA (free of DNA) from each organism was retrotranscribed to single-strand cDNA with the AMV reverse-transcriptase enzyme. The reaction conditions were as follows: 5 μl of annealing mix (1 μg of RNA, 10 pmol of the 9R primer, and H_2O DEPC to 4 μl) was held at 65°C for 10 min, transferred to 37°C for 10 min, and chilled on ice. Then 20 μl of reaction mix [1 μl of 10 m*M* dNTPs, 40 U of RNAse inhibitor, 25 U of AMV reverse transcriptase, 5 μl of AMV buffer (×5), and H_2O DEPC to 20 μl] was added, and the mixture was kept at 42°C for 45 min. The tube was then transferred to a water bath at 100°C for 5 min to inactivate the AMV enzyme. After that, 75 μl of distilled H_2O was added to the mix. PCR experiments were performed with 1 μl of this 18S cDNA single-strand solution following normal PCR conditions as described previously (Carranza et al. 1996). PCR products were sequenced as described above.

Southern Blot Analyses

Total genomic DNA from *D. subtentaculata* (3 mg), *S. mediterranea* (2 mg), and *P. ullala* (3 mg) was single-digested with *Eco*RI. DNAs were separated on a 0.9% agarose gel and transferred to Hybond-N membranes (Amersham) by capillary blotting. Probes were generated in the following way: 3F and 5R primers (Carranza et al. 1996) were used to PCR amplify a 570-bp region of the 18S gene. This region includes regions that are highly conserved among all the metazoa and highly variable regions, which differ in closely related species. The PCR was performed using two sources, DNA and RNA. The two PCR products for each species were visualized in an agarose gel, purified using the GenecleanR II kit and labeled together by random priming with ^{32}P. Hybridization was carried out at 42°C overnight in 2× SSC in the presence of 45% formamide and 10% sodium dextransulfate. Filters were washed successively with 2× SSC, 0.1% SDS at room temperature for 10 min, 2× SSC, 0.1% SDS at 65°C for 10 min, and 1× SSC, 0.1% SDS at 65°C for 10 min.

Sequence Alignment and Phylogenetic Analyses

Sequence data were aligned by hand with the help of a computer editor (GDE 2.2). Alignment gaps were inserted to account for putative length differences between sequences. A secondary structure model (Gutell et al. 1985) was used to optimize alignment of homologous nucleotide positions. These positions that could not be unambiguously aligned were subsequently excluded, resulting in a total of 1657 positions that could be used in the phylogenetic analysis (719 being variable and 274 being parsimony informative when all the taxa are compared). The sequence alignment and the weighting mask that determine the nucleotide positions taken for the analyses are available in GDE format from the ftp site, porthos.bio.ub.es/pub/incoming/phylogeny/S2phylogeny.gde. For estimation of the maximum-likelihood (ML) trees and for bootstrap analysis (100 replicates), we used the FastDNAml program, version 1.1.1a (with global rearrangements and reordering of species) (Felsenstein 1981; Olsen et al. 1994). Distance analyses were calculated using the PHYLIP program package, version 3.52 (Felsenstein 1993). A distance matrix was generated from the aligned sequences using the program DNADIST and corrected with the two-parameter method of Kimura (1980). The distances were then converted to phylogenetic trees using the neighbor-joining (NJ) method of Saitou and Nei (1987) provided by the NEIGHBOR program. Bootstrap resampling (Felsenstein 1985) was accomplished with the use of the programs SEQBOOT (1000 replicates) and CONSENSE. The most-parsimonious (MP) cladogram was reconstructed with the branch-and-bound search option in PAUP* (Swofford 1998). The reliability of the cladogram was assessed by the bootstrap test using the heuristic search option and 1000 replicates and by branch support (Bremer 1994). The numbers of transitions and transversions for all pairwise comparisons and for all six types of base change were calculated with the MEGA program (Kumar et al. 1994).

Relative Rate Test

The clade including families Dendrocoelidae and Planaridae was used as the outgroup since they are the sister group to the family Dugesiidae (Ball 1981; Sluys 1989; Riutort et al. 1992; Carranza et al. 1996). The nucleotide substitutions per site were computed as the total number of substitutions (*K*). For rate calculations, only base substitutions were used. A mean length of 1610 bp was used for all the 18S rDNA sequences. Standard errors for the relative rate test were calculated following the formulae given by Wu and Li (1985) and Li and Tanimura (1987).

Contingency Test

The frequencies of the six possible base changes among type I sequences versus those among type II sequences were tested for homogeneity using a contingency test (Sokal and Rohlf, 1994). Six comparisons were done in which the absolute number of each class of change (AT, AG, AC, GC, GT, CT) was compared for the two groups of sequences against the rest of the changes using a 5% χ^2 test. To calculate the expected value of a given class of change for either type I or type II sequences, the number of changes of that class (adding type I and type II) was multiplied by the number of changes of all classes in the given type of sequence, and the result divided by the total number of changes (of all classes and for both types of sequences) that had been observed.

