Characterization of the *Gyrodactylus salaris* Malmberg, 1957 (Platyhelminthes: Monogenea) ribosomal intergenic spacer (IGS) DNA

C. M. COLLINS[1] *and* C. O. CUNNINGHAM[2]*

[1] *Department of Molecular and Cell Biology, University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 3TZ*
[2] *FRS Marine Laboratory, PO Box 101, Victoria Road, Aberdeen AB11 9DB, UK*

SUMMARY

The intergenic spacer of the ribosomal RNA gene array from the monogenean *Gyrodactylus salaris* was isolated using PCR amplification. PCR products were cloned and sequenced. Three different fragments of 0·63, 1·0 and 2·62 kb, were consistently obtained. These showed homology at the 5′ and 3′ termini but differed in their overall size and intervening sequence. The 5′ end showed homology to various 28S ribosomal RNA gene sequences, suggesting that this represented the 3′ terminus of the *G. salaris* 28S ribosomal RNA gene. A number of features common to other eukaryotic intergenic spacers were found in the longest sequence, including A+T rich sequences, palindromic sequences and tandemly repeated elements. Two regions of 23 bp sequences arranged in non-identical tandem repeats were identified. There were 9 repeats in both regions, separated by 81 bp of non-repetitive sequence. The repeat units from the two regions shared some similarity at their 3′ ends. The *G. salaris* intergenic spacer sequence was examined for sequence motifs involved in the transcription of the ribosomal RNA genes in other species. Several regions with homology to transcription start sites were identified.

Key words: *Gyrodactylus salaris*, ribosomal RNA, intergenic spacer (IGS), repetitive DNA.

INTRODUCTION

The genes encoding ribosomal RNA, the ribosomal DNA (rDNA), of all eukaryotes studied to date consist of coding regions for the 28S or large subunit ribosomal RNA (rRNA) gene, the 18S or small subunit rRNA gene, the 5·8S rRNA gene and the 5S rRNA gene. The genes are separated by regions of spacer DNA that include both transcribed and non-transcribed sequences (Raué *et al.* 1990). The 18S, 5·8S and 28S rRNA genes are separated by transcribed spacers. Transcribed spacers consist of internal transcribed spacers (ITS) 1 and 2, which lie between the 18S and 5·8S and 5.8S and 28S rRNA genes, respectively, and the external transcribed spacer which lies upstream of the small subunit rRNA gene. All of these spacer regions are cleaved from the primary RNA transcript and do not form part of the mature rRNA. The non-transcribed spacer lies between the 3′ end of the large subunit rRNA gene and the 5′ terminus of the external transcribed spacer and separates individual units of the rRNA genes, which are tandemly repeated along the chromosome. The term intergenic spacer (IGS) is given to the non-transcribed and external trans-

cribed spacers combined, i.e. the region between the large and small subunit ribosomal RNA genes separating individual units within the rRNA gene array.

In the platyhelminths, and indeed, in eukaryotes in general, the 18S rRNA gene has been the most intensively studied part of the ribosomal RNA gene array, providing useful information for species identification and differentiation and phylogenetic analysis e.g. (Johnston, Kane & Rollinson, 1993; Kralova *et al.* 1997; Littlewood, Rohde & Clough, 1999). The internal transcribed spacers have also proved effective for species discrimination (Kane & Rollinson, 1994; Gasser & Hoste, 1995). In contrast to the 18S rRNA gene, the platyhelminth intergenic spacers have been poorly studied and, at present, sequences for only *Echinococcus* and *Schistosoma* IGS have been reported (Silva, Rodrigues & Zaha, 1991; Kane & Rollinson, 1998). Characterization of the rRNA gene array in the monogenean genus *Gyrodactylus* began with sequencing of the 18S rRNA gene and its use in phylogenetic analysis and species discrimination (Cunningham, McGillivray & MacKenzie, 1995 *a, b*). The internal transcribed spacers have also provided sufficient sequence variation for species identification and phylogenetic analysis within the genus (Cunningham, 1997; Cable *et al.* 1999) and predictive models of the secondary structures of these sequences have pro-

* Corresponding author: FRS Marine Laboratory, PO Box 101, Victoria Road, Aberdeen AB11 9DB, UK. Tel: +01224 295 634. Fax: +01224 295 620. E-mail: c.cunningham@marlab.ac.uk

vided insights into post-transcriptional processing of the monogenean rRNA (Cunningham, Aliesky & Collins, 2000).

The coding regions of ribosomal RNA give rise to transcripts that form an integral part of the ribosomal complex, facilitating the translation of messenger RNA into protein. Transcribed spacers may play a role in orientating the coding regions properly during ribosome assembly, before being cleaved from the mature transcript. The function of non-transcribed spacers is less certain. Three types of regulatory elements have been identified within the intergenic spacers of mouse and *Xenopus* ribosomal RNA gene arrays. These are promoters, enhancer sequences that stimulate transcription by RNA polymerase I, and transcription termination sites. Similar elements have been found in other organisms (Reeder, 1992). The promoter site for RNA polymerase I is duplicated several times within each copy of the ribosomal non-transcribed spacer of *Xenopus* and mouse. In addition to this, repetitive sequences within the non-transcribed spacer of these species have been shown to play a role in transcription enhancement (Labhart & Reeder, 1984; Pikaard *et al*. 1990). *Drosophila* has repetitive elements within the intergenic spacer that play a dual role of promoter and enhancer (Grimaldi, Fiorentini & Di Nocera, 1990).

