



The Structure of a Single Unit of Ribosomal RNA Gene (rDNA) Including Intergenic Subrepeats in the Australian Bulldog Ant *Myrmecia croslandi* (Hymenoptera: Formicidae)

Authors: Ohnishi, Hitoshi, and Yamamoto, Masa-Toshi

Source: Zoological Science, 21(2) : 139-146

Published By: Zoological Society of Japan

URL: <https://doi.org/10.2108/zsj.21.139>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

The Structure of a Single Unit of Ribosomal RNA Gene (rDNA) Including Intergenic Subrepeats in the Australian Bulldog Ant *Myrmecia croslandi* (Hymenoptera: Formicidae)

Hitoshi Ohnishi and Masa-Toshi Yamamoto*

Drosophila Genetic Resource Center, Kyoto Institute of Technology, Saga-Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan

ABSTRACT—A complete single unit of a ribosomal RNA gene (rDNA) of *M. croslandi* was sequenced. The ends of the 18S, 5.8S and 28S rRNA genes were determined by using the sequences of *D. melanogaster* rDNAs as references. Each of the tandemly repeated rDNA units consists of coding and non-coding regions whose arrangement is the same as that of *D. melanogaster* rDNA. The intergenic spacer (IGS) contains, as in other species, a region with subrepeats, of which the sequences are different from those previously reported in other insect species. The length of IGSs was estimated to be 7–12 kb by genomic Southern hybridization, showing that an rDNA repeating unit of *M. croslandi* is 14–19 kb-long. The sequences of the coding regions are highly conserved, whereas IGS and ITS (internal transcribed spacer) sequences are not. We obtained clones with insertions of various sizes of R2 elements, the target sequence of which was found in the 28S rRNA coding region. A short segment in the IGS that follows the 3' end of the 28S rRNA gene was predicted to form a secondary structure with long stems.

Key words: rDNA, IGS, R2, repeating unit, Hymenoptera

INTRODUCTION

The Australian bulldog ant *Myrmecia croslandi* is one of the better studied species of ants, and has the lowest chromosome number ($2n=2$) known in higher organisms. From this large metacentric chromosome, two smaller chromosomes were believed to be produced by a centric fission (Hirai *et al.*, 1996, and the references therein). These species exhibit chromosome number polymorphism, with $2n=2$, 3 and 4 (Crosland and Crozier, 1986; Taylor, 1991). *M. croslandi* belongs to the Myrmeciinae, which is thought to be an ancestral group of ants (Hölldobler and Wilson, 1990; Ogata, 1991). The phylogenetic position of the Myrmeciinae within Formicidae, however, has not been clarified (Taylor, 1987; Baroni Urbani *et al.*, 1992).

The ribosomal RNA genes (rDNAs) are essential for all organisms and contain a mosaic of DNA sequences showing extremely different evolutionary rates (Pélandakis *et al.*, 1991). 28S rDNA of eukaryotes consists of twelve divergent

domains (D1–D12) (Hassouna *et al.*, 1984) that have been widely used in molecular systematics. A phylogenetic tree of hymenopteran insects constructed by using the D2 region and the mitochondrial 16S rDNA sequences is not compatible with one constructed based on morphological data (Schmitz and Moritz, 1998). The analysis of 1100-bp sequences of 18S and 28S rDNAs produced a polydivergent tree for the ant subfamily relationships (Sullender, 1998).

There have been only a few studies on the complete rDNA unit structure in insects: the structures studied were from *Drosophila melanogaster* (Diptera) (Tautz *et al.*, 1988) and *Acyrtosiphon pisum* (Hemiptera) (Amako *et al.*, 1996). In the present study, we determined the complete nucleotide sequence of a single unit of rDNA of *M. croslandi*. This is the third report of such a study in insects and the first in the Hymenoptera. We characterize the unit of rDNAs by comparing it with that of *D. melanogaster* rDNAs, which has been widely used as a general reference.

MATERIALS AND METHODS

Ant materials and rDNA clones

Myrmecia croslandi male pupae (sample no. IH89-31, with $2n=3$), a gift from Hirotami T. Imai, were stored at -80°C . pMc.r1 and pMc.r2 are plasmid DNA clones each carrying an *EcoR* I frag-

* Corresponding author: Tel. +81-75-873-2660;

Fax +81-75-861-0881.

E-mail: yamamoto@ipc.kit.ac.jp

Accession numbers: AB052895, AB121786, AB121787, AB121788, AB121789 and AB121790

ment of the rDNA of *M. croslandi* (sample no. IH89-030, with $2n=2$) as described in Hirai *et al.* (1994, 1996).

DNA extraction, genomic library and hybridization probes

High-molecular-mass genomic DNAs were extracted from 10 frozen male pupae following the procedure described by Sambrook *et al.* (1989). Since the DNA fragments containing the rDNA region of *M. croslandi* were shown to be between 4.5 and 15 kb in length by *Hind* III digestion (Hirai *et al.*, 1994), genomic DNA digested partially by *Hind* III was ligated into Lambda DASH II using the Ligation High kit (Toyobo). They were then packaged with Gigapack III (Stratagene) and introduced into *Escherichia coli* XL1-Blue cells as described (Sambrook *et al.*, 1989). Two probes were used for plaque filter hybridization: the 1.5 kb *Sal* I-*Eco*R I fragment (pHO-8) containing a portion of the 3' half of the coding region of the 28S rRNA gene, derived from pMc.r2, and the 0.5 kb PCR-amplified fragment (pMc.18S-2pcr) containing a portion of the 18S rRNA gene near the 5' end (see Fig. 1). The phage DNAs were isolated as described (Sambrook *et al.*, 1989). Three probes, pHO-10 (a subclone of pHO-8), pHO-8 and pMc.18S-2pcr, were used for hybridization analyses to examine structures of λ phage DNAs (see Fig. 1).

