Marine Biology 110, 323–327 (1991)

Marine Biology
International Journal on Life in Oceans and Coastal Waters

Nucleotide sequencing of highly repetitive DNA from seven species in the coral genus *Acropora* (Cnidaria: Scleractinia) implies a division contrary to morphological criteria *

J. McMillan [1], T. Mahony [1], J. E. N. Veron [2] and D. J. Miller [1]

[1] Department of Chemistry and Biochemistry, James Cook University, Townsville, Queensland 4811, Australia
[2] Australian Institute of Marine Science, PMB No. 3, Townsville, Queensland 4810, Australia

Date of final manuscript acceptance: March 27, 1991. Communicated by G. F. Humphrey, Sydney

Abstract. Nucleotide sequences have been determined for 31 homologous 118 base-pair highly repeated DNA sequences from seven species of *Acropora*. A matrix was constructed from the sequence data and subjected to phylogenetic analysis using heuristic search routines in the PAUP (phylogenetic analysis using parsimony) program, Version 3.0L. These analyses confirm a close relationship between two species of one subgeneric group (*A. pulchra* and *A. millepora*), but identify a division in a group of six species which is contrary to taxonomic groupings based on morphological criteria.

Introduction

Acropora is the second-most widespread genus (next to *Porites*) of scleractinian coral, spanning both the Indian and Pacific Oceans and occurring also in the western Atlantic. It is by far the dominant coral of most Indo-Pacific reefs in terms of both abundance and diversity. This numerical superiority, combined with the wide environment-correlated growth form variations which all species show, has resulted in some 371 nominal (or described) species, far more than any other scleractinian genus. The number of true species is unknown, but 105 species are recognized in the Indo-west Pacific alone, and the total world complement may be as many as 152 (Veron in preparation).

So many species, covering such a wide morphological variation, has created formidable taxonomic as well as interpretive problems. *Acropora* as a genus is well defined and no species have doubtful generic designations. However, all attempts to create subdivisions within *Acropora* have failed, except for a single group of three species which comprise the subgenus *Isopora*. All taxonomic revisions of *Acropora* up to Wallace (1978) have resulted in subgeneric divisions of several types, but as none of these are based on comprehensive in situ study, all fail to recognize the extent of intraspecific growth-form variation. As a result, most species cross at least one subgeneric boundary and many have synonyms in two or more subgenera. Wallace recognized this and she (Wallace 1978), and subsequently Veron and Wallace (1984), divided the genus into a series of "species groups" without taxonomic status. These groups were created for the practical purpose (in lieu of subgenera) of providing a first step to identification, and are based on gross skeletal morphology as well as more fundamental similarities of skeleton detail. Some groups are relatively well defined, while others are composed of species which have a similarity which may be superficial.

The initial goal of this study was to explore the possibility of using molecular techniques where morphological studies have failed to create definable subdivisions within *Acropora*. To this end, we have characterized a highly repetitive sequence family of approximately 118 base-pair units which appears to be present throughout the subgenus *Acropora* but not in other genera within the family Acroporidae (McMillan et al. 1988, McMillan and Miller 1990). Nucleotide sequences of repeat units from two *Acropora* species have been reported (McMillan and Miller 1989), and hybridization experiments using cloned repeated sequence from *A. formosa* implied a species-relatedness series which differed in some respects to that based on morphological criteria (McMillan and Miller 1990). The present study reports the DNA sequences of repeat units from seven species of *Acropora*, and presents a phylogenetic analysis of the sequence data. Species were selected for analysis on the basis of preliminary hybridization experiments in which cloned repeated sequences from *A. formosa* were used as hybridization probes (McMillan et al. 1988, McMillan and Miller 1990).

