Mar Biotechnol (2008) 10:538–553
DOI 10.1007/s10126-008-9093-x

ORIGINAL ARTICLE

Unique Mitogenomic Features in the Scleractinian Family Pocilloporidae (Scleractinia: Astrocoeniina)

Chienhsun Chen · Chih-Yung Chiou · Chang-Feng Dai · Chaolun Allen Chen

Received: 15 December 2007 / Accepted: 2 March 2008 / Published online: 14 May 2008

Abstract The complete DNA sequences of three mitochondrial (mt) genomes were obtained from the scleractinian corals, *Stylophora pistillata*, *Pocillopora damicornis*, and *Madracis mirabilis*, and were compared to the published mt genomes to elucidate phylogenetically unique features of the family Pocilloporidae. The entire mt genomes of pocilloporid corals ranged from 16,951 to 17,425 bp with the A+T contents of their sense strands ranging from 68.4% to 70.2%. The gene order of protein-coding genes was identical to those of other scleractinian corals. The novel *atp8* gene, first described in confamilial *Seriatopora* corals, was also confirmed using reverse transcription-polymerase chain reaction (RT-PCR), Northern blot, and sequence analyses in other genera of the Pocilloporidae. The intergenic spacer between *atp6* and *nad4*, containing distinct repeated elements, conserved sequence blocks and domains, and functional structures, possesses typical characteristics of a putative control region for the four coral genera. A duplicated *trnW*, detected in the region close to the *cox1* gene and which shares the highly conserved primary and secondary structures of its original counterpart, was discovered in both *Seriatopora* and *Stylophora*. These molecular characteristics are unique and provide the phylogenetic information for future evaluation of the status of the family Pocilloporidae in the evolutionary history of scleractinian corals.

Keywords Scleractinian · Pocilloporidae · Mitochondrial genomes · Novel *atp8* · Putative control region · Duplicated *trnW*

Introduction

Mitochondrial (mt) genomes of higher animals are usually intron less, compact, closed circular DNA. Their gene contents are conserved with 12 or 13 protein-coding genes, one small and one large subunit rRNA genes, and 22 tRNA genes (Wolstenholme 1992). These genes are arranged with very few intervening nucleotides, except for a sequence that lacks genes but includes the molecule's major transcription promoter and origin of replication (Clayton 1991; Shadel and Clayton 1997). Because of its maturity in genetic-marker developmental history and natural characteristics of maternal inheritance, mtDNA showed some advantages over nuclear DNA markers in studies of population subdivision (Lin et al. 2006; Cho et al. 2007; Moriya et al. 2007; Azuma et al. 2008), species discrimination (Blair et al. 2006), and conservation implication (Ikeguchi et al. 2006; Mahidol et al. 2007), resolving population genetics conflicts (Burridge and Versace 2007). In the last two decades, studies of the complete mt genomes of several lower invertebrates have confronted the precepts of mitogenomic evolution (Boore 1999). Increasing evidences have shown that mt genomes of many invertebrates differ from the general patterns of higher animals, thus shedding new light on the mitogenomic evolution of metazoans and their application in phylogeny (Burger et al. 2003).

Electronic supplementary material The online version of this article (doi:10.1007/s10126-008-9093-x) contains supplementary material, which is available to authorized users.

C. Chen · C.-F. Dai · C. A. Chen
Institute of Oceanography, National Taiwan University, Taipei 106, Taiwan

C. Chen · C.-Y. Chiou · C. A. Chen (✉)
Research Center for Biodiversity, Academia Sinica, Nankang,
Taipei 11529, Taiwan
e-mail: cac@gate.sinica.edu.tw

The Anthozoa is one of the classes in the Cnidaria; their mt genomes possess several atypical features that differ from those of higher animals and closely related taxa (Table 1). First, a general form of the closed circular mt genome is exhibited in anthozoans, but it is present in a linear form in members of the subphylum Medusozoa (Bridge et al. 1992). Second, the cnidarian mt genome has a near-standard genetic code. It indicates that an adaptation of a highly biased mt codon-usage strategy was developed after the appearance of cnidarians (Beagley et al. 1998). Third, there are extensive intergenic spacers (IGSs) between genes and introns in *nad5* and *cox1* genes of hexacorals (Beagley et al. 1996, 1998; van Oppen et al. 2002). Fourth, the mt genomes of octocorals and sea