Results and Discussion

Characterization of Type I and Type II 18S Sequences

Preliminary evidence that a ribosomal duplication event took place in the genome of some dugesiids but not in the planariids was obtained from Southern blots of genomic DNA from *Phagocata ullala* (family Planariidae), *S. mediterranea* (family Dugesiidae), and *D. subtentacu-*

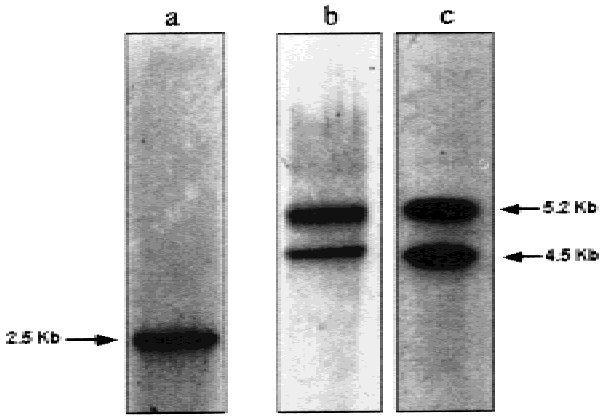

Fig. 1. Comparison of restriction patterns seen following digests of *P. ullala* **(a),** *D. subtentaculata* **(b),** and *S. mediterranea* **(c)** probed with their 3F-5R specific probes. In lanes b and c, the upper bands are around 5.2 kb. This is the expected size of the *Eco*RI type I rDNA segment for *S. mediterranea* (Carranza et al. 1996). The single band of approximately 2.5 kb of the *Eco*RI digestion for *P. ullala* supports the idea that there is only one type of rDNA cluster in this species.

lata (family Dugesiidae). Blots of *S. mediterranea* and *D. subtentaculata* showed two bands (Figs. 1b and c). This was expected, as *Eco*RI cuts the two rDNA clusters of *S. mediterranea* at different sites, giving the two-band pattern corresponding to the type I and type II 18S genes (Carranza et al. 1996). In contrast, the blot of *P. ullala* showed only a single band (Fig. 1a), which indicates that this species has only one type of 18S rDNA. If two or more types of 18S rDNA gene were present in the genome of *P. ullala,* this would have been detected by the universal region of the probe. We should bear in mind that, however unlikely, a single band may also result if, in *P. ullala, Eco*RI produces fragments of the same size for the two clusters.

To analyze further the presence or absence of the two types of 18S rDNA genes in other dugesiid, planariid, and dendrocoelid species, high molecular weight DNA was isolated from several triclads (see Table 1). The 18S rDNA molecule was PCR-amplified from total DNA using the 1F and 9R 18S universal primers (Carranza et al. 1996). The PCR products of four representatives of the family Dugesiidae (*S. polychroa, D. subtentaculata, D. japonica,* and *G. tigrina*), three representatives of the family Planariidae (*C. alpina, P. ullala,* and *P. nigra*), and one representative of the family Dendrocoelidae (*D. lacteum*) were cloned into pUC 18.

Full-length sequencing of the cloned 18S rDNA PCR products of *S. polychroa, D. subtentaculata,* and *G. tigrina* revealed that these sequences were different from the partial 18S rRNA sequences already published (Riutort et al. 1992), clustering with the type II sequence of *S. mediterranea* (Carranza et al. 1996) in preliminary phylogenetic trees (data not shown). We refer to these 18S rDNA sequences as type II. The 18S rDNA sequence of *D. japonica* was identical to that published by Katayama et al. (1995) and so it was only partially sequenced. In addition, this sequence from GenBank also clustered with the type II sequences in preliminary phylogenetic trees. *C. alpina* and *D. lacteum* 18S rDNA sequences were identical to the partial rRNA sequences already published (Riutort et al. 1992). The 18S rDNA sequences of *P. nigra* and *P. ullala* were the first records of the species and could not be compared.

As shown previously, whereas two types of 18S rDNA genes are present in the genome of *S. mediterranea,* only the type I sequence was recovered by direct sequencing from RNA (Riutort et al. 1992; Carranza et al. 1996). With this in mind, total RNA was extracted and retrotranscribed with the 9R universal primer (see material and methods) from a pool of live organisms for each of the species in which cloning into pUC 18 had been applied, the only exception being *D. japonica;* the only specimens available for this species had been fixed in 100% ethanol and total RNA could not be extracted. The 18S rRNA molecule was amplified using the RT-PCR technique. Full-length direct sequencing of the PCR products showed the RNA-derived 18S sequences of *S. polychroa, G. tigrina,* and *D. subtentaculata* to be identical to the partial 18S rRNA sequences already published (Riutort et al. 1992) and very similar to the type I sequence of *S. mediterranea.* We refer to these 18S rRNA sequences as type I. Moreover, two type I-specific primers together with two universal primers allowed us to amplify and sequence the type I 18S molecule of *D. japonica.* The 18S rRNA sequences of *C. alpina, P. ullala, P. nigra,* and *D. lacteum* were identical to the sequences from cloned 18S rDNA PCR products (see above) and were only partially sequenced from RNA. From this result it seemed that while in all the species of the families Planariidae and Dendrocoelidae there was only one type of sequence, all the representatives of the family Dugesiidae had two types.