This paper presents the first characterization of the rDNA intergenic spacer and provides the first step in defining ribosomal transcription within a monogenean species. The sequence containing the putative rRNA transcription initiation and termination sites is presented and features within the intergenic spacer of *G. salaris* that are present in the spacers of other organisms are identified.

MATERIALS AND METHODS

Preparation of Gyrodactylus *DNA*

Approximately 200 *G. salaris* specimens were obtained from Atlantic salmon (*Salmo salar* L.) parr collected from the river Rauma in Norway and preserved in 70% (v/v) ethanol. The animals were placed in 500 μl of lysis buffer (Tris 0·1 M, pH 8·5, NaCl 0·2 M, EDTA 0·05 M, SDS 1%, proteinase K 60 μg/ml) and incubated at 65 °C for 3·5 h, followed by inactivation of residual proteinase K at 95 °C for 10 min. Total genomic DNA was separated from tissue proteins and lipids using phenol and chloroform extraction, ethanol precipitated, and resuspended in TE buffer (Tris–HCl 10 mM, EDTA 1 mM, pH 8·0). DNA concentration was estimated by subjecting an aliquot to electrophoresis in a 0·8% (w/v) agarose gel alongside markers of known DNA concentration (Low DNA Mass Ladder, Gibco–BRL).

Amplification of the ribosomal intergenic spacer

The intergenic spacer (IGS) region of the ribosomal RNA gene array (rDNA) was amplified using primers 3′10 (5′-GCTGATTTAATGAGCC-3′) which is complementary to sequence at the 5′ terminus of the *G. salaris* small subunit (18S) ribosomal RNA gene (Cunningham *et al*. 1995*a*), and 28AAF (5′-AGGTTAGTTTTACCCTACT-3′) which lies at the 3′ terminus of the large subunit (28S) ribosomal RNA gene (Hillis & Dixon, 1991) and was found to be conserved in *Schistosoma* (Kane & Rollinson, 1998). A sample of 100 ng of genomic DNA was used as template in a total amplification volume of 100 μl. The amplification reaction contained primers 1 μM each, dNTPs 0·25 mM each, MgCl$_2$ 1·0 mM, AmpliTaq Gold Buffer 1 × (Perkin Elmer), and distilled H$_2$O. Amplification was carried out on a Techne Genius thermocycler using conditions of 96 °C for 10 min, at which point 2·6 U AmpliTaq Gold enzyme mix (Perkin Elmer) were added to the reaction mix, followed by 35 cycles of 94 °C for 1 min, 50 °C for 1 min, 72 °C for 2·5 min, and a final incubation at 72 °C for 5 min. Results of the amplification were checked on a 0·8% agarose gel.

PCR amplifications were also carried out using single *Gyrodactylus* specimens, incubated under the same conditions as above, in 7·5 μl of lysis buffer (proteinase K 60 μg/ml, NP40 0·45%, Tween 20 0·45% in TE buffer). The total lysate was added to the amplification mix.

Cloning and sequencing of the intergenic spacer

PCR products were excised from agarose gel, purified using the Geneclean III Kit (Bio 101), and cloned using the pGEM-T Vector System (Promega). Sequencing reactions used the dRhodamine Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer) and an ABI Prism 377 automatic sequencer. Full-length sequences were obtained from plasmid purified by resin column (Promega) and excised inserts that had been gel purified, using plasmid and internal primers. In addition, sequences of the region containing repetitive DNA were obtained directly from PCR products.

Sequence analysis

Sequences were examined and contigs assembled using Sequencher (Gene Codes Corporation). BLAST searches were carried out to identify sequences from other organisms with homology to the *G. salaris* sequence. Alignment of sequences for interspecies comparison of IGS regions was carried out using ClustalX (Thompson *et al*. 1997).
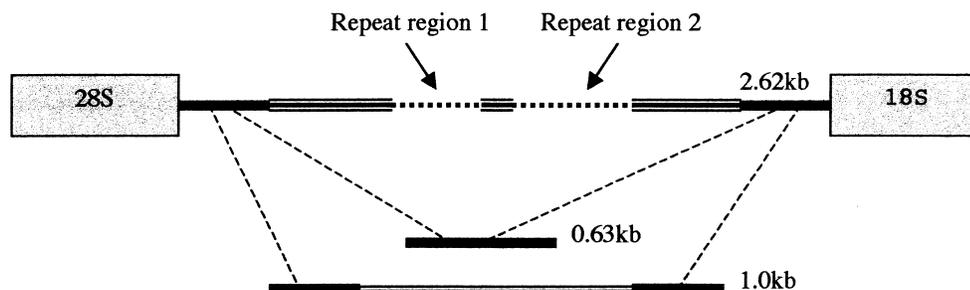
Fig. 1. PCR products (2·62, 1·0 and 0·63 kb) obtained following amplification of *Gyrodactylus salaris* intergenic spacer. All 3 products contain identical sequences at the 5′ and 3′ termini (solid line). The 1·0 kb and 2·62 kb products contain different sequences between these termini (parallel line) and the 2·62 kb product contains 2 regions of repetitive DNA (dashed lines).