Materials and methods

DNA was extracted from coral sperm collected from single colonies during the mass coral spawning events on the Great Barrier Reef as

* Please address all correspondence to Dr. D. J. Miller at the Department of Biochemistry, James Cook University

previously described (McMillan et al. 1988). DNA was prepared from the following species: *Acropora longicyathus*, *A. latistella*, *A. tenuis*, *A. digitifera*, *A. pulchra*, *A. millepora* and *A. formosa*. After preparative electrophoresis of MboI-digested genomic DNA, highly repetitive DNA sequences were cloned into pUC18 and the recombinant plasmids were characterized as previously described (McMillan and Miller 1990). The nucleotide sequences of cloned repeats were determined by the dideoxy method (Sanger et al. 1977, Hattori and Sakaki 1986) using either Klenow enzyme or modified T7 DNA polymerase (Tabor and Richardson 1987). In most cases clones were completely sequenced in both directions. Other general methods, including the construction of a genome library for *A. formosa* in λEMBL3, were as described in Maniatis et al. (1982) and Davis et al. (1986).

Results

General characteristics of the repeated sequences

Fig. 1 shows the nucleotide sequences of 31 homologous, cloned, repeated sequences from seven species of *Acropora*. The sequences were aligned manually. The consensus sequence was derived by first establishing a consensus for each species, and then from this calculating the sequence shown in Fig. 1. The consensus is 118 base pairs (bp) long, as were the majority of the individual repeated sequences. However, for some species the repeat units were frequently variable in length. For example, three of five *A. longicyathus* repeat units and three of six *A. digifera* repeats were 119 or 120 bp in length.

The repeated sequences are AT-rich; for example, the consensus contains 64.1% A + T. Certain sequence motifs appear to be highly conserved. In particular the sequence GTTTTGGTGGTTTTT at Positions 11 to 25 is present in 27 of the 31 repeat units, the exceptions being two repeats from *Acropora longicyathus* and two from *A. digitifera*. In general, the 5′ region of the repeat unit was relatively constant; 19 of the first 27 positions were invariant, and a further 6 positions varied in two or less cases. As we have previously stated (McMillan and Miller 1989), there is no obvious internal homology within the repeat unit. However, the triplet TTG occurs seven times in the consensus, and the tetranucleotide TTGC four times. Alignment of blocks of sequence around the TTG triplets implies that the 118 bp repeats may comprise a family of shorter related sequences which share the consensus TTTTGCTG.

Although the repeat sequence comprises approx 5% of the genome of *Acropora formosa* (McMillan and Miller 1989), clones containing the MboI repeat units were relatively rare in genome libraries of *A. formosa*. Of approx 5000 λEMBL3 clones screened (each containing