Polymerase chain reaction (PCR)

PCR amplification was conducted using the following primer pairs with Ex Taq polymerase (Takara), according to the manufacturer's protocol. The primer pair 18S-1U (5'-AGTAGTCATATGCT-TGTCTC-3') and 18S-1L (5'-AATCATTCATCGGTAGTAG-3') amplified most of the 18S rDNA (see Fig. 1, pMc.18S-1pcr). The primer pair 18S-1U and 18S-2L (5'-TGCTGCCCTTCCTTGATGTG-

3') amplified a quarter of the 18S rDNA (pMc.18S-2pcr) (see Fig. 1). For amplification of a portion of the IGS region including the 3'-most end of 28S rDNA and extending downstream (pMc.IGSpcr), the primer set HO-X-2(U) (5'-CCTGGCGGGGTGTTGTACTC-3') and IGS-2L (5'-ACCGAATATCAGAAGGAAAGAC-3') was used (see Fig. 1). The amplification conditions were 30 cycles of 1 min at 96°C, 1 min at 58°C, and 2 min at 74°C, following an initial denaturation at 96°C for 2 min. After electrophoresis, the amplified fragment was purified using SUPREC-01 (Takara) and inserted into pCR2.1 using a TA cloning kit (Invitrogen).

Southern hybridization

A half microgram each of ant genomic DNA was singly digested with *Eco*R I, *Hind* III and *Ban* III and separated by electrophoresis on a 1.0% agarose gel. The DNA fragments were transferred to a nylon membrane with 0.4 N NaOH and allowed to hybridize with a probe. The probe was a 0.3 kb *Ban* III-digested fragment, derived from the 4.0 kb *Hind* III fragment of λ Mc.8 (see Fig. 1), labeled with DIG-11-dUTP using a DNA labeling kit (Boehringer Mannheim). Hybridization was carried out in a solution of 5 \times SSC, 0.1% (w/v) N-lauroylsarcosine, 0.02% SDS and 1% (w/v) casein at 68°C overnight. Filters were washed twice in 2 \times SSC, 0.1% SDS at room temperature for 5 min and twice in 0.1 \times SSC, 0.1% SDS at 68°C for 15 min. Hybridization was detected with CDP-Star (Boehringer Mannheim) according to the manufacturer's protocol.

Subcloning and sequencing

Isolation of bacteriophage DNAs, subcloning of fragments into plasmids (pUC 118, 119 and pBluescript), transformation into *E. coli*

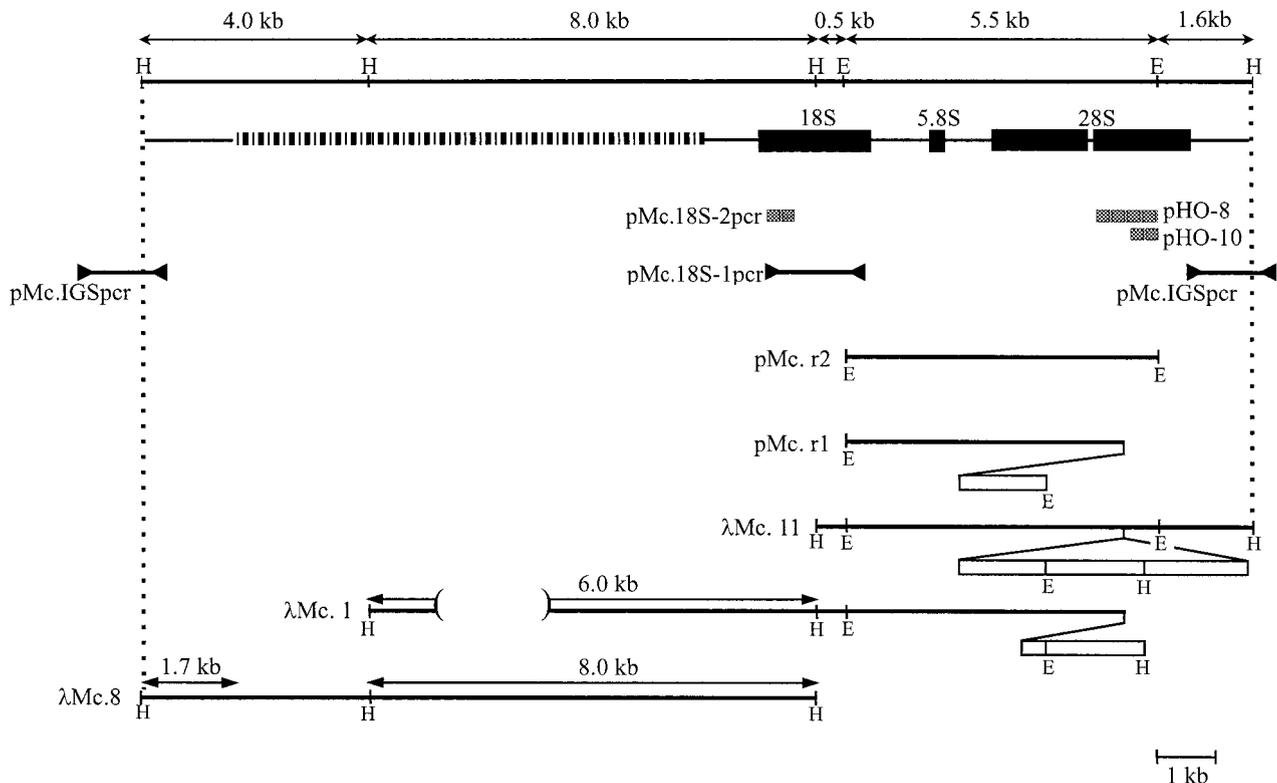


Fig. 1. Map of the *M. croslandi* rDNA repeating unit (top) and the clones analyzed (bottom). The rDNA repeating unit is defined by a *Hind* III site within the IGS. The conserved rRNA coding regions are indicated by filled boxes, spacer regions without repeats are indicated by thin lines, and those with repeats by broken bars. The probes used for plaque hybridization, pMc.18S-2pcr and pHO-8, are shown by gray boxes. Clones are indicated by thin lines with individual clone designations. There are non-28S rDNA sequence (open boxes) at the R2 insertion site in the pMc.r1, λ Mc.11 and λ Mc.1 clones. The λ Mc.8 *Hind* III fragment is 2kb longer than the λ Mc.1 fragment due to the IGS length variation. E, *Eco*R I; H, *Hind* III