RESULTS

PCR amplification of the intergenic spacer

Three PCR products were obtained using primers 3′10 and 28AAF. These were 2·62, 1·0 and 0·63 kb in size. Based on intensity of ethidium bromide staining in gels, the 1·0 kb product appeared to be amplified at a lower concentration than the other fragments. These 3 products were always obtained, despite modifications of parameters in the amplification reaction. Changes in the concentrations of DNA template, MgCl₂ and dNTPs and changes in duration and temperature of annealing and extension steps did not result in single products. Three products were obtained in both reactions using pooled DNA and those with single parasites as template.

Intergenic spacer sequences

The intergenic spacer (IGS) PCR products are shown schematically in Fig. 1. The 0·63 kb IGS PCR product contains the first 229 nucleotides of the 5′ end and the last 400 nucleotides of the 3′ end of the 2·62 kb IGS sequence with the intervening sequence missing. The 1·0 kb IGS PCR product is identical to the first 235 and the last 485 nucleotides of the 2·62 kb IGS sequence. The remaining intervening sequence shows no homology to the 2·62 kb intergenic spacer. This entire 1·0 kb sequence has been submitted to the EMBL nucleotide sequence database under accession number AJ276033.

The 5′ terminus of each product from nucleotides 19 to 159, shows homology to ribosomal sequences from several species, including 90 % similarity with *Schistosoma* ribosomal intergenic spacer region (EMBL accession numbers AJ223840, AJ223838, AJ223842) and *Spirometra erinacei* 28S ribosomal DNA (EMBL accession number AB027760), 80 % similarity with mouse 28S ribosomal DNA (EMBL accession number AV355665) and human 28S ribosomal DNA (EMBL accession number M11167).

The 2·62 kb IGS PCR product contains 2 regions of tandem repeats (see Fig. 1), separated from each other by a non-repeated sequence of 81 bp. The first region of repeats extends from nucleotides 1381 to 1587 and consists of 9 repeat units, each 23 bp in length. These repeats are not identical but differ from each other at 1–3 positions (Fig. 2A). The second set of tandem repeats starts at nucleotide 1669 and ends at nucleotide 1875. This array also consists of 9 repeats, each 23 bp in length and differing from each other in 3–4 positions (Fig. 2B). The first and second repeat arrays have a high degree of similarity in 12 nucleotides at the 3′ ends of their individual repeats (Fig. 2). The same sequence was obtained from clones of the entire IGS amplified from pooled DNA and from PCR products spanning only the region of repetitive DNA that had been amplified from single specimens.

Immediately upstream of the first array of repeats, the sequence (5′-GTACA-3′) occurs. This comprises the last four nucleotides of the 23 bp repeat unit in this first array. Another truncated repeat (5′-CACTATTACCGTGG-3′) which corresponds to the 5′ portion of the 23 bp unit found within the second repeat array occurs downstream of this array (Fig. 3).

Three palindromic sequences were found in the 2·62 kb IGS sequence at nucleotides 138 to 150 (5′ CTAAGTCTGAATC 3′), 515 to 528 (5′ CTGGT-ACCATGGTC 3′) and 2374 to 2388 (5′ TCCTT-AATAATTCCT 3′). An inverted repeat was found at nucleotide 1144 (5′ AGCTCACTATAGTGAG-CT 3′). Three chi-like sequences occur at positions 828 (5′ ACTGGTGG 3′), 971 (5′ GCTGGTGC 3′) and 2353 (5′ GTTGGTGG 3′).

The overall composition of the 2·62 kb IGS sequence is 56·8 % AT. The occurrence of AT and GC base pairing shows a distinct distribution with areas of high AT content flanking the regions of tandem repeats and at the 3′ end of the IGS (Fig. 4). The 2·62 kb sequence has been submitted to the EMBL nucleotide sequence database under accession number AJ276032.