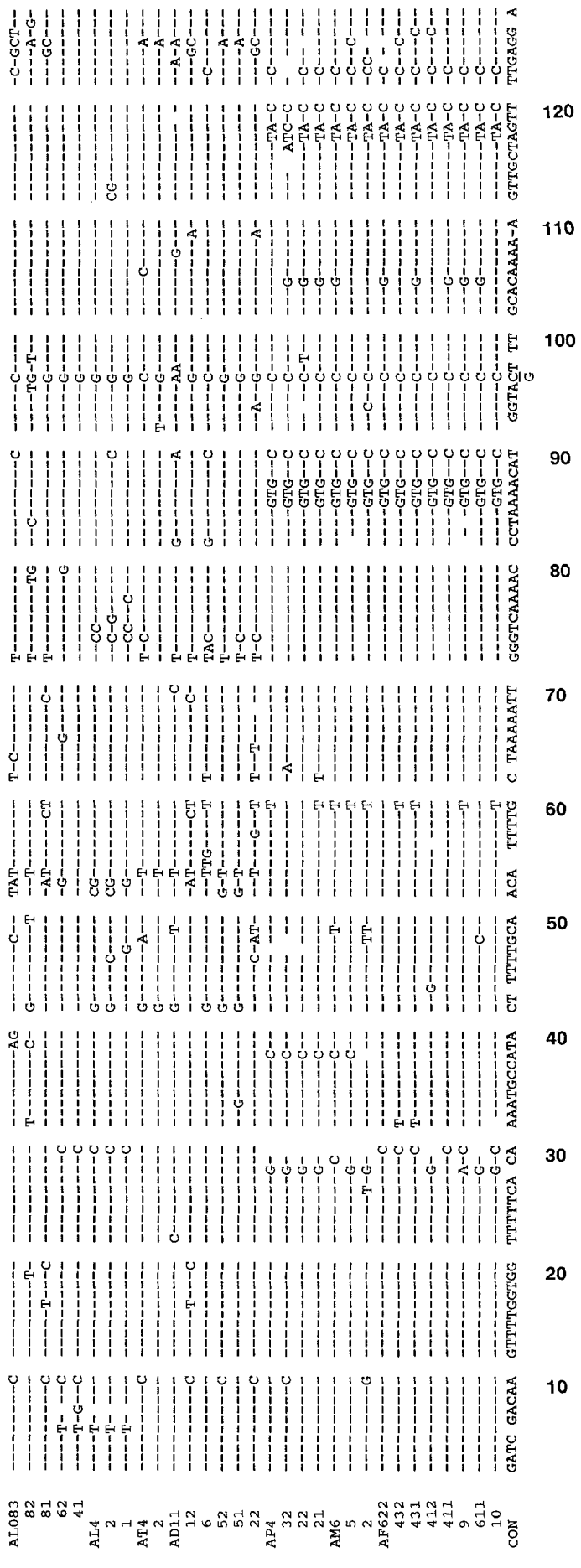

Fig. 1. *Acropora* spp. Nucleotide sequences of 31 individual repeat-sequence units from seven species. Consensus sequences for each species were established, and from these the overall consensus (CON) derived. Only differences from consensus are shown. Hyphens indicate identity with consensus, gaps indicate absence of a corresponding base. Numbers below the sequence indicate the nucleotide position in the consensus. Clone identities are: ALO, *A. longicyathus*; AL, *A. latistella*; AT, *A. tenuis*; AD, *A. digitifera*; AP, *A. pulchra*; AM, *A. millepora*; AF, *A. formosa*

inserts of average size 18 to 20 kbase pairs), only 5 positives were identified. Within these clones the repeats were organized in long tandem arrays; when two of these clones were subjected to partial digestion with Sau 3A, both yielded ladders upon hybridization with cloned repeat (Clone AF10). The proportion of the inserts accounted for by repeat units relative to unique- or low copy-number sequence in two of these lambda clones was assessed by comparison of the hybridization patterns given after probing blots of digested λ DNA with genomic *A. formosa* DNA or cloned repeated sequence (AF10). On this basis, only a small proportion of the inserts of these clones appeared to be unique; approx 2 to 3 kbase (of 18 to 20 kbase total) of unique sequence were present in each of two lambda clones.

Phylogenetic analysis

Alignment of the individual repeat sequences required adjusting the length of the 118 bp consensus to 128 positions by insertion of spaces (Fig. 1). When aligned in this way the sequences were identical at 50 positions, and a further 36 positions were not informative for parsimony analysis (i.e., variations were not shared by at least two individual clones). Sequence data for the remaining 42 positions were used to construct a matrix and analysed using Swofford's (1990) PAUP 3.0L. For these purposes, base changes or deletions were treated as unordered character-state transitions, and the data matrix was analysed using heuristic search routines. Fig. 2 shows the 50% majority-rule consensus tree generated from 1582 equally parsimonious trees (length = 112) found when a simple addition sequence was followed and branch-swapping was performed. Use of random addition sequences (10 replications) gave identical results. The minimum length of 1582 trees selected at random from the universe of all possible trees for the same data was 211 (one tree), the mean length of this set of random trees being 273.2 (SD = 15.6). Certain groupings were supported by these analyses with a high degree of confidence. The *Acropora pulchra*, *A. millepora*, and *A. formosa* repeat units always clustered together, and formed a cluster which was always well separated from a second cluster which comprised the *A. longicyathus*, *A. latistella* and *A. tenuis* repeat units. While the repeat units from *A. digitifera* generally clustered with the second group, the location of this species on the tree is equivocal. Features which clearly resolved the individual repeat units into two groups corresponding to the two species groups included five positions (# 85–87, 117 and 120) at which all the representative repeat units from the two groups of species were invariant within the group but differed between groups, and at Positions # 90, 96, 118 and 122 the species consensuses within the groups were identical but differed between groups.