Table 1 A list of the Cnidaria mitogenomes available on GeneBank

Taxa	Protein no.	tRNA no.	Intron location	Note	Accession no. (Reference)[a]
Anthozoa					
Hexacorallia					
Antipatharia					
Chrysopathes formosa	13	2	*nad5*	C	NC008411 (C)
Actiniaria					
Metridium senile	14	2	*cox1*, *nad5*	C	NC000933 (A)
Nematostella sp.	13	2	*nad5*	C	NC008164 (G)
Corallimorpharia					
Discosoma sp. CASIZ 168915	13	2	*cox1*, *nad5*	C	NC008071 (G)
Discosoma sp. CASIZ 168916	13	2	*cox1*, *nad5*	C	NC008072 (G)
Rhodactis sp.	13	2	*cox1*, *nad5*	C	NC008158 (G)
Ricordea florida	13	2	*cox1*, *nad5*	C	NC008159 (G)
Scleractinia					
Acropora tenuis	13	2	*nad5*	C	NC003522 (M)
Agaricia humilis	13	2	*nad5*	C	NC008160 (G)
Anacropora matthai	13	2	*nad5*	C	NC006898 (L)
Astrangia sp.	13	2	*nad5*	C	NC008161 (G)
Colpophyllia natans	13	2	*nad5*	C	NC008162 (G)
Montastraea annularis	13	2	*nad5*	C	NC007224 (F)
Montastraea faveolata	13	2	*nad5*	C	NC007226 (F)
Montastraea franksi	13	2	*nad5*	C	NC007225 (F)
Montipora cactus	13	2	*nad5*	C	NC006902 (L)
Mussa angulosa	13	2	*nad5*	C	NC008163 (G)
Pavona clavus	13	2	*nad5*	C	NC008165 (G)
Pocillopora damicornis	14	2	*nad5*	C	NC009797 (E)
Pocillopora eydouxi	14	2	*nad5*	C	NC009798 (E)
Porites porites	13	2	*cox1*, *nad5*	C	NC008166 (G)
Seriatopora caliendrum	13	3	*nad5*	C	EF633601 (D)
Seriatopora hystrix	13	3	*nad5*	C	EF633600 (D)
Siderastrea radians	13	2	*cox1*, *nad5*	C	NC008167 (G)
Zoantharia					
Savalia savaglia	14	1	*cox1*, *nad5*	C	NC008827 (K)
Octocorallia					
Briareum asbestinum	14	1	No intron	C	NC008073 (G)
Pseudopterogorgia bipinnata	14	1	No intron	C	NC008157 (G)
Sarcophyton glaucum	14	1	No intron	C	AF063191 (B)
					AF063192 (B)
					AF064823 (H)
Hydrozoa					
Hydra oligactis	13	2	No intron	L	EU237491 (I)
Scyphozoa					
Aurelia aurita	15	2	No intron	L	NC008446 (J)

C Circular mitochondrial molecule, *L* linear molecule

[a] Reference: (*A*) Beagley et al. (1998), (*B*) Beaton et al. (1998), (*C*) Brugler and France (2007), (*D*) Chen et al. (2008), (*E*) Flot and Tillier (2007), (*F*) Fukami and Knowlton (2005), (*G*) Medina et al. (2006), (*H*) Pont-Kingdon et al. (1998), (*I*) Kayal and Lavrov (2008), (*J*) Shao et al. (2006), (*K*) Sinniger et al. (2007), (*L*) Tseng et al. (2005), and (*M*) van Oppen et al. (2002)

anemones, respectively, feature the occurrence of a bacterial MutS homologue and a homing endonuclease except for the ordinary 13 mt protein-coding genes (Pont-Kingdon et al. 1995; Beagley et al. 1996; Pont-Kingdon et al. 1998). Fifth, the loss of most tRNA genes was reported for cnidarians, with only two tRNA genes detected in most cases and one tRNA identified for octocorals (Beaton et al. 1998; Pont-Kingdon et al. 1998; van Oppen et al. 2002; Shao et al. 2006). Despite these unique characteristics, nucleotide-substitution rates in anthozoans are considered to be relatively slow, approximately 10~20-fold slower than the standard molecular clock rate of higher animals (van Oppen et al. 1999; Shearer et al. 2002; Hellberg 2006).