Molecular Phylogenetic Analysis

To analyze the evolutionary history of the duplication, a phylogenetic analysis using all the sequences listed in Table 1 was performed using maximum-likelihood (ML), maximum-parsimony (MP), and neighbor-joining (NJ) methods. All gave very similar results, differing only in minimal internal rearrangements within the families Dugesiidae and Planariidae. Figure 2 shows the ML tree with the bootstrap values for the three methods and the branch support value. The topology of the tree clearly shows that the 18S sequences of the representatives of the families Dendrocoelidae and Planariidae do not belong either to type I or to type II 18S sequences. Type I and type II sequences of the dugesiids cluster together with high bootstrap and branch support values (see Fig. 2). Consequently, the phylogenetic analysis confirms

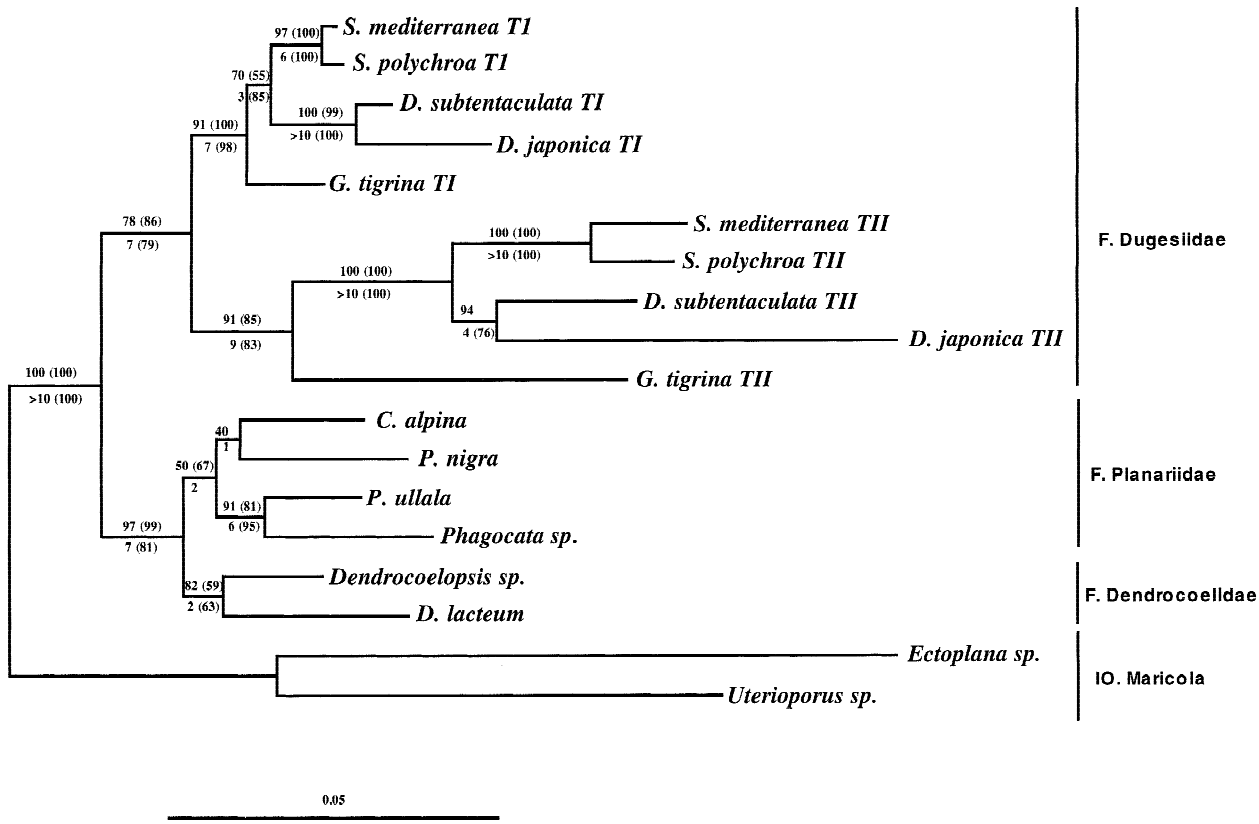

Fig. 2. Maximum-likelihood tree for the 18 18S sequences studied. Parsimony-informative sites, 274; variable sites, 719. *Numbers above the nodes* indicate the bootstrap proportions for the ML tree (1000 replicates); *numbers in parentheses above the nodes* indicate the bootstrap proportions for the NJ tree (1000 replicates); *numbers below the nodes* indicate Bremmer support values; *numbers in parentheses below the nodes* indicate the bootstrap proportions for the MP tree (1000 replicates). We show the bootstrap proportions for the NJ and MP analyses only on those nodes in which their topology is compatible with the ML tree.

that the duplication of the rDNA array occurred after the ancestor of the Dendrocoelidae and Planariidae branched off from the ancestor of the Dugesiidae. The same branching order is observed within the type I and type II sequences of the dugesiids, suggesting that although type II sequences evolve 2.3 times more rapidly than type I sequences (see below), the former are not saturated and have not lost its phylogenetic signal. From a phylogenetic point of view, the ribosomal duplication presented here is a valuable taxonomic molecular character within the Tricladida that clearly defines the family Dugesiidae and differentiates it from its sister families Planariidae and Dendrocoelidae. The analysis of this gene duplication in other triclad groups such as the Terricola and some doubtful representatives of the family Planariidae can contribute to the better knowledge of this group of platyhelminthes.