(JM109, INVa-F' and XL1-Blue), plasmid growth, isolation of plasmid DNAs, and restriction enzyme digestions were performed according to standard methods (Sambrook *et al.*, 1989). Sequences of subcloned DNAs were determined using a DyeDeoxy terminator Cycle Sequencing kit (Perkin Elmer) and an ABI PRISM 310 Genetic Analyzer (PE Applied Biosystems). Subclones containing a portion of the internal IGS (λ Mc.1–7 and λ Mc.11–12) were sequenced using a GPS (genome priming system) –1 kit (BioLabs). The sequences determined were submitted to DDBJ/EMBL/GeneBank (accession numbers: AB052895, AB121786, AB121787, AB121788, AB121789 and AB121790).

Sequence data analyses

Alignment of sequences obtained from subcloned plasmids, secondary structure prediction based on Zuker method, and dot-matrix analyses were all performed with the aid of a personal computer program DNASIS-Mac version 3.6 (Hitachi Software Engineering).

RESULTS

Cloning and alignment of *M. croslandi* rDNA

The clone pMc.r2, containing a portion of 28S rDNA, was previously used for fluorescence *in situ* hybridization to chromosomes of species of *Myrmecia* (Hirai *et al.*, 1994). We first determined the complete nucleotide sequence of pMc.r2 (accession number AB052895) (Fig. 1). This clone contained part of the 18S rRNA gene, ITS 1, the 5.8S rRNA gene, ITS 2 and part of the 28S rRNA gene, leaving the sequence including the intergenic spacer (IGS) lying between the right portion of the 28S rRNA gene and the left portion of the 18S rRNA gene to be cloned and sequenced. In order to clone and sequence the entire single unit of the rDNA, we amplified the 18S rDNA region that was not contained in pMc.r2, using a forward primer designed using the 18S rRNA gene sequences common to *D. melanogaster* (Tautz *et al.*, 1988) and *Polistes dominulus* (Chalwatzis *et al.*, 1995), and a reverse primer designed using the sequence of the 18S rDNA at the 5' end of pMc.r2. A 1.8 kb fragment obtained was cloned (pMc.18S-1pcr) and sequenced (accession number AB121786).

pMc.r1 is another ant rDNA clone prepared previously and is longer than pMc.r2 (Hirai *et al.*, 1994). Restriction mapping showed that the region around the 3' end of pMc.r1 was different from that of pMc.r2 (Hirai *et al.*, 1994). We determined the sequence of this region and found that pMc.r1 contained an inserted sequence in the 28S rRNA gene region (Fig. 1).

We then screened a library constructed from genomic DNAs partially digested with *Hind* III with pMc.18S-2pcr and/or pHO-8 (Fig. 1) as probes. We obtained two clones which hybridized with both pMc.18S-2pcr and pHO-8, three clones with only pHO-8 and two clones with only pMc.18S-2pcr.

These clones were digested with *Hind* III and probed with pMc.18S-2pcr, pHO-8 and pHO-10. After restriction mapping, subcloning and partial sequencing of both the 5'- and 3'-end regions, the clones were aligned with pMc.r2 and pMc.18S-1pcr (Fig. 1). λ Mc.1 contained the 18S and 28S regions, λ Mc.8 contained the upstream sequence of 18S

and λ Mc.11 contained the sequence spanning from part of 18S to downstream of 28S. There is a single *Hind* III site within the 18S rRNA gene; a *Hind* III site was also found at the 5' end of λ Mc.11 and at the 3' end of the λ Mc.8.

We designed a forward primer toward the 3' direction based on the sequence near the 3' end of λ Mc.11 and a reverse primer toward the 5' direction from the sequence near the 5' end of λ Mc.8 in order to determine if any *Hind* III fragments exist between the 3' end of λ Mc.11 and the neighbouring *Hind* III site at the 5' end of λ Mc.8 within a portion of the intergenic spacer (IGS) region. An amplified fragment from genomic DNA was cloned (pMc.IGSpcr) and sequenced. We found that the DNA fragment contained the 3' end portion of λ Mc.11 and the 5' end portion of λ Mc.8 and only one *Hind* III site (Fig. 1). The DNA sequence data clearly indicate that the rDNA illustrated in Fig. 1 is a complete unit of the tandemly arranged repeated gene, and that we obtained clones covering the entire unit.

The sequence of the rDNA repeating unit

To extend the sequence analyses in the either direction from the core rDNA coding region, we further examined subclones λ Mc.1, λ Mc.8 and λ Mc.11. The 6.0 kb *Hind* III-*Hind* III fragment from λ Mc.1 and the 1.6 kb *Eco*R I-*Hind* III fragment at the 3' end of λ Mc.11 were sequenced (accession numbers AB121787, AB121788, respectively). The 1.7 kb *Hind* III-*Ban* III fragment of the 5' end of λ Mc.8 was also sequenced (accession number AB121789), but the remaining 2.3 kb *Ban* III-*Hind* III fragment was not sequenced completely, because it contained small subrepeat structures (see below).

The 5' and 3' ends of each rRNA gene were determined based on sequence comparison using the sequences of *D. melanogaster* as a reference: the 18S rRNA gene was 1,922 bp, the 5.8S rRNA gene was 109 bp and 28S rRNA gene was 4,243 bp long.