To date, 28 anthozoan mt genomes have been completely sequenced, more than half of which are from the order Scleractinia (Table 1). The scleractinian mt genome was first revealed in *Acropora tenuis* (van Oppen et al. 2002). Mitogenome of *Ac. tenuis* contains only two tRNA genes, an interruption of the *nad5* gene by a large group I intron, a putative control region in the IGS between *rns* and *cox3*, and the same orientation for all of the genes (van Oppen et al. 2002). Although it has been shown that all of the scleractinians mt genomes contained similar features and shared a uniform gene order, some significant differences exist among them (Fukami and Knowlton 2005; Tseng et al. 2005; Chuang 2006; Medina et al. 2006; Chen et al. 2008). For example, comparisons of the mt genomes between the Acroporidae (complex-clade scleractinians) and those of the *Montastraea annularis* complex (robust-clade scleractinians) showed that the mitogenomic organison is loose compact in the former (17,887~18,338 bp with 14 IGSs) than in the latter (16,138 bp with nine IGSs). While the putative control region is located in the *rns-cox3* IGS in the Acroporidae (van Oppen et al. 2002; Tseng et al. 2005), it was identified as being located in the *atp6-nad4* IGS in *Seriatopora* spp. (Chen et al. 2008). These differences might affect their phylogenetic utilities of revealing mt genomic evolution in scleractinians.

In this study, we conducted a comprehensive analysis of the mt genomes of four genera, *Madracis*, *Pocillopora*, *Seriatopora*, and *Stylophora*, of the family Pocilloporidae (Veron and Pichon 1976). The *Madracis* was remaining considered a genus in the Pocilloporidae, although it was assigned to the family Astrocoeniidae by unknown reason in a later revision (Veron 2000). The Pocilloporidae comprises globally distributed scleractinian corals in shallow-water reefs. Most of its members are major reef-building and pioneering species of many Indo-Pacific coral reefs (Grigg and Maragos 1974; Glynn 1976; Loya 1976; Wallace 1985). Despite its ecological importance, the evolutionary affinity of the Pocilloporidae remains unresolved. The Pocilloporidae is conventionally assigned to the suborder Astrocoeniina (or Archaeocoeniia) which includes the Acropordiae and Astrocoeniidae based on simple structures of corallite morphology (Vaughan and Wells 1943; Wells 1956; Veron 2000; Stolarski and Roniewicz 2001). In contrast to morphological similarities, there was a closer relationship between the Pocilloporidae and the Faviidae rather than the Acroporidae as indicated by molecular phylogenetic studies (Romano and Palumbi 1996, 1997; Chen et al. 2002).

Two recently completed sequences of mt genomes in *Seriatopora* corals revealed several unique characteristics which were first reported in scleractinians, as well as in anthozoans (Chen et al. 2008). These features included an idiosyncratic *atp8*, a duplicated *trnW* ($tRNA^{TRP}$), and a putative control region located between *atp6* and *nad4*, and highlighted the need for a further survey of mt genomes in confamilial genera of the Pocilloporidae. In this study, we extended genomic sequencing approaches to examine the mt genomes of *Pocillopora damicornis*, *Stylophora pistillata*, and *Madracis mirabilis*. The question particularly addressed is whether the molecular characteristics found in *Seriatopora* mt genomes are also conserved in the family Pocilloporidae.

Materials and Methods

Coral Collection and DNA Extraction

All scleractinians were collected from Taiwan, except for those of *Ma. mirabilis* which were collected from the Bocas del Toro Islands in the Caribbean off Panama (9°19′18″ N; 82°11′54″ W). Samples of *St. pistillata* and *Poc. damicornis* were collected from Chimei Island of the Penghu Archipelago (23°12′27″ N; 119°25′39″ E). The mt genome of *Seriatopora hystrix* was adopted from our previous study for further comparison (Chen et al. 2008). Apical fragments (1~2 cm long) of corals were kept in a modified guanidine solution (CHAOS solution) before DNA extraction as described by Fukami et al. (2004).

Total genomic DNA of the preserved sample was extracted with an optimized protocol by modifying a phenol–chloroform extraction (Fukami et al. 2004). An equal volume of phenol extraction buffer (PEB containing 100 mM TrisCl (pH 8.0), 10 mM ethylenediaminetetraacetic acid, and 0.1% sodium dodecyl sulfate) was added to the CHAOS solution just before DNA extraction. The mixed solution was extracted twice with an equal volume of phenol–chloroform (1:1). The DNA of the supernatant from the second extraction was precipitated by adding an equal volume of isopropanol. The precipitated DNA was washed with 70% EtOH and air-dried for 24 h with the tube inverted. The DNA was resuspended and dissolved in an appropriate volume of double deionized distilled water with RNaseA (10 mg/ml).