The species most closely related to *Gyrodactylus* for which intergenic spacer sequences were found in nucleotide sequence databases were *Echinococcus*

First repeat section

Repeat

A	G T C C T T C A G T	G T A G A A C C G T A C	A
B	G T C C T T C A G T	G T A G A G C C G T A C	A
E	G T C C T T T A G T	G T A G A G C C G T A C	A
F	G T C A T T C A G G	G T A G A G C C G T A C	A

(a)

Second repeat section

Repeat

P	T A C T A A T A C C	G T G T A G C C G T A G	G
Q	T A T T A T T A C C	G T A G A G C C G T A C	G
R	C A C T A T T A C C	G T G G A G C C G T A G	G
T	T A C T T A T A C C	G T A G A G C C G T A C	G
V	T A C T T T T A C C	G T G A A G C C G T A G	G

(b)

Fig. 2. Differences in sequences between tandem repeats within the intergenic spacer region of *Gyrodactylus salaris*. Nucleotides which vary from the consensus are shaded. Regions of homology between repeats are boxed.

```
TTCCAAGTTCACATAGCATTCGTACA|GTCCTTCAGTGTAGAACCGTACA|GTCCTTCAGTGTAGAGCCGTA
                                      A                           B

CA|GTCCTTCAGTGTAGAGCCGTACA|GTCCTTCAGTGTAGAGCCGTACA|GTCCTTCAGTGTAGAGCCGTA
        B                           B                           B

CA|GTCCTTCAGTGTAGAGCCGTACA|GTCCTTCAGTGTAGAGCCGTACA|GTCATTCAGGGTAGAGCCGTA
        B                           B                           F

CA|GTCCTTTAGTGTAGAGCCGTACA|CACTAACACTGACGTACGCACTAACGCCTATACTTTGAACAAATT
        E                          Non-repetitive region

TGGAGTGCTCTAAAATTTTTTATCACGAGACGAAGG|TACTAATACCGTGTAGCCGTAGG|TACTAATACCG
                                                 P

TGTAGCCGTAGG|TACTAATACCGTGTAGCCGTAGG|TATTATTACCGTAGAGCCGTACG|CACTATTACCG
     P                     P                     Q

TGGAGCCGTAGG|TACTTATACCGTAGAGCCGTACG|CACTATTACCGTGGAGCCGTAGG|TACTTTTACCG
     R                     T                     R

TGAAGCCGTAGG|TATTATTACCGTAGAGCCGTACG|CACTATTACCGTGGTAACTAAGATTTTCACACCG
     V                     Q
```

Fig. 3. Sequences of repeat regions 1 and 2 within the intergenic spacer of *Gyrodactylus salaris* ribosomal RNA genes. Letters underneath sequences indicate repeat sequences denoted in Fig. 2.

granulosus (EMBL accession number U26429), *Schistosoma intercalatum* (EMBL accession number AJ223840), *S. haematobium* (EMBL accession num-ber AJ223838), and *S. mansoni* (EMBL accession number AJ223842). These sequences were aligned with the *G. salaris* IGS sequence and examined for
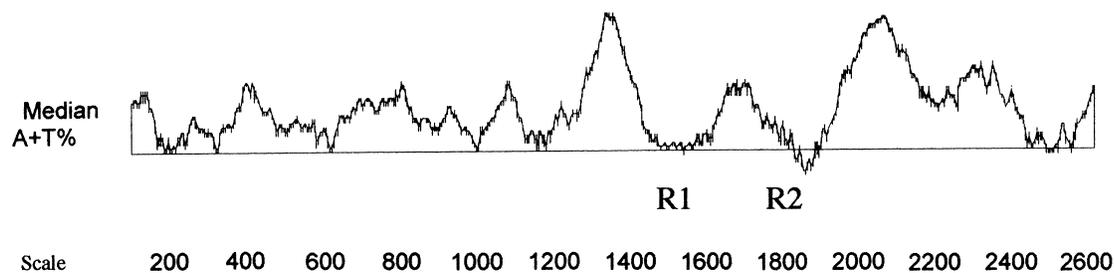
Fig. 4. AT content in the *Gyrodactylus salaris* intergenic spacer DNA. Areas marked R1 and R2 indicate positions of repetitive DNA regions 1 and 2.

common motifs. Apart from sequence representing the 3′ end of the 28S gene, no significant similarities were seen between the intergenic spacer of *G. salaris* and those of the other species.

The IGS sequences obtained for *G. salaris* were examined for the presence of sequence motifs found in other organisms. A summary of results is given in Table 1.

DISCUSSION

The intergenic spacer of *G. salaris* shows similarities to those of other organisms in that it has variable length and contains tandem repeats and long stretches of AT-rich sequences.

Three different PCR products were consistently found on amplification of the intergenic spacer of *G. salaris*. Similar results have been seen in other organisms. Kane & Rollinson (1998) reported different sized spacer regions on amplification of the *Schistosoma* IGS. One of these products had a large deletion of approximately 1·5 kb, resembling the difference between the 2·62 and 0·63 kb PCR products found in *G. salaris*. Spacer length heterogeneity occurs in *Leishmania*, *Drosophila*, mouse and human, with a 3-fold range in sizes occurring in *Xenopus* (Long & Dawid, 1980). Very often the length heterogeneity in spacers is due to variation in repetitive elements within the spacer but in *Gyrodactylus* it appears to be due to large deletions, including the region containing both the regions of repetitive sequence.