λ Mc.1, λ Mc.11 and pMc.r1 each contained an inserted sequence with a different size at the R2 target site (Burke *et al.*, 1999) that we found in 28S rDNA. R2 element (a non-LTR retrotransposon) recognizes 54 bp-long sequence highly conserved in Arthropoda at the target site. These insertions were partially sequenced and aligned with the known R2 element sequences of insect and other arthropod species (Burke *et al.*, 1999). The results indicated that the insertions indeed represent the R2 element. Among the 3 clones, only λ Mc.11 had a duplication of the 26 bp target sequence.

A fragment of λ Mc.1 (6.0 kb, see Fig. 1) was completely sequenced and a fragment of λ Mc.8 (8.0 kb, see Fig. 1) was partially sequenced. These sequences showed that the difference in these two clones was due to the number of the internal subrepeat sequences of IGS (see below).

Dot-matrix analyses of rDNA

The rDNA unit of *M. croslandi* was compared with that of *D. melanogaster* (Tautz *et al.*, 1988) by dot-matrix analy-

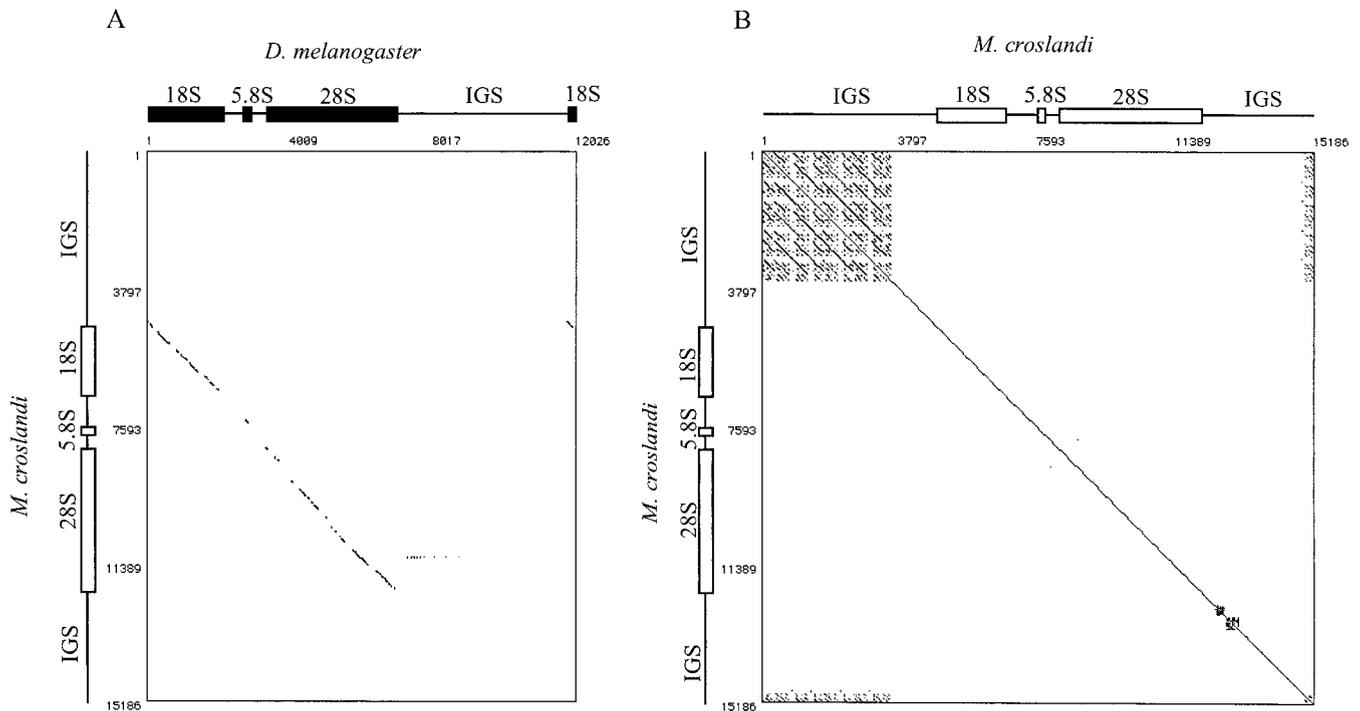


Fig. 2. Dot matrix comparisons of rDNA sequences. All comparisons were made as described in the text. (A) Comparison between *M. croslandi* and *D. melanogaster* sequences. The filled boxes indicate rDNA coding regions of *D. melanogaster*. The open boxes indicate predicted rDNA coding regions of *M. croslandi*. (B) Comparison of *M. croslandi* sequences.

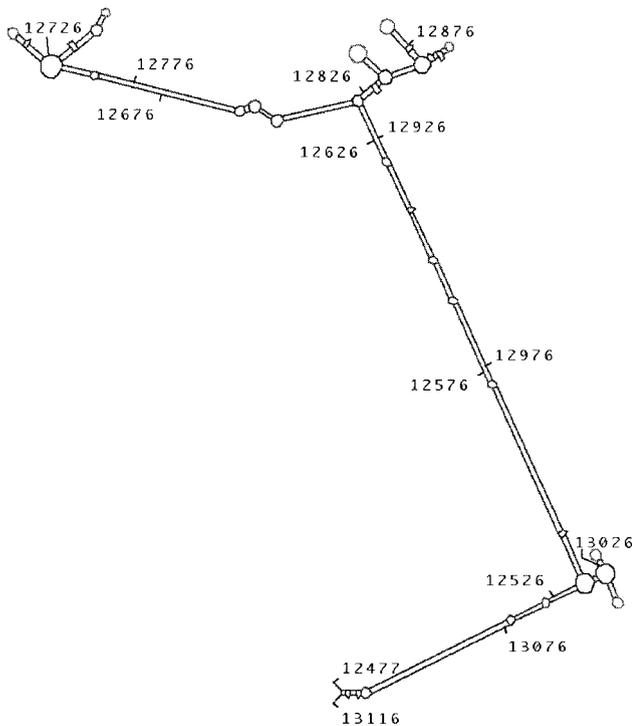


Fig. 3. A potential secondary structure that can be formed by the IGS sequence immediately following the 3' end of 28S rDNA in *M. croslandi* (see Fig. 2B). The nucleotide sequence (1.6 kb) in the *EcoR* I–*Hind* III region is shown at the top right in Fig. 1 and numbered 1–1,607.