Long Polymerase Chain Reaction, Cloning, and Sequencing

The complete mt genome was amplified by means of two overlapping long polymerase chain reactions (PCR)s from extracted DNA of the collected samples. To accomplish this, partial sequences of *rns* and *rnl* were initially obtained by a standard PCR using coral-specific primers (Chen and Yu 2000; Romano and Palumbi 1996). From the *rns* and *rnl* sequences, we designed primers for the long PCR: 5′-GAC TTG GCT GTT CGG TTG TTA ATT AGA GGA GCG CG-3′ (F25) and 5′-CGC GCT CCT CTA ATT AAC AAC CGA ACA GCC AAG TC-3′ (R21); and 5′-TAC CCT GGG GAT AAC AGC GCA ATA ACG-3′(F02) and 5′-AAG GCC CAA TAA CCT TCC ATT GCA TCC GGT AGC-3′ (R06), respectively. According to the gene orientation of previously published mt genomes of scleractinian corals, two sets of primer combinations, F02 with R21 and F25 with R06, were used for the long PCR amplifications.

Long PCRs were conducted in a Px2 Thermocycler (Thermo Electron Corp.) and performed using the Long PCR Enzyme Mix (TaKaRa LA *Taq*) under the conditions recommended by the manufacturer. Long PCR reactions were setup in a volume of 50 μl: 1×PCR buffer, 1.5 mM $MgCl_2$, 0.2~0.4 mM of each deoxyribonucleotide triphosphate (dNTP), 0.5 μm of each primer, 2% of dimethyl sulfoxide, 1.875 U of the PCR enzyme mix, and approximately 0.5 μg of genomic DNA. Long PCR profiles were 1 cycle of 94°C for 2 min; 10 cycles of 96°C for 20 s, 65°C for 30 s, and 68°C for 8 min; followed by 25 cycles of 96°C for 20 s, 65°C for 30 s, and 68°C for 8 min, with a 10-s extension per cycle; then 1 cycle of 68°C for 10 min. The PCR products were electrophoresed in a 0.8% agarose gel to inspect their quantity and were then recovered from the agarose gels using Topo XL gel purification kit according to manufacturer's protocol (Invitrogen). Purified long PCR products were cloned into a pCR-XL-TOPO vector system (Invitrogen) and transformed into *Escherichia coli* (Top10). The sequence for each fragment was obtained in both directions by primer walking on the same purified PCR product on an ABI 377 automated DNA sequencer. The M13 forward and reverse primers were used to obtain the initial sequences from the ends of each fragment. The consensus sequences from 2~6 sequenced clones were present for each species.

Reverse-Transcription-PCR and Northern Blot Analyses of the Putative *atp8*

The *atp8* gene turned out to be a novel and functional gene in *Se. hystrix* using reverse transcription (RT)-PCR in our previous study (Chen et al. 2008). To examine whether the *atp8* gene in other pocilloporid corals is functional, RT-PCR and Northern blot analyses were conducted within the predicted *atp8* from the complementary DNAs (cDNAs) of *St. pistillata* and a Northern blot from *Se. hystrix*, respectively.

Total RNA was prepared from approximately 5 cm^2 of *St. pistillata* using the TRIzol reagent (Invitrogen) for the RT-PCR experiment. Coral samples were preserved in 2 ml TRIzol reagent for 10~20 min, and total RNA was extracted once with chloroform and treated with DNase I (Promega), then extracted again with 1:1 phenol–chloroform. The RNA was precipitated in isopropanol with the addition of sodium acetate (0.25 M), then dissolved in diethylpyrocarbonate-treated water.

The first-strand cDNAs of the total RNA were synthesized by RT-PCR using SuperScript III Reverse Transcriptase (Invitrogen) under conditions recommended by the manufacturer in a Px2 thermocycler. RT-PCRs were setup in a volume of 50 μl: 1×RT-PCR buffer, 1 μl of RNase block ribonuclease inhibitor (40 U/μl), 1 mM of each dNTP, and 1 μM of primer oligo(dT)17. RT-PCR profiles consisted of one cycle of 60°C for 10 min and one cycle of 42°C for 1.5 h.