The origin of the smaller IGS fragments is unclear. It may be that they are non-functional, in that they have lost their promoter and transcription start sites, as was suggested for the truncated intergenic spacers found in *S. haematobium* (Kane & Rollinson, 1998). It has been found in *Drosophila* and *Xenopus* that elimination of half the promoter sites in rDNA does not result in an appreciable reduction in rDNA transcription and ribosomal activity (Fedoroff, 1979). Therefore it is possible that non-functional spacers may be tolerated and allowed to accumulate within ribosomal gene arrays. Several spacers may be present if there are processes causing elimination and accumulation of spacers that

exist in some form of equilibrium (Fedoroff, 1979). However, as discussed below, the processes of concerted evolution appear to act in *G. salaris* as in other organisms, and it therefore seems unlikely that there would be the rapid gain or loss of spacers necessary to give rise to such disparate PCR products. If more than 1 cluster of rRNA genes occurs, perhaps even on different chromosomes, it may be the case that different intergenic spacers are found in each cluster while the genes and internal transcribed spacer sequences are maintained.

Determining the site of 28S rRNA gene transcription termination in the IGS is difficult. Comparison of all 3 IGS PCR fragments obtained from *G. salaris* with the IGS of schistosomes shows high homology for the first 159 nt of the IGS sequence. This most likely represents the 3′ end of the coding region for the 28S ribosomal gene, as the coding regions of the termini of *Gyrodactylus* ribosomal RNA genes have been shown to be highly conserved when compared with a wide range of divergent species (Cunningham *et al.* 2000). However, homology is significantly lower in the area defined as the 3′ terminus of the *Schistosoma* 28S gene by Kane & Rollinson (1998). Further evidence will be required to precisely identify the terminus of this gene in *Gyrodactylus*.

It is not known if a 3′ trailer is transcribed which is later cleaved from the mature *Gyrodactylus* 28S product. A putative termination sequence, TTTT, is found at position 447 in the 5′ end of the *G. salaris* IGS sequence. T clusters (on non-coding strands) and palindromic sequences are regularly found at transcription termination sites in both eukaryotes and prokaryotes (Fedoroff, 1979). Inverse termination signals (AAAA) and palindromic sequences occur between nucleotides 138 and 515. As position 138 falls within the putative 28S gene sequence, transcription termination is likely to occur within the region of nucleotides 284–515. Sequencing the IGS from other closely related gyrodactylid species may highlight similarities in sequence downstream of the 28S gene, improving the delineation of this gene. However, it may prove necessary to study cDNA reverse transcribed from *Gyrodactylus* RNA to unequivocally pinpoint the terminus of the 28S rRNA.

Table 1. Regions of sequence similarity between the *Gyrodactylus salaris* intergenic spacer (IGS) and functional motifs within the IGS of other organisms

Motif	Sequence (5′–3′)	*G. salaris* sequence (5′–3′)	Similarity (%)	Position in *G. salaris* IGS
Schistosoma chi-like*1	GCTGGTGG	ACTGGTGG	88	828–836
	CGACCACC	GCTGGTGC	88	971–978
		GTTGGTGG	88	2377–2384
Transcription termination signal[2]	TTTT	AAAA	100	285–288
		AAAA	100	332–335
		TTTT	100	447–450
Human transcription factor UBF1 binding site[3]	CTCCGAGTCG	CTCAGGGTCG	80	934–944
		AGGGTCGACGG	73	937–948
		CTTGCGGCCTA	73	1222–1232
		CACCGAGTCA	80	1929–1938
		TGGGCCGACTA	73	2304–2314
Xenopus laevis rDNA transcription start[4]	AGGTAGGGA (−4 to +5)	TCACTACTT	78	2256–2264
		AGGTAGCGA	89	2594–2601
Dictyostelium rDNA transcription start[4]	ATATACAAG (−4 to +5)	GAATACAAG	78	2006–2014
		CTAGTAAAT	78	2023–2031
		AAATACTAG	78	2092–2099
		CTTGTATGG	78	2166–2173
		CATCTATAT	78	2349–2356
		CCTGTAGAT	78	2410–2418
Saccharomyces rDNA transcription start[5]	AGGAACTTCA-TGCGAAAGC (−8 to +10)	TCTTCCGGTTT-GAGTTTCT	63	2121–2140
Mouse rDNA transcription start[4]	AGGTACTGA (−4 to +5)	ACGCACTAA	70	1625–1635
		AGGTACTAA	80	1688–1697
		AGGTACTAA	90	1712–1721
		AGGTACTAA	90	1735–1744
		AGGTACTTA	90	1757–1767
		AGGTACTTA	90	1781–1790
		ASGCACTATT	70	1804–1812
		AGGTACTTT	80	1827–1836
		AGGTAYTWT	80	1850–1858
		ASGYAYTAT	80	1873–1882
		GAGTATCTT	70	1940–1950
		AAGTAGTAA	70	2066–2075
		CACTACTTG	70	2256–2265
		CAGAACCGA	70	2528–2537
		AGGTAGCGA	70	2593–2602
Human rDNA transcription start[4]	ATATGCTGA (−4 to +5)	ATATTGTGA	78	2354–2362
Bombyx rDNA transcription start[4]	GTGAATACA (−4 to +5)	ACGTACGCA	78	1623–1631
		GTGAATACA	78	2005–2013
Drosophila melanogaster rDNA transcription start[4]	CTATAGGTA (−4 to +5)	CCTTGTATG	78	2164–2173
		CATTGAAAG	78	2586–2595
Autonomously replicating sequences of yeast[6]	A/TAAAT/CATAAAAT	TAATTATTAAA	82	2316–2327
		TTAATATAACA	82	2452–2462

* 1. Kane & Rollinson (1998); 2. Fedoroff (1979); 3. Lewin (1997); 4. Fujiwara & Ishikawa (1987); 5. Sommerville (1984); 6. Safrany & Hidvegi (1989).