The extent of repeat-sequence heterogeneity within species was highly variable. The repeated sequences in *Acropora latistella* were relatively homogenous, as were those across the *A. pulchra*/*A. millepora*/*A. formosa* group, whereas by several criteria those in *A. longicyathus*

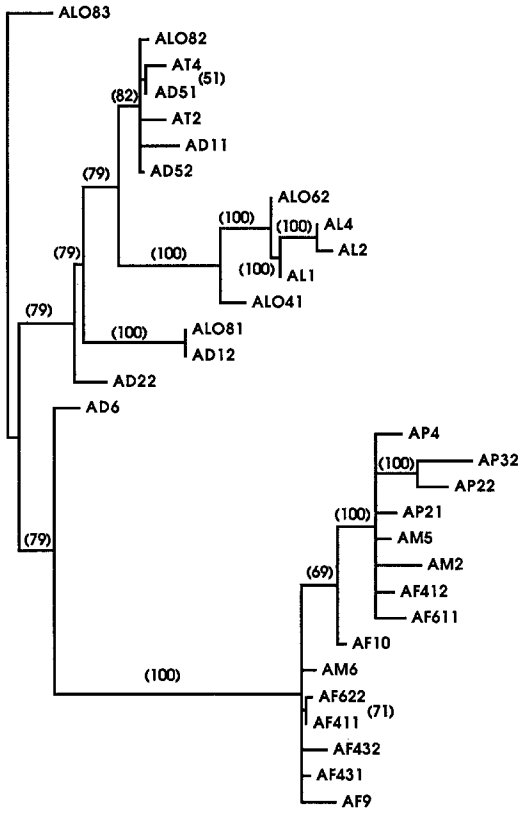

Fig. 2. *Acropora* spp. Phylogenetic analysis of repetitive DNA using PAUP 3.0 L. The unrooted phylogram shown is a 50% majority-rule consensus-tree generated from the 1582 equally parsimonious shortest trees found using heuristic search protocols, in which base changes were treated as unordered character-state transitions. Numbers in parentheses above or beside branches indicate percentage of trees in which that branching pattern occurs. Clone identities as in legend to Fig. 1

and *A. digitifera* were significantly more heterogenous. For example, for *A. formosa* eight individual repeat sequences have been determined, and the total number of positions at which variations occurs was 12. For *A. longicyathus* and *A. digitifera* the numbers of variable positions were 34 and 35, respectively (five *A. longicyathus* repeats and six *A. digitifera* repeats sequenced). Thus, more positions were variable within these species than across the whole "*A. formosa* group" (26 variable positions across *A. pulchra*, *A. millepora* and *A. formosa*).

Discussion

The nucleotide-sequence data confirm the interpretation of the Taq I digestion pattern presented in our earlier paper (McMillan and Miller 1990). The Taq I sites are homologous, that at the 5′-end of the repeat unit being present in a much wider range of species than that at the 3′-end, the latter being confined (with the exception of one repeat unit from *Acropora digitifera*) to *A. pulchra*, *A. millepora* and *A. formosa*.

In *Acropora formosa*, a significant proportion of the repeat units are organized in long tandem arrays, and

throughout the genus there appears to have been strong selection for the overall length of the individual repeat units to be 118 bp. Certain sequence motifs have been highly conserved; the core of the most highly conserved motif (GTTTTGGTGGTTTTT) is highly homologous with the chi sequence (GCTGGTGG), the target for the general recombination system in *Escherichia coli* (Kobayashi et al. 1984, Ponticelli et al. 1985). A sequence like chi is present at the core of many human minisatellites, and is thought to promote initial tandem duplication of unique-sequence DNA and/or to stimulate the unequal exchanges necessary to amplify the initial duplication (Jeffreys et al. 1985). It is possible therefore that the chi-like sequence in *Acropora* plays (or has played) a similar role in recombination processes.