Because the sequence contains several poly(A) regions which results in nonspecific products in the following PCR reaction, three pairs of primers located before and after the largest poly(A) region were used to detect the 3′-end of the transcribed *atp8*. Aliquots of first-strand cDNA were used in the PCRs (reagents as above) using primer pairs of primer PF1ST (5′-AGT GCC TCA GTT AAA AGT AA-3′) with primer PR1ST (5′-CCA AAT TTT TGT TAA GAT GG-3′), primer PF1ST with primer PR3ST (5′-TAC TTT AAC TTT ACT CCT TCA AAA AT-3′), and primer PF1ST with primer oligo(dT)17, while the negative control used the F30 (5′-TGG TTA TCC CCC TCA GGT GGT C) and PR1ST primers. PCRs were setup in a volume of 50 μl: 1× PCR buffer, 2.5 mM $MgCl_2$, 1 mM of each dNTP, 0.2 μm of each primer, 1 U of *Taq* polymerase, and 5 μl of cDNA aliquots. The PCR profiles consisted of 1 cycle of 94°C for 5 min; 35 cycles of 94°C for 30 s, 50 or 46°C for 30 s, and 72°C for 1 min; then 1 cycle of 72°C for 5 min. The PCR products were electrophoresed and inspected in 1.5% agarose gels, purified using the Montage PCR purification kit (Millipore) then cloned into a Topo TA cloning system (Invitrogen) and transformed into *E. coli* (Top10). The sequences of inserts were determined in both directions by M13 forward and reverse primers on an ABI 377 automated DNA sequencer.

For the Northern blot analysis, total RNA was prepared from *Se. hystrix* per a previously described protocol, except for modification of overnight DNAaseI treatment. Samples of 20 μg total RNA were separated in a 1.2% (*w*/*v*) denaturing formaldehyde agarose gel and then transferred onto a Hybond-N plus membrane (Amersham Pharmacia Biotech). The coral *atp8* cDNA (186 bp) and *cob* cDNA (330 bp) were used as DNA-specific probes. Labeling, hybridization, and

washing were carried out with the ECL Direct nucleic acid labeling and detection system (RPN3000; Amersham Biosciences) according to the manufacturer's protocol. Prehybridization was performed for 2 h at 42°C. Hybridization was performed at 42°C for 24 h. Membranes were washed to a final stringency of 15 min at 42°C in 1×SSC/0.1% sodium dodecyl sulfate after exposure to X-ray film at room temperature. The probes and the previously amplified long PCR product were used as positive controls.

Genome Annotation and Sequence Analysis

The DNA sequences were assembled using the software, Sequencher 4.2 (Gene Codes Corp.), then analyzed in Vector NTI 6.0 (InforMax). Open reading frames (ORFs) of considerable length (>50 amino acids) in the sequences were initially translated using cnidarian mt genetic codes (the 4th NCBI translation table), then compared with the databases using the BLASTX program (Gish and States 1993). The identified putative ORFs including protein coding and rRNA genes were aligned with the corresponding genes of previously published scleractinian mt genomes by using MEGA 3.1 (Kumar et al. 2004) for final recognition, with the weighted matrix of ClustalW (Thompson et al. 1994) and Gonnet (Gonnet et al. 1992). The nucleotide composition was calculated using BioEdit 7.01 (Hall 1999).

The hydropathy profiles of the unassigned ORFs were generated using the general method of Kyte and Doolittle (1982) and compared with those of the identified *atp8* of other scleractinian corals. Except for *atp8*, the overall codon usage of protein-coding genes was compared among pocilloporid corals.

Intergenic spacers (IGSs) were examined for the presence of canonical tRNAs with tRNAscan-SE search server 1.21 (Lowe and Eddy 1997) or DOGMA (Wyman et al. 2004), using the default search mode and specifying mt/chloroplast DNA as the source and the mold and protozoan mt genetic codes for tRNA structure prediction. According to these predicted structures, corresponding tRNA genes of pocilloporid corals were aligned, and positions with nucleotide substations were annotated.

According to known features of mt control region of *Ac. tenuis* (van Oppen et al. 2002) and *Seriatopora* spp. (Chen et al. 2008), the presence of potential secondary structures and tandem repeats in IGSs were used to identify their putative control regions. Tandem repeat sections in the IGSs were identified using the Tandem Repeat Finder 4.0 with default parameters (Benson 1999), then adjusted manually.