Did the Duplication Event Affect the Whole Cluster?

Analysis of the sequence region between the 18S and the 28S ribosomal genes in *S. mediterranea* showed that in both types of sequences, an ITS-1 spacer, a 5.8S gene, and an ITS-2 spacer exist (accession numbers for the ITS-1/5.8S/ITS-2 type I and type II ribosomal regions are AF047853 and AF047854, respectively). These regions are different for the two types, the differences including length variability among the spacers, which preclude their alignment—952 bp of the ITS-1 type I vs 520 bp of the ITS-1 type II and 578 bp of the ITS-2 type I vs 379 bp of the ITS-2 type II—and equal lengths (124 bp) but 6.79% divergence (Kimura two-parameter correction) between the two 5.8S rRNA genes. The finding of two types of ITS-1, 5.8S and ITS-2, each specifically associated with either the type I or the type II 18S rDNA, suggests that the duplication event affected the whole ribosomal cluster.

Differences in DNA composition were analyzed in 18S (type I and II), ITS-1 (type I and II), and ITS-2 (type I and II). Because there have been reports on compositional differences between stem and loop regions in rRNA (Vawter and Brown 1993; Friedrich and Tautz 1997), these two classes were analyzed also separately. As suggested by Sueoka (1988), changes in DNA composition in neutrally evolving sequences should reflect changes in directional mutation pressure. Friedrich and Tautz (1997) clearly show that the DNA composition in

Table 2. Nucleotide composition of the sequences (percentages)[a]

	A	T	C	G
S. mediterranea				
Type I	28.8 (29.0)	27.6 (38.2)	18.8 (14.7)	24.8 (17.8)
Type II	30.1 (30.1)	28.8 (38.6)	18.0 (14.1)	23.2 (17.0)
S. polychroa				
Type I	28.8	27.7	18.8	24.7
Type II	30.2	28.4	18.2	23.2
D. subtentaculata				
Type I	28.8	27.7	19.0	24.5
Type II	30.7	28.3	18.3	22.8
D. japonica				
Type I	29.5	27.6	19.2	23.8
Type II	30.8	27.1	18.0	24.1
G. tigrina				
Type I	29.0	27.3	19.1	24.7
Type II	29.6	27.9	18.3	24.2
Stems				
Type I	23.1	28.0	20.0	28.9
Type II	25.1	28.7	18.9	27.3
Loops				
Type I	37.4	26.7	17.4	18.5
Type II	37.7	27.2	16.7	18.4

[a] The numbers in parentheses represent the composition for the ITS regions. For stems and loops the average values for all type I and all type II sequences are given. PUZZLE was used to compare base compositions for the sequences using a 5% χ^2 test on the average composition; no significant difference was found between type I and type II sequences, even when stem and loop regions were anlyzed independently.

28S rRNA expansion segments (likely to follow a neutral evolution model) of dipteran vs all other nondipteran hexapod taxa is biased toward A+T, suggesting a different mutation pressure acting on dipterans. In our case, the results shown in Table 2 clearly indicate that neither in the most conserved parts of the 18S rDNA nor in the supposedly neutrally evolving regions of the ITS-1 and ITS-2 are there significant differences in base composition. This result suggests that even if slight differences in the types of substitutions between type I and type II sequences occur (see below), these do not affect the base composition of the rDNA clusters as is the case in other invertebrate taxa (Friedrich and Tautz 1997).

Molecular Evolution of Type I and Type II Sequences

The percentage of the six types of base changes (AG, TC, AT, AC, TG, CG), the percentage of transitions and transversions, and the transition (Ts)/transversion (Tv) ratio were calculated independently for the type I and type II sequences and all the representatives of the family Planariidae + Dendrocoelidae 18S sequences. The results are shown in Table 3. Type II sequences have a very low Ts/Tv ratio (below 1; see Table 3). The Ts/Tv ratio is known to decrease as the number of changes among sequences increases (Brown et al. 1982); consequently, a low Ts/Tv ratio has been considered as an indicator that sequences could be saturated (Holmquist 1983). However, plotting the observed substitutions (total, transitions, or transversions) among these sequences against the number of real changes estimated from the maximum-parsimony tree (Philippe, 1993) shows they are not saturated (data not shown; linear correlation, $R = 0.938$).

The distribution of the various types of change in the three groups of sequences is very heterogeneous (Table 3). A contingency test showed that for all types of changes (except TG), type I and type II sequences had a significantly different frequency. The reason for this deviation is not known, although it seems to indicate that type I and type II sequences, although part of the same genome, are subject to different substitution pressures.