The repeats in some species, such as *Acropora latistella*, were remarkably homogenous, whereas in species such as *A. longicyathus* and *A. digitifera* the degree of heterogeneity was much higher. In those species with higher sequence heterogeneity, the frequency of repeat-unit length-heterogeneity was also high, and these species included the only instances of variation in the most highly conserved region (GTTTTGGTGGTTTTT at Positions 11 to 25 in the consensus). Since the copy number for the MboI repeat unit is high (80 to 300×10^3; McMillan and Miller 1989), it is possible that the individual units sequenced do not accurately reflect sequence heterogeneity in the genome. However, the nucleotide sequences presented here are consistent with the available genomic digestion data. For example, genomic digestion patterns implied that the repeat units from *A. digitifera*, *A. millepora* and *A. formosa* contained one, one or two and two Taq I sites, respectively (McMillan and Miller 1990). The repeat units sequenced show that five of six *A. digitifera* repeats contained the predicted one Taq I sites, two of three *A. millepora* clones contained two Taq I sites and one clone contained one site, and six of eight *A. formosa* repeats contained one Taq I site. The implication is, therefore, that the sequences reported here are typical representatives of the genomic populations of repeat units.

Unequal crossing-over, or "molecular drive" (Dover 1982) is thought to be responsible for the progressive homogenization of reiterated sequences which appears to occur within species. Many variables influence the homogenization rate for homologous sequences, including unequal cross-over rates, genome sizes and copy numbers (Strachen et al. 1985). The copy number for the repeat-sequence units is relatively constant throughout the genus (McMillan and Miller 1990). However, no data on crossover rates or genome sizes are available. Other possible causes for the observed differences in sequence heterogeneity between species of *Acropora* include differences in time since speciation and/or differences in effective population sizes. Some species may be considerably older than others, therefore the repeats have had longer to become homogenized. The fossil record indicates that *Acropora* is amongst the oldest of extant corals, having been recorded from the Eocene of Indonesia and Somaliland. At the species level, however, the record is poor, as diagenesis usually destroys the skeletal detail needed for identification. Of the species selected in the present study, only *A. formosa* has any fossil record, having been recorded from the Pliocene of Papua New Guinea (Veron and Kelley 1988). Smaller effective population sizes and/or shorter generation times may have lead to more recombinational events per unit time in those species displaying low-sequence heterogeneity than in those with higher heterogeneity. However, there is no evidence to support major differences in present effective population sizes or generation times across the genus. All the species studied are known to take part in the annual mass coral-spawning. All are common in at least some biotopes, and all occur over a wide depth range, except for *A. digifera* which is restricted to shallow water exposed to strong wave action. All have wide geographic ranges, extending north to the Ryuku Islands and occurring in both the eastern Indian and western Pacific oceans.

Phylogenetic analysis of the repeat-sequence data clearly implies a major divergence between a group of species comprising *Acropora pulchra*, *A. millepora* and *A. formosa* and three of the other species, *A. longicyathus*, *A. latistella* and *A. tenuis*. A similar divergence is apparent in the ultrastructure of the sperm of these taxa (Harrison 1988 and personal communication). Whilst a detailed comparison with the species-group scheme of Veron and Wallace (1984) is not appropriate at this stage, the following comparisons can be made. Only two species (*A. pulchra* and *A. millepora*) occur in the same subgeneric group of Veron and Wallace (the *A. aspera* group), and these two species are closely grouped in the present study. Although the relative positions of groups in Veron and Wallace have a superficial identification function only, there is one area of agreement and two areas of substantial difference between the present results and morphological similarity. The agreement is in the wide separation of *A. latistella* from *A. pulchra*/*A. millepora*/*A. formosa*; the differences are in the grouping of *A. formosa* with *A. pulchra*/*A. millepora* (where both growth form and corallite structures are substantially different) and in the wide separation of *A. tenuis* and *A. pulchra*/*A. millepora* (where corallite structures are essentially similar). Possession of nearly identical repeat units by *A. longicyathus* and *A. digitifera* (Clones AL081 and AD12 differ at one position only) also implies a closer relationship between these two species than does the morphological data. Relationships within *Acropora* and between the Acroporidae and other scleractinian families, are presently under investigation in our laboratory, using the polymerase chain reaction to amplify the histone gene spacers and the ribosomal genes and spacers. Preliminary results support the divergence between two groups of *Acropora* species implied by the repeat-sequence analysis.