The most striking feature of the IGS sequence reported here is the presence of 2 regions where 23 bp sequences are tandemly repeated. Repetitive elements have been described in the IGS of ribosomal genes from numerous species, sometimes as several large repeats, sometimes interspersed with non-repeat regions (Challoner *et al.* 1985; Kane & Rollinson, 1998) or in the form of long runs of simple sequence (Kane & Rollinson, 1998). Blocks of tandemly repeating units have been described in organisms as diverse as *Acanthamoeba* (Yang, Zwick & Paule, 1994), mouse (Kuhn, Deppert & Grummt, 1990; Pikaard *et al.* 1990), *Xenopus* (Labhart & Reeder, 1984) and plant species (Reeder, 1992).

In repetitive sequence, such as the rRNA gene array, the processes of concerted evolution act to

homogenize the array, resulting in conservation of sequence within species and greater divergence between species (Challoner *et al*. 1985; Elder & Turner, 1995). The sequences of the 5·8S rRNA gene, the V4 region of the small subunit rRNA gene, and the internal transcribed spacer region of the rRNA gene array are highly conserved within *G. salaris*, *G. derjavini* and *G. truttae*, even in widely dispersed populations (unpublished), demonstrating that homogenization has occurred in this genus.

Homogenization is facilitated by frequent recombination and unequal crossing over that are promoted at regions of, or surrounding, repetitive DNA (Challoner *et al*. 1985). This can lead to gain or loss of repeats, including those within the IGS (Jemtland *et al*. 1986; Safrany & Hidvegi, 1989). Sequences of the repeat regions of *G. salaris* IGS from different individuals within the same population show no variation in the number and sequences of repeats. Further analysis will be carried out to examine variation in the IGS repeat arrays between individuals from geographically distinct populations of *G. salaris* to investigate if the process of homogenization applies to all regions of the rRNA equally, or if different patterns occur in the transcribed and intergenic sequences, as demonstrated for *Drosophila* (Polanco *et al*. 1998).

AT-rich sequences occur in the *G. salaris* IGS as in other species, including *Caenorhabditis elegans* (Ellis, Sulston & Coulson, 1986) and *Schistosoma* (Kane & Rollinson, 1998). AT-rich regions can also promote recombination by acting as DNA unwinding elements (DUEs) (Sinden, 1994) and aiding binding of enzymes involved in recombination (Safrany & Hidvegi, 1989). The distribution of AT-rich sequence in the *G. salaris* IGS shows a distinct pattern in relation to the tandem repeats, occurring before and after each run of repeats.

Three chi-like sequence motifs exist within the IGS of *G. salaris*. The octamer chi sequence facilitates recombination in *E. coli* (Lewin, 1997). Chi-like sites were also found in the IGS of *Schistosoma* where it was suggested that they may have been involved in promoting recombination (Kane & Rollinson, 1998). Only 1 chi-like sequence falls within a region of high AT content. As it is likely that recombination will fall within an area of high AT, 2 of these motifs may be coincidental.

A high percentage of A and T nucleotides is also found towards the 3′ end of the *G. salaris* intergenic spacer, possibly indicating the rRNA gene promoter region. AT-rich sequences have been noted surrounding promoter sites where again, a high AT content may aid unwinding of the DNA duplex prior to transcription (Gruendler *et al*. 1991).

Sequence motifs showing similarity to mouse rDNA transcription start sites are found throughout the *G. salaris* intergenic spacer. Four such motifs (1 inverse and complement) are found within the last 600 bases of the IGS. As 1 of these motifs occurs within the second repeated sequence, it is duplicated at constant intervals. This may be comparable to the situation seen within the intergenic spacers of *Xenopus* and *Drosophila* where the repetitive sequences may have arisen by duplication of the promoter. The entire promoter appears to be repeated in *Drosophila* while only 1 promoter domain has been repeated in *Xenopus* (Reeder, 1992). The mouse promoter site is also repeated (Kuhn & Grummt, 1987).

There are numerous sequence motifs found within the IGS of *G. salaris* that show similarity to sites involved in transcription start within the IGS of other organisms. However, no motif was found within *G. salaris* IGS with 100 % sequence similarity to those reported to date. This was not unexpected, as the intergenic spacer is highly variable except between the most closely related species, and even then the functional sites within these sequences do not show 100 % homology (Challoner *et al*. 1985). This is clearly demonstrated in studies where transcription of ribosomal DNA can only be achieved with polymerase and cell extract (transcription co-factors) from the same species (Grummt, Roth & Paule, 1982; Miesfeld & Arnheim, 1984).