A relative rate test for type I and type II sequences using the families Planariidae + Dendrocoelidae as the outgroup is shown in Table 4. Type I and type II sequences were compared within each species; thus, differences in their *K* value cannot be due to an increase in one lineage vs. another. All comparisons showed significant differences between type I and type II sequences, above two standard errors, the only exception being *D. subtentaculata.* The mean *K* value for the type I and type II lineages was calculated (Table 5) following Li and Tanimura (1987). Mean *K* values of 0.0264 substitution per site (sub/site) and 0.0604 sub/site for type I and type II 18S genes, respectively, were found. Our calculation shows that, on average, type II genes evolve 2.3 times more rapidly than type I. Acceleration in the evolutionary rate may occur after gene duplications. The lack of a functional constraint allows an increased substitution rate for the duplicated genes, until one of them acquires a new function or becomes a pseudogene (Ohta 1991, 1994). A number of such cases have been described for both coding and regulatory regions: immunoglobulins, cytochrome P450, hemoglobins, growth hormone, interferon α, chorion, vitellogenin, esterase, lactic dehydrogenase, and others (Otha 1991). Goodman (1981) and Czelusinak et al. (1982) showed that extremely high rates of amino acid substitutions occurred following the gene duplications separating α and β hemoglobins. These high rates were due to advantageous mutations improving the function of hemoglobins. Multigene families such as histones and rDNA have been considered to be special cases of gene duplication since their members are rendered uniform by concerted-evolution mechanisms (Ohta 1991). Type I and type II 18S rDNA in dugesiids are an exception to this rule. Chromosome rearrangements, such as translocation of part of the rDNA cluster to a new chromosome or duplication of a whole chromosome bearing that cluster, may have shifted the duplicated genes (or part of the old cluster) to a new chromosome or a new location within the same chromosome. Preliminary results with fluorescence in situ hybridization (FISH) on metaphase chromosome of *S. mediterranea*

Table 3. Percentage values for the six possible base changes, percentage of transitions and transversions, and transition/transversions ratio[a]

	AG	TC	AT	AC	TG	CG	Ts	Tv	Ts/Tv
Type I sequence (Dugesiidae)	32.5	24.7	13.2	13.4	10.4	5.8	57.2	42.8	1.33
Type II sequence (Dugesiidae)	25.8	18.6	25	8.6	11.3	10.7	44.4	55.6	0.8
Planariidae + Dendrocoelidae	25.4	28.4	23.1	10.1	7.2	5.7	53.8	46.1	1.16

[a] Each base change value was calculated by dividing the number of a particular base change within a given group of sequences by the total number of transitions and transversions within the same group of sequences and multiplying by 100. The percentage of transitions and transversions was calculated as the sum of two (AG+TC) and four (AT+AC+TG+CG) base change percentages, respectively. The Ts/Tv ratio was calculated by dividing the total number of transitions by the total number of transversions within each particular group of sequences.

Table 4. Relative rate test between type I and type II sequences for each dugesiid species[a]

	K_{12}	$K_{13} - K_{23}$
S. mediterranea	0.0827	0.0395 ± 0.0139*
S. polychroa	0.0834	0.0368 ± 0.0143*
D. subtentaculata	0.0683	0.0259 ± 0.0144
D. japonica	0.1199	0.0337 ± 0.0166*
G. tigrina	0.0797	0.0340 ± 0.0142*

[a] K = nucleotide substitutions per site; 1 = 18S type I sequence; 2 = 18S type II sequence; 3 = 18S sequence of Planariidae + Dendrocoelidae. In all comparisons, the number of sites = 1610.
* >2 SE.

Table 5. Rate of nucleotide substitutions per site for type I and type II lineages since their separation for each species[a]

	Type I	Type II
S. mediterranea	0.0216	0.0611
S. polychroa	0.0233	0.0601
D. subtentaculata	0.0212	0.0471
D. japonica	0.0431	0.0768
G. tigrina	0.0228	0.0569
Mean	0.0264 ± 0.0083	0.0604 ± 0.0096

[a] The number of substitutions per site was calculated as $[K_{12} + (K_{13} - K_{23})]/2$ for type I sequences and $[K_{12} - (K_{13} - K_{23})]/2$ for type II.

show that both type I and type II rDNA clusters are located in nonhomologous loci (Carranza 1997). This may change the substitution rate of the duplicated gene (Wolfe et al. 1989) and decrease the possibility of their being uniformized with their paralogous relatives by concerted evolution, although concerted evolution may be operationally but independently active in both clusters.

Absolute Evolutionary Rate of the Two 18S rDNA Genes

To compare the rate of evolution of the two 18S rDNA genes to that of genes from other organisms or to other genes, it is necessary to calculate absolute evolutionary rates. These can be calculated only on the basis of known fossils or biogeographical events. Freshwater planarians, and Platyhelminthes as a whole, are small, soft-bodied organisms that lack a fossil record. Several biogeographical features of the family Dugesiidae may help to calibrate this age. Ball (1974) has convincingly argued that whereas the genus *Girardia* has a North American distribution, the closely related genera *Schmidtea* and *Dugesia* have a mainly European and Euroasian distribution, respectively. Hence, Ball (1974) hypothesized that the ancestor of the family was present in Laurasia before the splitting of the two continents around 70 million years ago (MYA). We used these data to calculate the rate of change per million years (Myr) for the two types of sequences independently (mean K values for *G. tigrina* vs the other dugesiids/2 × 70 MYA). The results give estimates of the evolutionary rate of 2.2×10^{-4} substitutions/site/Myr for type I sequences and 7.5×10^{-4} substitutions/site/Myr for type II sequences. The values found here for the type I 18S gene can be compared with those found in other studies. Larson and Wilson (1989) estimated a rate of 3×10^{-4} changes/site/Myr for divergent nuclear rDNAs in ambistomid salamanders, Spears et al. (1993) found a rate of 2×10^{-4} changes/site/Myr for certain crustaceans, Friedrich and Tautz (1997) estimated rates between 0.7 and 3.3×10^{-4} changes/site/Myr for different groups of the Diptera, and Pawlowsky et al. (1997) estimated a rate of $0.2–5 \times 10^{-4}$ changes/site/Myr for benthic foraminifera. Both type I and type II sequences are also within the range found for the rate of nonsynonymous substitutions in coding regions of nuclear genes for mammalian species [$<10 \times 10^{-4}$ changes/site/Myr on average (Li and Graur 1991 and references therein)] and are clearly below the evolutionary rate estimated for the dipteran stem lineage [19.2×10^{-4} changes/site/Myr (Friedrich and Tautz 1997)] and for the planktonic foraminifera [$20–30 \times 10^{-4}$ changes/site/Myr (Pawlowsky et al. 1997)].