ses (Fig. 2A). For this purpose, the sequence of the ant rDNA was arranged so that the 5' end of *Hind* III site of the 6.0 kb sequence of λ Mc.1 within the IGS subrepeat was numbered as the first nucleotide, and the 3' end of *Ban* III site of the 1.7 kb sequence of λ Mc.8 within the IGS subrepeat was numbered as nucleotide 15,186. Twenty bases were compared at a time, requiring 18 bases to be identical to register a dot. Linear-like plots appeared in the rDNA coding regions. The sequences of the rRNA genes were highly conserved with respect to those of *D. melanogaster* (identity; 18S: 76%, 5.8S: 86%, the first half of 28S: 52%, the latter half of 28S: 65%). No sequence similarity was detected between the IGS and ITS regions, as has been noted in previous studies (Long and Dawid, 1980; Hillis and Dixon, 1991).

The ant rDNA was compared with itself using the same method (Fig. 2B). Extensive stitch-like clusters were observed in the IGS region, suggesting that the same sequence arrays are present as tandemly arrayed subrepeats. Minor stitch-like clusters were also observed near the 5' end of IGS. Analysis of the sequences corresponding to these minor clusters, nucleotide position 12477–13116, based on Zuker method for prediction of RNA structure (Zuker, 1989), showed that they could form a secondary structure with long stems (Fig. 3).

Sequences of subrepeats inside the IGS region

The 6.0 kb fragment of λ Mc.1 contained two different kinds of subrepeats with lengths of 293 bp and 229 bp. The

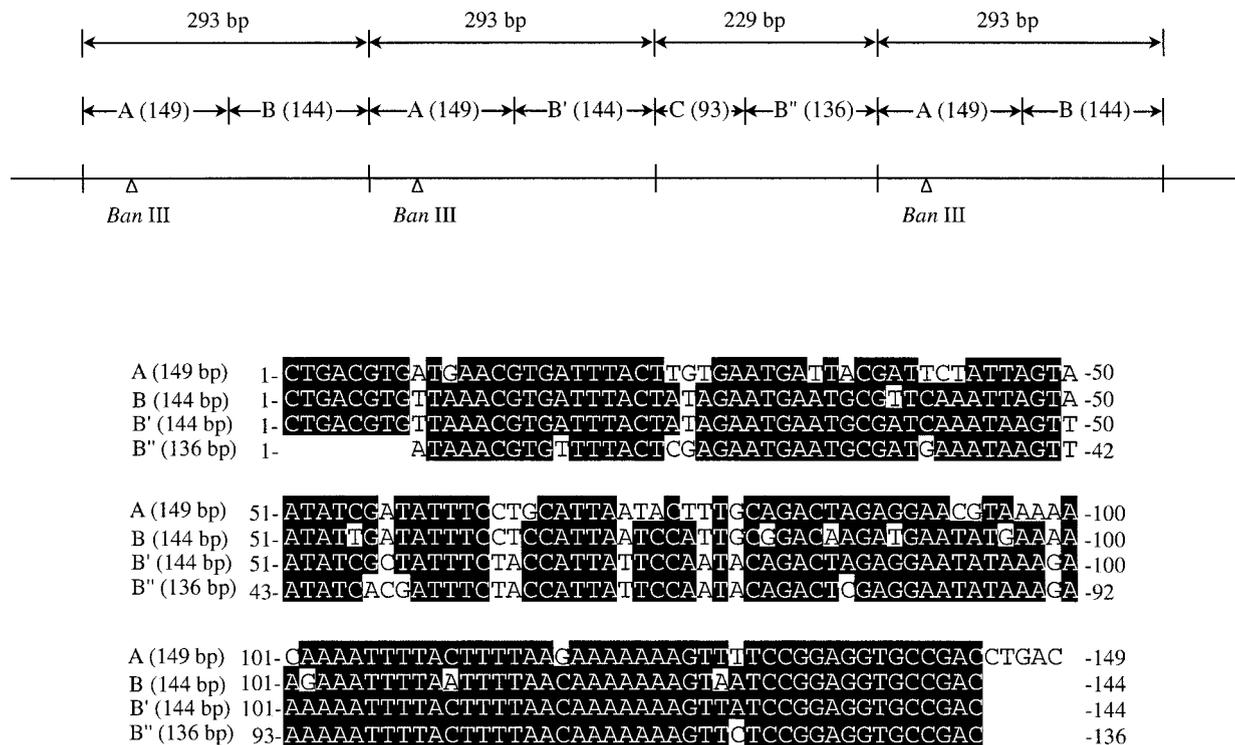
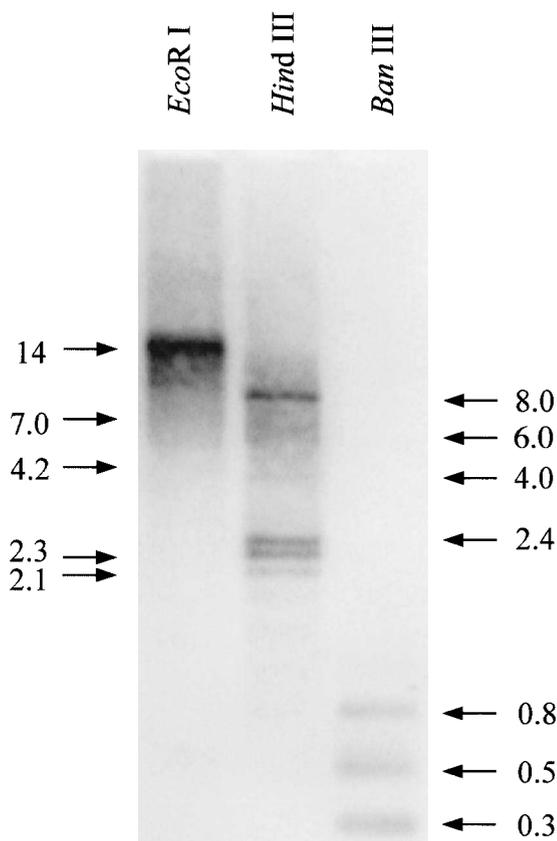


Fig. 4. Structure of the repeat region in the IGS. Sequence analysis of the *Ban* III-digested fragments first demonstrated the presence of two kinds of repeats, of 293 bp and 229 bp. Closer examination revealed the existence of the smaller subrepeats shown on top. Individual subrepeat sequences, except for C, which has no sequence similarity to the others, are shown below.