Acknowledgements. The authors thank Dr C. Wallace for identifying corals, Dr R. Babcock, Dr B. Willis and Dr P. Harrison for assistance in collection of coral gametes, and Dr D. Blair and Dr J. Pandolfi for their comments on the manuscript. The involvement of Dr J. Blok in the initial stages of nucleotide sequencing is also gratefully acknowledged. This work was supported by grants from the Marine Sciences and Technologies Grants Scheme (Grant No. 86/0866) and The Australian Museum.

Literature cited

Davis, L. G., Dibner, M. D., Battey, J. F. (1986). Basic methods in molecular biology. Elsevier, New York

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

Harrison, P. L. (1988). Comparative ultrastructure of scleractinian spermatozoa and its evolutionary implications. Ph. D. thesis. James Cook University of North Queensland, Townsville

Hattori, M., Sakaki, Y. (1986). Dideoxy sequencing method using denatured plasmid templates. Analyt. Biochem. 152: 232–238

Jeffreys, A. J., Wilson, V., Thein, S. L. (1985). Hypervariable 'minisatellite' regions in human DNA. Nature, London. 314: 67–73

Kobayashi, I., Stahl, M., Stahl, F. (1984). The mechanism of chi-*cos* interaction in RecA-RecBC-mediated recombination in phase λ. Cold Spring Harb. Symp. quant. Biol. 49: 497–506

Maniatis, T., Fritsch, E. F., Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor Press, New York

McMillan, J., Miller, D. J. (1989). Nucleotide sequences of highly repetitive DNA from scleractinian corals. Gene (Amsterdam, Elsevier) 83: 185–186

McMillan, J., Miller, D. J. (1990). Highly repeated DNA sequences in the scleractinian coral genus *Acropora*: evaluation of cloned repeats as taxonomic probes. Mar. Biol. 104: 483–487

McMillan, J., Yellowlees, D., Heyward, A., Harrison, P., Miller, D. J. (1988). Preparation of high molecular weight DNA from hermatypic corals and its use for DNA hybridization and cloning. Mar. Biol. 98: 271–276

Ponticelli, A. S., Schultz, D. W., Taylor, A. F., Smith, G. R. (1985). Chi-dependent DNA strand cleavage by RecBC enzyme. Cell 41: 145–151

Sanger, F., Nicklen, S., Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. Proc. natn. Acad. Sci. U.S.A. 74: 5463–5476

Strachen, T., Webb, D., Dover, G. (1985). Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. EMBO J 4: 17071–1708

Swofford, D. L. (1990). PAUP: phylogenetic analysis using parsimony, version 3.0. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois

Tabor, S., Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. natn. Acad. Sci. U.S.A. 84: 4767–4771

Veron, J. E. N., Kelley, R. (1988). Species stability in reef corals of Papua New Guinea and the Indo Pacific. Association of Australasian Palaeontologists, Sydney (Mem.; ISBN O 949466 042)

Veron, J. E. N., Wallace, C. C. (1984). Scleractinia of eastern Australia. Part V. Family Acroporidae. Australian National University Press, Canberra

Wallace, C. C. (1978). The coral genus *Acropora* (Scleractinia: Astrocoeniina: Acroporidae) in the central and southern Great Barrier Reef province. Mem. Qd Mus. 18: 43–103

Note added to proof

Since submission of the manuscript we have analysed the data matrix using Farris's HENNIG 86 software package, which gave essentially the same results as those reported above.