These results also provide a rough estimate of the time of the duplication event by dividing the mean K values for type I and type II 18S rDNA genes (see Table 5) by their respective evolutionary rates (see above). This calculation gives an estimate of 120 MYA for the duplication event using the information from the type I genes and 80 MYA using the information from the type II genes. The differences in the estimates of the age of the duplication event may be simply a statistical error or, more interestingly, may be due to variations in the rate of change of type I or type II genes between the 18S du-

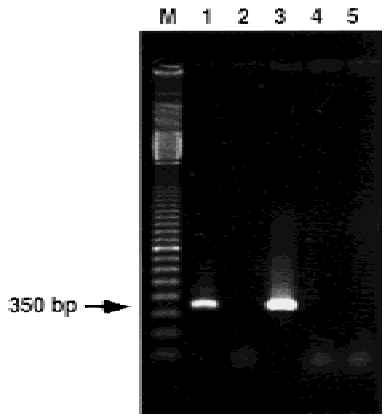

Fig. 3. Agarose gel (1.5%) stained with ethidium bromide, showing the RT-PCR and PCR results using the RNA and DNA of *S. polychroa* as a source. Lanes **1** and **3** show amplification of a type II gene segment (3F-polyR) from RNA and from DNA, respectively. Lane **2** shows that the attempt to amplify the ITS-1 molecule from the *S. polychroa* RNA source (previously treated with DNase) was unsuccessful. Lanes **4** and **5** are the negative controls for the two pairs of oligonucleotides used. The molecular size marker is the 100-bp ladder.

plication and the speciation events. The difference found between the relative rate of type I and of type II genes obtained from the duplication event (Table 5) and that obtained since *G. tigrina* speciation (2.3 and 3.4 times faster for type II, respectively) also seems to reinforce the possibility that the rate of change of type II sequences continued to increase long after the duplication event.

Type II rDNA Is Functional

To explore whether type II is functional, we amplified this sequence from RNA of *S. polychroa.* Figure 3 shows the result of this experiment. High molecular weight DNA and total RNA of *S. polychroa* were independently extracted from a pool of live animals. The RNA was treated with DNAse. The genomic DNA and the RNA were used as a template to PCR-amplify a fragment of the type II 18S sequence using a universal primer and a type II 18S rDNA-specific primer of *S. polychroa* (polyR: 5′-ATC TCA ATG AAG TGA CTT TC-3′). To check whether traces of "contaminating" genomic DNA were present in the RNA sample, PCR of the internal transcribed spacer-1 (ITS-1; approximately 600 bp) using the primers 9F and ITSR was also carried out from the RNA. Lane 1 in Fig. 3 shows the result of the RT-PCR from RNA with the *S. polychroa* type II-specific primer. PCR direct sequencing of this band confirmed that it was type II. This result seems to indicate that the type II 18S gene is functional, in contrast to findings using the Northern blot technique (Carranza et al. 1996). Lane 3 in Fig. 3 shows the PCR result with the same primers as in lane 1 but using the genomic DNA as a template. PCR direct sequencing of the product also confirmed that it was of type II. Lane 2 (Fig. 3) shows no amplification when RT-PCR was performed to amplify the ITS-1 from the RNA source, indicating that no traces of genomic DNA were present in the RNA sample. Lanes 4 and 5 (Fig. 3) are the negative controls for the type II fragment and the ITS-1. This does not support the hypothesis of Buckler et al. (1997) that the *Schmidtea* type II gene might be a pseudogene from a recently inactivated array, based on information on the predicted free energy and the *K* value of both type I and type II sequences of *S. mediterranea* (Carranza et al. 1996).

However, this evidence for low-level transcription of type II rDNA does not prove processing, folding, and involvement in protein synthesis. Moreover, the evidence presented above indicates that type II is transcribed, but the fact that we cannot differentiate an 18S gene fragment amplified from rDNA vs rRNA (the ribosomal genes do not have introns) does not allow us to rule out the possibility of amplification of the 18S fragment from traces of genomic DNA even if it has been treated with DNAse and a negative control has been performed. To explore further the functionality of the type II gene, a secondary structure model was used to localize in the sequences some positions previously defined as functionally important for small-subunit rRNA in prokaryotes (Gutell et al. 1985). This allowed us to identify unambiguously in the planarian 18S rRNA sequence 15 positions defined as potential contacts to the 50S subunit (only 1 had a different nucleotide for type II sequences) and 6 positions that possibly act as tRNA or mRNA contacts, all of them conserved among type I and type II sequences (Table 6).