The repetitive regions were excluded from the following analysis in identifying conserved sequences, as the insertion/deletion processes of mt tandemly repeated sequences were not well understood in anthozoans. The modified sequences were aligned using MEGA 3.1 (Kumar et al. 2004) with a weighted matrix of ClustalW (Thompson et al. 1994) and opening and extension penalties of 10 and 3, respectively. The conserved sequence block (CSB) was then determined using Gblocks (Castresana 2000) with default parameters, and closely positioned CSBs were arbitrarily assigned a conserved sequence domain (CSD). To access the stabilities of the potential secondary structures of the putative control region, free energies of the CSD and their random sequences were examined using the DNA Mfold server (Zuker 2000).

Results

Composition and Organization of Mitochondrial Genomes in the Pocilloporidae

The entire mt genomes of pocilloporid corals range from 16,951 bp in length for *Ma. mirabilis* to 17,425 bp for *Poc. damicornis* (Table 2). The A+T contents of their sense strands were about 70% (68.4~70.2%; Table 2). Each of the mt genomes contained 12 identified ORFs, the small and large subunits of ribosomal RNA genes (*rns* and *rnl*), and two or three tRNA genes (Table 2), all of which were transcribed on the same strand (Fig. 1). The arrangements of protein-coding genes were identical to those of other scleractinian corals. The *nad5* was interrupted by an intron, which contained ten protein-coding genes, *rns*, and a putative control region (Fig. 1).

Protein-Coding Genes

Except for *atp8*, these identified ORFs corresponded to 12 of 13 typical protein-coding genes for pocilloporid corals (*cox1~cox3*, *nad1~nad6*, *nad4L*, *cob*, and *atp6*). High levels of nucleotide and amino acid similarities, respectively, ranging 82.5~99.6% and 83.3~100%, were detected among corresponding protein-coding genes of pocilloporid corals. The most divergent protein-coding gene among pocilloporid corals was *atp6*, with the significantly highest intergeneric pairwise genetic distances than those of other protein-coding genes ($p<0.05$ in all comparisons, Mann–Whitney U test; Fig. 2). In addition, no amino acid difference was detected among corresponding *nad4L* of these four species. Differences in divergences may indicate an unequal evolutionary rate among protein-coding genes of pocilloporid corals.

Idiosyncratic *atp8* Gene

Although the searching strategies of BLASTN and BLASTX did not detect *atp8* coding sequences or *atp8*-

Table 2 Mitogenomic organizations of *Ma. mirabilis*, *Poc. damicornis*, *Se. hystrix* and *St. pistillata*

Region	Position (intergenic nucleotides)				Start/stop codon
	Ma. mirabilis	*Poc. damicornis*	*Se. hystrix*	*St. pistillata*	
trnM	1–71 (0)	1–71 (0)	1–71 (0)	1–71 (0)	
rnl	72–2,008 (0)	72–1,988 (0)	72–1,975 (0)	72–2,007 (0)	
nad5(5′)	2,009–2,719 (0)	1,989–2,699 (0)	1,976–2,686 (0)	2,008–2,718 (0)	AUG
Group I intron(5′)	2,720–2,843 (0)	2,700–2,823 (0)	2,687–2,810 (0)	2,719–2,842 (0)	
nad1	2,844–3,821 (32)	2,824–3,801 (51)	2,811–3,788 (19)	2,843–3,820 (51)	AUG/UAA
cob	3,854–4,993 (222)	3,853–4,992 (220)	3,808–4,947 (220)	3,872–5,011 (220)	GUG/UAA
nad2	5,216–6,307 (1)	5,213–6,304 (1)	5,168–6,259 (1)	5,232–6,323 (1)	AUU/UAA
nad6	6,309–6,872 (−1)	6,306–6,869 (−1)	6,261–6,824 (−1)	6,325–6,888 (−1)	AUG/UAA
atp6	6,872–7,549 (0)	6,869–7,546 (0)	6,824–7,501 (0)	6,888–7,565 (0)	AUG/UAA
Putative control region	7,550–8,406 (0)	7,547–8,936 (0)	7,502–8,524 (0)	7,566–8,670 (0)	
nad4	8,407–9,852 (−1)	8,937–10,382 (0)	8,525–9,970 (12)	8,671–10,116 (0)	AUG/UAG (UAA)
rns	9,852–10,761 (0)	10,383–11,292 (0)	9,983–10,898 (0)	10,117–11,030 (0)	
cox3	10,762–11,541 (−25)	11,293–12,072 (−25)	10,899–11,678 (−25)	11,031–11,810 (−23)	AUG/UAA
cox2	11,517–12,455 (−19)	12,048–12,884 (−19)	11,654–12,472 (−19)	11,786–12,622 (−19)	AUG/UAG
nad4L	12,437–12,736 (−1)	12,866–13,165 (−1)	12,454–12,753 (−1)	12,604–12,903 (−1)	AUG/UAA
nad3	12,736–13,080 (0)	13,165–13,509 (0)	12,753–13,097 (0)	12,903–13,247 (0)	AUG/UAA
Group I intron(3′)	13,081–13,134 (0)	13,510–13,563 (0)	13,098–13,151 (0)	13,248–13,301 (0)	
nad5(3′)	13,135–14,238 (12)	13,564–14,691 (12)	13,152–14,279 (12)	13,302–14,429 (12)	/UAA
trnW	14,251–14,320 (1)	14,704–14,773 (1)	14,292–14,361 (1)	14,442–14,511 (1)	
Putative atp8	14,322–14,546 (668)	14,775–14,987 (851)	14,363–14,599 (91)	14,513–14,761 (−20)	AUG (GUG)/UAA
trnW’			14,691–14,760 (713)	14,742–14,803 (788)	
cox1	15,215–16,801 (150)	15,839–17,398 (27)	15,474–17,021 (38)	15,592–17,139 (38)	AUG/UAA
Genome size (bp)	16,951	17,425	17,059	17,177	
Intergenic spacer (bp)	1,086	1,163	1,106	1,111	
A+T content (%)	68.4	69.8	69.9	70.2	