229 bp subrepeats were found sandwiched between the 293 bp subrepeats sporadically and not in any uniform manner. The 293 bp subrepeats had a *Ban* III site within the sequence, but the 229 bp subrepeats did not (data not shown but see the database, accession number AB121787) (Fig. 4). Thus, the *Ban* III fragment sizes are expected to be either 0.3 kb (293 bp) from [293 bp+293 bp] subrepeats or 0.5 kb (522 bp) from [293 bp and 229 bp] subrepeats. When the *Ban* III digest of the 6.0 kb fragment of λ Mc.1 was separated by electrophoresis on an agarose gel, a 0.3 kb and a 0.5 kb band were observed. On the other hand, when the *Ban* III digest of the 8.0 kb fragment of λ Mc.8 was similarly examined, a 0.8 kb-long band was observed in addition to the bands of 0.3 kb and 0.5 kb (data not shown). Cloning and sequencing of the 0.8 kb fragments showed that they contained tandemly arrayed 0.3 kb and 0.5 kb subrepeats with 293 bp subrepeats that had lost a *Ban* III site (accession number AB121790).

Closer examination showed that the 293 bp subrepeat could further be divided into two shorter subunits, 149 bp (A) and 144 bp (B or B'), with similar sequences (identity: A/B

Fig. 5. Southern hybridization for IGS subrepeats. The *M. croslandi* genomic DNAs were singly digested with the indicated restriction enzymes and probed with a DNA fragment corresponding to the IGS subrepeat sequence. The fragment sizes (kb), estimated by comparison with marker DNAs (1 kb ladder (Biolabs)), are indicated.

81%, A/B' 81%, B/B' 88%). There was a *Ban* III site within the 149 bp subunit but not within the 144 bp subunit. Similarly, the 229 bp subrepeat could be divided into two shorter subunits of 93 bp (C) and 136 bp (B'') (identity: A/B'' 79%, B/B'' 81%, B'/B'' 94%) (Fig. 4).

The DNA of λ Mc.8 was digested with *Hind* III and delivered two fragments of 4.0 kb and 8.0 kb. The results indicated that there is a single *Hind* III site within the region of subrepeats in the IGS. Digesting the 4.0 kb fragment with *Ban* III, we obtained 0.3 kb, 0.5 kb and 1.7 kb fragments. The sequence of the 1.7 kb fragment showed that this fragment contained neither 293 bp nor 229 bp subrepeats (see Fig. 1).

Genomic Southern hybridization

The length of the IGS was estimated by genomic Southern hybridization using a 0.3 kb *Ban* III fragment of the IGS subrepeats from λ Mc.8 as a probe. Ant genomic DNAs were completely digested with *Eco*R I, *Hind* III or *Ban* III and probed (Fig. 5). In the *Eco*R I digest, major bands were detected at and around 14 kb. In the *Hind* III digest, a major band at 8.0 kb, two less prominent bands at 2.4 kb and 2.2 kb, and minor bands at 6.0 kb, 4.0 kb and 2.0 kb were observed, while in the *Ban* III digest three distinct bands (0.8 kb, 0.5 kb and 0.3 kb) were detected. Taken together (see also Figs. 1 and 4) the results indicate that the IGS region varies in length (7–12 kb) due to variation in the length of subrepeats.

DISCUSSION

The present study revealed the complete structure and sequence of a single rDNA repeating unit of *M. croslandi*. The boundaries of the individual coding regions and associated ITSs of the 18S, 5.8S and 28S rRNA genes were determined by using the sequences of *D. melanogaster* rRNA genes as references (Tautz *et al.*, 1988). As is typical for eukaryotic rDNA units (Long and Dawid, 1980), the unit of *M. croslandi* rDNA contains, in addition to the core genes, a long intergenic spacer (IGS) sequence. The *M. croslandi* IGS consists of terminal regions at both ends with unique sequences and a long internal region with repeated short stretches.

Eukaryotic rRNA genes are generally composed of a few hundred to a few thousand copies of tandemly arrayed repeated units, and those of insects are no exception (Long and Dawid, 1980). In *D. melanogaster*, there are 200–250 copies of rDNA units that are tandemly arrayed (Glover, 1981). Hymenopteran insects are also known to have multiple copies of rDNAs (Bigot *et al.*, 1992). Numbers of *Myrmecia* species, including *M. croslandi*, were shown to carry multiple copies of rDNA in the significantly vast area on the chromosome complements (Hirai *et al.*, 1994, 1996). The fact that we could PCR amplify a fragment (cloned as pMc.IGSpcr, see Fig. 1), whose sequence was aligned at its 3' end with the 5' end of the rDNA unit, and at its 5' end with

the 3' end of the rDNA unit (Fig. 1), indicates that the copies were arrayed in tandem fashion.