Nevertheless, to confirm the expression of the type II gene and to establish its function, it is necessary to analyze the expression of 18S rDNA by in situ hybridization

Table 6. Alignments of the six regions that contain conserved tRNA or mRNA contact sites[a]

	616–631	982–987	1273–1280	1667–1676	1732–1742	1886–1889
E. coli	CAGCAGCCGCG**G**TAA	GT**G**AAA	AT**G**CAACG	TGAAGTCG**G**A	ACACAC**C**GCCC	G**G**AA
Type I	CAGCAGCCGCG**G**TAA	GT**G**AAA	AC**T**CAACA	TGAACGAG**G**A	ACACAC**C**GCCC	T**G**AA
Type II	CAGCAGCCGCG**G**TAA	GT**G**AAA	AC**T**CAACA	TGAACGAG**G**A	ACACAC**C**GCCC	T**G**AA

[a] Each fragment corresponds to a single-stranded region in the secondary structure of the 18S rRNA; the positions suggested in *E. coli* to be possible tRNA and/or mRNA contacts are in boldface. The numbers correspond to the positions in the alignment available at the anonymous ftp site, porthos.bio.ub.es/pub/incoming/phylogeny/S2phylogeny.gde.

experiments in the intact organism as well as during development and regeneration. It is likely that type II is expressed in very low amounts in some minority tissues or cell types in intact nonregenerating organisms, which could be detected only with highly sensitive methods such as RT-PCR. Moreover, its potential roles in processes such as embryonic development and regeneration are also worth investigation.

Acknowledgments. We wish to thank our colleagues Iñaki Ruiz, Gonzalo Giribert, Miguel Angel Arnedo, Carles Ribera, Julio Rozas, Joan Balanyà, and Josep Maria Comerón for helpful comments and discussion. This study was supported by DGI-CYT Grant PB90-0477 and CIRIT Grant GRQ93-1044 to J.B. We also thank three anonymous referees for comments and corrections that clearly improved the manuscript. David Swofford kindly provided a prerelease version of PAUP* (version 4.0d55)

References

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

Arnheim N, Treco D, Taylor B, Eicher E (1982) Distribution of ribosomal gene length variants among mouse chromosomes. Proc Natl Acad Sci USA 79:4677–4680

Ball IR (1974) A contribution to the phylogeny and biogeography of the freshwater triclads (Platyhelminthes: Turbellaria). In: Riser NW, Morse MP (ed) The biology of the Turbellaria. McGraw–Hill, New York, pp 339–401

Ball IR (1981) The phyletic status of the Paludicola. Hydrobiologia 84:7–12

Bremer K (1994) Branch support and tree stability. Cladistics 10:295–304

Brown G, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J Mol 18;225–239

Buckler ES IV, Ippolito A, Holtsford TP (1997) The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics 145:821–832

Carranza S (1997) Taxonomia molecular mitjancant la sequenciacio del DNA ribosomic 18S. Aplicacio a l'origen I filogenia dels Platihelmints. PhD thesis, University of Barcelona, Barcelona

Carranza S, Giribet G, Ribera C, Baguñà J, Riutort M (1996) Evidence that two types of 18S rDNA coexist in the genome of *Dugesia* (*Schmidtea*) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida). Mol Biol Evol 13:824–832

Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294–5299

Coen ES, Dover GA (1983) Unequal exchanges and the coevolution of X and Y rDNA arrays in *Drosophila melanogaster*. Cell 33:849–855

Czelusniak J, Goodman M, Hewett-Emmett D, Weiss ML, Venta PJ, Tashian RE (1982) Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature 298:297–300

De Vries DE, Sluys R (1991) Phylogenetic relationships of the genus *Dugesia* (Platyhelminthes, Tricladida, Paludicola). J Zool London 223:103–116

Dover GA (1982) Molecular drive: A cohesive model of species evolution. Nature 299:111–117

Dover GA, Coen E (1981) Spring cleaning ribosomal DNA: A model for multigene evolution? Nature 290:731–732

Felsenstein J (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. J Mol Evol 17:368–376

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

Felsenstein J (1993) PHYLIP: Phylogeny inference package, Version 3.5c. University of Washington, Seattle

Friedrich M, Tautz D (1997) An episodic change of rDNA nucleotide substitution rate has occurred during the emergence of the insect order Diptera. Mol Biol Evol 14:644–653

García-Fernández J, Baguñà J, Saló E (1993) Genomic organization and expression of the planarian homeobox genes Dth-1 and Dth-2. Development 118:241–253

Goodman M (1981) Decoding the pattern of protein evolution. Prog Biophys Mol Biol 38:105–164

Gunderson JH, Sogin ML, Wollett G, et al. (1987) Structurally distin, stage-specific ribosomes occur in Plasmodium. Science 238:933–937

Gutell RR, Weibser B, Woese CR, Noller HF (1985) Comparative anatomy of 16S-like ribosomal RNA. Prog Nucleic Acid Res Mol Biol 32:155–216