Numbers in squares represent intergenic nucleotides between the gene on the same line and the gene on the line underneath, with a negative number indicating an overlap of that length

like motifs in pocilloporid corals, an unidentified ORF detected in the sequences might correspond to *atp8* of other scleractinians. This ORF is located between the two *trnW* of *Se. hystrix* and *St. pistillata*, and 1 bp downstream of *trnW* in *Ma. mirabilis* and *Poc. damicornis* (Table 2). The ORF was heterogeneous in length among pocilloporid corals, with 71~83 amino acid in pocilloporid corals.

Transcription of the unidentified ORF was confirmed by the fact that an RT-PCR product could be amplified in the predicted partial *atp8* using total RNA prepared from *Se. hystrix* (Chen et al. 2008) and *St. pistillata* (Fig. 3). The identity of the RT-PCR product was confirmed by sequencing (Fig. 3a). The expression of the ORF was also confirmed by the Northern blot analysis (Fig. 4). By probing with the coral *atp8* cDNA, a single band was detected in the total RNA prepared from *Se. hystrix*. The size of the mRNA was about 400 bp long, when extrapolated using the sizes of the probes (Fig. 4). The result strongly indicated that the presumptive *atp8* was expressed. The larger molecule should be a mature mRNA converted from precursor mRNA during posttranscriptional modification because the primary transcript is unstable and exists only briefly before being processed into mature mRNA. This transcript may contain extended untranslated regions at the 5′- and 3′-ends, and this phenomenon has also been reported in marine invertebrate mt systems (e.g. Gissi and Pesole 2003; Milbury and Gaffiney 2005). Hence, even though the size of the probed signal was larger than that of the presumptive *atp8* of *Seriatopora*, this result strongly indicated that the *atp8* was expressed.

This ORF was inferred to begin with Val-Pro-Gln-Leu in *Se. hystrix* and *St. pistillata* but with Met-Pro-Gln-Leu in *Ma. mirabilis* and *Poc. damicornis*, and these agreed with *atp8* from 13 released scleractinian mt genomes (Fig. 5). Although the overall amino acid similarities were low among the putative *atp8* and *atp8* of other scleractinians (25.6~34.6%), high similarities were observed at the beginning of the aligned sequences. This resulted in a well-conserved N-terminal motif (Fig. 5).

In addition to sequence similarity, the hydropathy profile of the ORF-encoded protein was similar to those of *atp8* in other scleractinians (Fig. 6). For example, the mean hydrophobicity of their first 12 amino acids was slightly neutral, then positively charged until the 36th amino acid.