Preliminary sequencing of the 3′ end of the IGS of a second gyrodactylid species, *G. derjavini*, shows high homology with the last 300 bases of *G. salaris* IGS. Upstream of this region, the sequences diverge significantly (T. Neilson, FRS Marine Laboratory, personal communication). This strongly indicates that the signals for gyrodactylid RNA polymerase I binding and transcription initiation lie within the last 300 nucleotides of the IGS sequence. Many motifs that share similarity with transcription factor binding sites and transcription start sites occur within this region. AT content, homology with the external transcribed spacer of another *Gyrodactylus* species and the occurrence of many motifs with homology to transcription start sites of other organisms all suggest that transcription starts at the end of the IGS, and downstream of the repeat regions. This sequence is identical in all 3 IGS PCR products, and raises the possibility that all 3 variants of IGS may be able to initiate transcription.

In summary, the intergenic spacer of *G. salaris* shares many features with those of other organisms. However, at a more detailed level it shows little similarity to other organisms in its sequence, the position, number and size of repeats, and the conservation of motifs involved in transcription initiation and termination. Studies on transcriptional activation, or at least, further intergenic spacer sequences obtained from closely related gyrodactylid species, should enable better definition of the functional elements of the IGS spacer within *Gyrodactylus*.

REFERENCES

CABLE, J., HARRIS, P. D., TINSLEY, R. C. & LAZARUS, C. M. (1999). Phylogenetic analysis of *Gyrodactylus* spp. (Platyhelminthes: Monogenea) using ribosomal DNA sequences. *Canadian Journal of Zoology* **77**, 1439–1449.

CHALLONER, P. B., AMIN, A. A., PEARLMAN, R. E. & BLACKBURN, E. H. (1985). Conserved arrangement of repeated DNA sequences in non-transcribed spacers of ciliate ribosomal RNA genes: evidence for molecular coevolution. *Nucleic Acids Research* **13**, 2661–2680.

CUNNINGHAM, C. O. (1997). Species variation within the internal transcribed spacer (ITS) region of *Gyrodactylus* (Monogenea; Gyrodactylidae) ribosomal RNA genes. *Journal of Parasitology* **83**, 215–219.

CUNNINGHAM, C. O., ALIESKY, H. & COLLINS, C. M. (2000). Sequence and secondary structure variation in the *Gyrodactylus* (Platyhelminthes: Monogenea) ribosomal RNA gene array. *Journal of Parasitology* **86**, 567–576.

CUNNINGHAM, C. O., McGILLIVRAY, D. M. & MacKENZIE, K. (1995*a*). Phylogenetic analysis of *Gyrodactylus salaris* Malmberg, 1957 based on the small subunit (18S) ribosomal RNA gene. *Molecular and Biochemical Parasitology* **71**, 139–142.

CUNNINGHAM, C. O., McGILLIVRAY, D. M., MacKENZIE, K. & MELVIN, W. T. (1995*b*). Discrimination between *Gyrodactylus salaris*, *G. derjavini* and *G. truttae* (Platyhelminthes: Monogenea) using restriction fragment length polymorphisms and an oligonucleotide probe within the small subunit ribosomal RNA gene. *Parasitology* **111**, 87–94.

ELDER, J. F. & TURNER, B. J. (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. *The Quarterly Review of Biology* **70**, 297–320.

ELLIS, R. E., SULSTON, J. E. & COULSON, A. R. (1986). The rDNA of *C. elegans*: sequence and structure. *Nucleic Acids Research* **14**, 2345–2364.

FEDOROFF, N. V. (1979). On spacers. *Cell* **16**, 697–710.

FUJIWARA, H. & ISHIKAWA, H. (1987). Structure of the *Bombyx mori* rDNA initiation site for its transcription. *Nucleic Acids Research* **15**, 1245–1258.

GASSER, R. B. & HOSTE, H. (1995). Genetic markers for closely-related parasitic nematodes. *Molecular and Cellular Probes* **9**, 315–320.

GRIMALDI, G., FIORENTINI, P. & DI NOCERA, P. P. (1990). Spacer promoters are orientation dependent activators or pre-rRNA transcription in *Drosophila melanogaster*. *Molecular and Cell Biology* **10**, 4667–4677.

GRUENDLER, P., UNFRIED, I., PASCHER, K. & SCHWEIZER, D. (1991). rDNA intergenic region from *Arabidopsis thaliana*. Structural analysis, intraspecific variation and functional implications. *Journal of Molecular Biology* **221**, 1209–1222.

GRUMMT, I., ROTH, E. & PAULE, M. R. (1982). Ribosomal RNA transcription *in vitro* is species specific. *Nature, London* **296**, 173–174.

HILLIS, D. M. & DIXON, M. T. (1991). Ribosomal DNA: Molecular evolution and phylogenetic inference. *Quarterly Review of Biology* **66**, 411–453.