Hillis DM, Dixon MT (1991) Ribosomal DNA: Molecular evolution and phylogenetic inference. Q Rev Biol 66:411–453

Hillis DM, Moritz C, Porter CA, Baker RJ (1991) Evidence for biased gene conversion in concerted evolution of ribosomal DNA. Science 251:308–310

Holmquist R (1983) Transitions and transversions in evolutionary descent: an approach to understanding. J Mol Evol 19:134–144

Katayama T, Wada H, Furuya H, Satoh N, Yamamoto M (1995) Phylogenetic position of the dicyemid Mesozoa inferred from 18S rDNA sequences. Biol Bull 189:81–90

Katayama T, Nishioka M, Yamamoto M (1996) Phylogenetic relationships among turbellarian orders inferred from 18S rDNA sequences. Zool Sci 13:747–756

Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Krystal M, D'Eustachio P, Ruddle FH, Arnheim N (1981) Human nucleolus organizers on nonhomologous chromosomes can share the same ribosomal gene variants. Proc Natl Acad Sci USA 78: 5744–5748

Kumar S, Tamura K, Nei M (1994) MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Comput Appl Biosci 10:189–191

Larson A, Wilson AC (1989) Patterns of ribosomal RNA evolution in salamanders. Mol Biol Evol 6:131–154

Lassner M, Dvorak J (1986) Preferential homogenization between adjacent and alternate subrepeats in wheat rDNA. Nucleic Acids Res 14:5499–5512

Li WH, Graur D (1991) Fundamentals of molecular evolution. Sinauer, Sunderland, MA

Li WH, Tanimura M (1987) The molecular clock runs more slowly in men than in apes and monkeys. Nature 326:93–96

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, pp 256–276

Ohta T (1991) Multigene families and the evolution of complexity. J Mol Evol 33:34–41

Ohta T (1994) Further examples of evolution by gene duplication revealed through DNA sequence comparisons. Genetics 138:1331–1337

Olsen GJ, Matsuda H, Hangstrom R, Overbeek R (1994) Fast DNAml: A tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput Appl Biosci 10:41–48

Pawlowsky J, Bolivar I, Fahrni JF, Vargas C de, Gouy M, Zaninetti L (1997) Extreme differences in rates of molecular evolution of foraminifera revealed by comparison of ribosomal DNA sequences and the fossil record. Mol Biol Evol 14(5):498–505

Philippe H (1993) Etude de la robustesse des phylogénies moléculaires. Université Paris-Sud. Centre d'Orsay, France. PhD thesis

Riutort M, Field KG, Turbeville JM, Raff RA, Baguñà J (1992) Enzyme electrophoresis, 18S rRNA sequences, and levels of phylogenetic resolution among several species of freshwater planarians (Platyhelminthes, Tricladida, Paludicola). Can J Zool 70:1425–1439

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

Sluys R (1989) Phylogenetic relationships of the triclads (Platyhelminthes, Seriata, Tricladida). Bijdr Dierk 59:3–25

Sluys R, Ribas M, Baguñà J (1995) Taxonomy, ecology and karyology of a new species of *Phagocata* from Spain, with a discussion of the phylogenetic systematics of the genus *Phagocata* s.1. (Platyhelminthes, Tricladida, Paludicola). Can J Zool 73:557–568

Sokal RR, Rohlf J (1994) Biometry: The principles and practice of statistics in biological research, 3rd ed. WH Freeman, New York

Spears T, Abele L, Kim W (1993) The monophyly of brauchyurans crabs: A phylogenetic study based on 18S rDNA. Syst Biol 41: 446–461

Spencer DF, Collings JC, Schnare MN, Gray MW (1987) Multiple spacer sequences in the nuclear large subunit ribosomal RNA gene of *Crithidia fasciculata*. EMBO J 6:1063–1071

Stothard JR, Frame IA, Miles MA (1997) Use of polymerase chain reaction-based single strand conformational polymorphism and denaturing gradient gel electrophoresis methods for detection of sequence variation of ribosomal DNA of *Trypanosoma cruzi*. Int J Parasitol 27(3):339–343

Sueoka N (1988) Directional mutation pressure and neutral molecular evolution. Proc Natl Acad Sci USA 85:2653–2657

Swofford DL (1998) PAUP*: Phylogenetic analysis using parsimony (*and other methods), Version 4.0. Sinauer, Sunderland, MA

Telford MJ, Holland PWH (1997) Evolution of 28S ribosomal DNA in chaetognaths: Duplicate genes and molecular phylogeny. J Mol Evol 44:135–144

VanCampenhout S, Aert R, Volckaert G (1998) Orthologous DNA sequence variation among 5S ribosomal RNA gene spacer sequences on homoeologous chromosomes 1B, 1D, and 1R of wheat and rye. Genome 41(2):244–255

Vawter L, Brown WM (1993) Rates and patterns of base change in the small subunit ribosomal RNA gene. Genetics 134:597–608

Wolfe KH, Sharp PM, Li W-H (1989) Mutation rates differ among regions of the mammalian genome. Nature 337:382–285

Wu CI, Li WH (1985) Evidence for higher rates of nucleotide substitution in rodents than in men. Proc Natl Acad Sci USA 82:1741–1745