There have been only a few reports describing the complete molecular structure and the sequence of the rDNA unit in insects (Tautz *et al.*, 1988; Amako *et al.*, 1996). The order of the coding and non-coding (spacer) regions, as is typical of that in the eukaryotes, is shown to be the same among *M. croslandi* (Hymenoptera), *D. melanogaster* (Diptera) and *A. pisum* (Hemiptera) by comparing the sequence results of the entire single repeating unit of their rDNAs (present results; Tautz *et al.*, 1988; Amako *et al.*, 1996). High sequence similarity between *M. croslandi* and *D. melanogaster* in the entire coding regions (and for the 28S coding region of *A. pisum* (Amako *et al.*, 1996)) was demonstrated as expected (data not shown; the *A. pisum* sequence data for the 18S and 5.8S regions have not been submitted to the database and hence are not available for comparison). The sequences in the coding regions are well conserved among various invertebrate and vertebrate species (Hillis and Dixon, 1991).

In contrast, the sequences in the non-coding regions are quite dissimilar among the three insect species. The subrepeat region within the IGS in *M. croslandi* consists of two different repeats (293 bp and 229 bp), that in *D. melanogaster* of three different repeats (95 bp, 330 bp and 240 bp) (Tautz *et al.*, 1988), and that in *A. pisum* of identical 247 bp repeats (Kwon and Ishikawa, 1992). The sequences of the subrepeats have no similarity among the three species, each belonging to a different order. Tautz *et al.* (1987) compared the IGS sequences of four *Drosophila* species and presented the results showing the overall evolutionary relationships between these species.

The complete sequence of the rDNA of *M. croslandi* revealed in the present study provides the means to design PCR primers anywhere inside the rDNA of *M. croslandi*. Primers prepared based on the unique sequence of the IGS region amplified a DNA fragment of *M. croslandi*, but did not amplify any DNA fragment in two other ant species (*Myrmecia pilosula* and *Nothomyrmecia macrops*, our unpublished data). This highly specific amplification suggests that the sequence information of IGSs may even be species-specific.

It will thus be of considerable interest to examine the IGS sequences, both unique and subrepeat, among species of the same Genera, Subfamilies and Families in various insect Orders. If the subrepeat patterns and sequences reflect phylogenetic relationships at the Genus, Subfamily, or Family levels, the information should be very useful for the clarification of evolutionary relationships. Even in the cases in which the subrepeat patterns and sequences do not reflect phylogenetic relationships, the information should at least be useful for the identification of a species in the social insects, such as ants, in which mating experiments are often quite difficult, if not impossible, to perform. In any event, for the comparison of distantly related groups of species, only the sequences of the conserved regions appear

to be useful, and thus PCR primers must be designed accordingly to amplify only conserved regions.

In the present study, the sequence of the 1.5 kb *EcoR*I–*Hind* III fragment around the 3' end of the 28S rRNA gene obtained from λ Mc.11 was refractory to determination by the ordinary dye-termination method. The sequence was finally determined by the GPS (genome priming system) method. The GPS method interrupts stem construction by the insertion of a transposable element, allowing us to determine the nucleotide sequence (Biery *et al.*, 2000). We found a relatively short sequence containing both tandemly and non-tandemly arranged stretches following the end of the coding region of 28S rRNA gene (Fig. 2B). The sequence represents a characteristic feature found in the ant rDNA unit, and previously unknown in other species such as *D. melanogaster* (Tautz *et al.*, 1988) and *A. pisum* (Amako *et al.*, 1996). This short sequence is expected to form a secondary structure with long stems (Fig. 3), which may disturb the polymerase extension reaction in the sequencing process. Whether this structure would affect the rDNA transcription process *in vivo* remains to be determined.

We found clones carrying an R2 element (a site-specific non-LTR retrotransposon) inserted at the 54 bp-long consensus R2 target sequence found in the 28S coding region (Fig. 1). One clone had an R2 insertion next to the target sequence and a duplication (26 bp) of the target sequence, supporting the R2 transposing model (Burke *et al.*, 1999). The proportion of rDNA copies carrying R2 elements was high; four of the eight clones isolated from the library constructed with *EcoR*I digest of the genomic DNA (pMc.1 to 8), and all four clones isolated from the *Hind* III library, contained R2 elements (data not shown). It has been said that the copy number of functional rDNAs is not large, since the R2 insertion may disturb transcription (Luan and Eickbush, 1996). The large bands hybridized with rDNA fragments detected in some *Myrmecia* species by FISH analyses (Hirai *et al.*, 1994) might simply represent an extensive array of rDNAs that became inactive due to the insertion of R2 elements (Jakubczak *et al.*, 1991). The number of active rDNAs may be comparable to that observed in other species, compensated to the normal level by amplification of rDNA without R2 insertion.

Knowledge of the complete unit of rDNA of ants may contribute not only to the elucidation of the molecular phylogeny of Hymenoptera, but also may provide materials for studies of rDNA functions. The subrepeat region in IGS has been shown to control the transcription of rDNA in *Drosophila* (Kohorn and Rae, 1983; Grimaldi and Nocera, 1988), and in other organisms (Grummt, 1982; Mougey *et al.*, 1996). The rate of rDNA transcription is efficiently regulated according to the growth rate of the cells and body size (Zafraan *et al.*, 1998; Giordano *et al.*, 1999; Frank *et al.*, 2002). The mechanisms to be studied include the differences in gene expression between haploid (male) and diploid (female) cells in Hymenoptera, and differentiation of body sizes in different castes of eusocial insects such as ants.

ACKNOWLEDGMENTS

We thank Hirotami T. Imai of the National Institute of Genetics for his introduction to the studies and generous provision of the ant material. We are grateful to Hajime Ishikawa and Hiromichi Makita of the University of the Air, Zhi-Hui Su and Kazunori Yamazaki of the JT Biohistory Research Hall, and Kyoichi Sawamura of the University of Tsukuba for their valuable advice.