JEMTLAND, R., MAEHLUM, E., GABRIELSEN, O. S. & ØYEN, T. B. (1986). Regular distribution of length heterogeneities within non-transcribed spacer regions of cloned and genomic rDNA of *Saccharomyces cerevisiae*. *Nucleic Acids Research* **14**, 5145–5158.

JOHNSTON, D. J., KANE, R. A. & ROLLINSON, D. (1993). Small subunit (18S) ribosomal RNA gene divergence in the genus *Schistosoma*. *Parasitology* **107**, 147–156.

KANE, R. A. & ROLLINSON, D. (1994). Variation in the ribosomal DNA internal transcribed spacer regions of *Schistosoma haematobium*, *S. intercalatum* and *S. mattheei* is due to repeating elements. *Molecular and Biochemical Parasitology* **63**, 153–156.

KANE, R. A. & ROLLINSON, D. (1998). Comparison of the intergenic spacers and 3′ end regions of the large subunit (28S) ribosomal RNA gene from three species of *Schistosoma*. *Parasitology* **117**, 235–242.

KUHN, A. & GRUMMT, I. (1987). A novel promoter in the mouse rDNA spacer is active *in vivo* and *in vitro*. *EMBO Journal* **6**, 3487–3492.

KUHN, A., DEPPERT, U. & GRUMMT, I. (1990). A 140 bp repetitive sequence element in the mouse ribosomal gene spacer enhances transcription by RNA polymerase I in a cell-free system. *Proceedings of the National Academy of Sciences, USA* **87**, 7527–7531.

KRALOVA, I., VAN DE PEER, Y., JIRKU, M., VAN RANST, M., SCHOLZ, T. & LUKES, J. (1997). Phylogenetic analysis of a fish tapeworm, *Protecephalus exigus*, based on the small subunit rRNA gene. *Molecular and Biochemical Parasitology* **84**, 263–266.

LABHART, P. & REEDER, R. H. (1984). Enhancer-like properties of the 60/81 bp elements in the ribosomal gene spacer of *Xenopus laevis*. *Cell* **37**, 285–289.

LEWIN, B. (1997). *GenesVI*. Oxford University Press, New York.

LITTLEWOOD, D. T. J., ROHDE, K. & CLOUGH, K. A. (1999). The interrelationships of all major groups of Platyhelminthes: phylogenetic evidence from morphology and molecules. *Biological Journal of the Linnean Society* **66**, 75–114.

LONG, E. O. & DAWID, I. B. (1980). Repeated genes in eukaryotes. *Annual Review of Biochemistry* **49**, 727–764.

MIESFELD, R. & ARNHEIM, N. (1984). Species-specific rDNA transcription is due to promoter-specific binding factors. *Molecular and Cell Biology* **4**, 221–227.

PIKAARD, C. S., PAPE, L. K., HENDERSON, S. L., RYAN, K., PAALMAN, M. H., LOPATA, M. A., REEDER, R. H. & SOLLNER-WEBB, B. (1990). Enhancers for RNA Polymerase I in mouse ribosomal DNA. *Molecular and Cell Biology* **10**, 4816–4825.

POLANCO, C., GONZALEZ, A. I., DE LA FUENTE, A. & DOVER, G. A. (1998). Multigene family of ribosomal DNA in *Drosophila melanogaster* reveals contrasting patterns of homogenization for IGS and ITS spacer regions: a possible mechanism to resolve this paradox. *Genetics* **149**, 243–256.

RAUÉ, H. A., MUSTERS, W., RUTGERS, C. A., RIET, J. V. T. &

PLANTA, R. J. (1990). rRNA: from structure to function. In *The Ribosome: Structure, Function and Evolution*, (ed. Hill, W. E., Dahlberg, A., Garrett, R. A., Moore, P. B., Sclessinger, D. & Warner, J. R.), pp. 217–235. American Society for Microbiology, Washington D. C.

REEDER, R. H. (1992). Regulation of transcription by RNA polymerase I. In *Transcriptional Regulation* Vol. 1, (ed. McKnight, S. L. & Yamamoto, K. R.), pp. 315–347. Cold Spring Harbor Laboratory Press, New York.

SAFRANY, G. & HIDVEGI, E. J. (1989). New tandem repeat region in the non-transcribed spacer of human ribosomal RNA gene. *Nucleic Acids Research* **17**, 3013–3022.

SILVA, S. C., RODRIGUES, J. J. S. & ZAHA, A. (1991). Molecular cloning and characterization of the ribosomal RNA genes of the cestode *Echinococcus granulosus. Brazilian Journal of Medical and Biological Research* **24**, 345–357.

SINDEN, R. R. (1994). *DNA Structure and Function*. Academic Press, London.

SOMMERVILLE, J. (1984). RNA Polymerase I promoter and transcription factors. *Nature, London* **310**, 189–190.

THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. & HIGGINS, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876–4882.

YANG, Q., ZWICK, M. G. & PAULE, M. R. (1994). Sequence organisation of the *Acanthamoeba* rRNA intergenic spacer: identification of transcriptional enhancers. *Nucleic Acids Research* **22**, 4798–4805.