REFERENCES

- Amako D, Kwon OY, Ishikawa H (1996) Nucleotide sequence and presumed secondary structure of the 28S rRNA of pea aphid: implication for diversification of insect rRNA. *J Mol Evol* 206: 469–475
- Baroni Urbani C, Bolton B, Ward PS (1992) The internal phylogeny of ants (Hymenoptera: Formicidae). *Syst Entomol* 17: 301–329
- Biery MC, Stewart FJ, Stellwagen AE, Raleigh EA, Craig NL (2000) A simple *in vitro* Tn7-based transposition system with low target site selectivity for genome and gene analysis. *Nucleic Acids Res* 28: 1067–1077
- Bigot Y, Lutcher F, Hamelin MH, Périquet G (1992) The 28S ribosomal RNA-encoding gene of Hymenoptera: inserted sequences in the retrotransposon-rich regions. *Gene* 121: 347–352
- Burke WD, Malik HS, Jones JP, Eickbush TH (1999) The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. *Mol Biol Evol* 16: 502–511
- Chalwatzis N, Baur A, Stetzer E, Kinzelbach R, Zimmerman FK (1995) Strongly expanded 18S rRNA genes correlated with a peculiar morphology in the insect order of Strepsiptera. *Zoology* 98: 115–126
- Crosland MWJ, Crozier RH (1986) *Myrmecia pilosula*, an ant with only one pair of chromosomes. *Science* 231: 1278
- Frank DJ, Edgar BA, Roth MB (2002) The *Drosophila melanogaster* gene *brain tumor* negatively regulates cell growth and ribosomal RNA synthesis. *Development* 129: 399–407
- Giordano E, Peluso I, Senger S, Furia M (1999) *minify*, a *Drosophila* gene required for ribosome biogenesis. *J Cell Biol* 144: 1123–1133
- Glover DM (1981) The rDNA of *Drosophila melanogaster*. *Cell* 26: 297–298
- Grimaldi G, Nocera PPD (1988) Multiple repeated units in *Drosophila melanogaster* ribosomal DNA spacer stimulate rRNA precursor transcription. *Proc Natl Acad Sci USA* 85: 5502–5506
- Grummt I (1982) Nucleotide sequence requirements for specific initiation of transcription by RNA polymerase I. *Proc Natl Acad Sci USA* 79: 6908–6911
- Hassouna N, Michot B, Bachellerie JP (1984) The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucleic Acids Res* 12: 3563–3583
- Hillis DM, Dixon MT (1991) Ribosomal DNA: Molecular evolution and phylogenetic inference. *Quart Rev Biol* 66: 411–453
- Hirai H, Yamamoto MT, Ogura K, Satta Y, Yamada M, Taylor RW, Imai HT (1994) Multiplication of 28S rDNA and NOR activity in chromosome evolution among ants of the *Myrmecia pilosula* species complex. *Chromosoma* 103: 171–178
- Hirai H, Yamamoto MT, Taylor RW, Imai HT (1996) Genomic dispersion of 28S rDNA during karyotypic evolution in the ant genus *Myrmecia* (Formicidae). *Chromosoma* 105: 190–196
- Hölldobler B, Wilson EO (1990) *The Ants*. The Belknap Press of Harvard University Press, Cambridge, Mass
- Jakubczak JL, Burke WD, Eickbush TH (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc Natl Acad Sci USA* 88: 3295–3299
- Kohorn BD, Rae PM (1983) A component of *Drosophila* RNA poly-

- merase I promoter lies within the rRNA transcription unit. *Nature* 304: 179–181
- Kwon OY, Ishikawa H (1992) Unique structure in the intergenic and 5' external transcribed spacer of the ribosomal RNA gene from the pea aphid *Acyrtosiphon pisum*. *Eur J Biochem* 206: 935–940
- Long EO, Dawid IB (1980) Repeated genes in eukaryotes. *Annu Rev Biochem* 49: 727–764
- Luan DD, Eickbush TH (1996) Downstream 28S gene sequences on the RNA template affect the choice of primer and the accuracy of initiation by the R2 reverse transcriptase. *Mol Cell Biol* 16: 4726–4734
- Mougey EB, Pape LK, Sollner-Webb B (1996) Virtually the entire *Xenopus laevis* rDNA multikilobase intergenic spacer serves to stimulate polymerase I transcription. *J Biol Chem* 271: 27138–27145
- Ogata K (1991) Ants of the genus *Myrmecia* Fabricius: a review of the species groups and their phylogenetic relationships (Hymenoptera: Formicidae: Myrmeciinae). *Syst Entomol* 16: 353–381
- Pélandakis M, Higgins DG, Solignac M (1991) Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. *Genetica* 84: 87–94
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, New York
- Schmitz J, Moritz RFA (1998) Molecular phylogeny of Vespidae (Hymenoptera) and the evolution of sociality in wasps. *Mol Phylogenet Evol* 9: 183–191
- Sullender BW (1998) A molecular phylogeny for the Formicidae. http://www.ruf.rice.edu/~bws/molecular_phylogeny.html
- Taylor RW (1978) *Nothomyrmecia macrops*: A living-fossil ant rediscovered. *Science* 201: 979
- Taylor RW (1991) *Myrmecia croslandi* sp. n., a karyologically remarkable new Australian jack-jumper ant (Hymenoptera: Formicidae: Myrmeciinae). *J Aust Entomol Soc* 30: 288
- Tautz D, Tautz C, Webb D, Dover GA (1987) Evolutionary divergence of promoters and spacers in the rDNA family of four *Drosophila* species: Implications for molecular coevolution in multigene families. *J Mol Biol* 195: 525–542
- Tautz D, Hancock JM, Webb DA, Tautz C, Dover GA (1988) Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Mol Biol Evol* 5: 366–376
- Zaffran S, Chartier A, Gallant P, Astier M, Arquier N, Doherty D, Gratecos D, Sémériva M (1998) A *Drosophila* RNA helicase gene, *pitchoune*, is required for cell growth and proliferation and is a potential target of d-Myc. *Development* 125: 3571–3584
- Zuker M (1989) Computer prediction of RNA structure. *Methods in Enzymology* 180: 262–288

(Received July 29, 2003 / Accepted October 11, 2003)