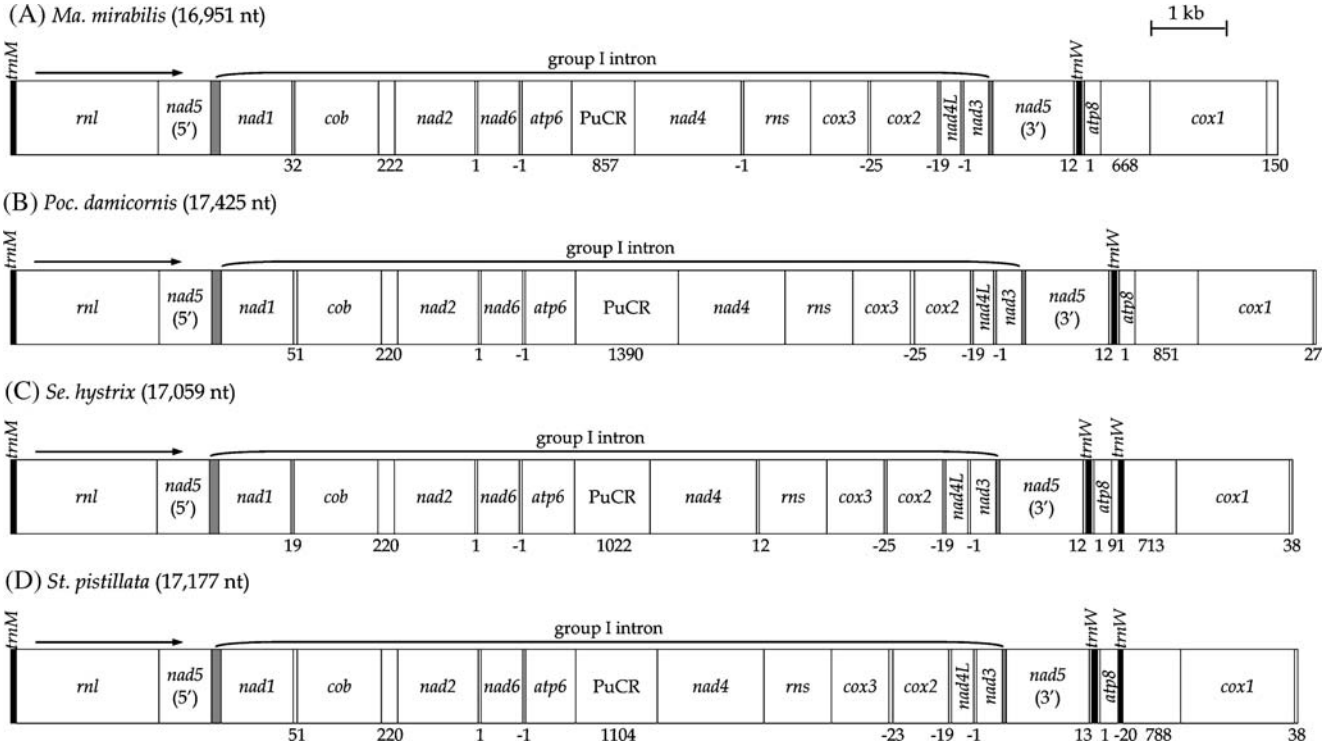

Fig. 1 Schematic overview of the gene map of pocilloporid corals **a** *Ma. mirabilis*, **b** *Poc. damicornis*, **c** *Se. hystrix*, **d** *St. pistillata*. Protein-coding genes and tRNA and rRNA genes are abbreviated as in the text. *Numbers below the figure* represent the sizes of the intergenic regions. The *black and shaded portions* indicate tRNA genes and group I introns, respectively. The putative control region is denoted PuCR. *Arrows* indicate the direction of transcription. Scale bar: 1 kb

However, the hydropathy varied at the C-terminals, which became negatively charged at the 52nd amino acid position of the presumptive *atp8* for *Ma. mirabilis* and *Poc. damicornis* and at the 58th amino acid position of the presumptive *atp8* for *Se. hystrix* and *St. pistillata* (Fig. 6a). Near the end of the C-terminal, the hydropathy gradually became positively charged in *Ma. mirabilis*, *Se. hystrix*, and *St. pistillata*, but not in *Poc. damicornis* (Fig. 6a). Taking the analyses of amino acid similarity and hydropathy together, this unidentified ORF is suggested to be the presumptive *atp8* of pocilloporid corals that ended with an extended 3′-end.

Codon Usage

Due to size differences and low amino acid similarities among the putative *atp8* and the corresponding gene in other scleractinians, the *atp8* was excluded from the following comparisons of codon usage. The 12 protein-coding genes of pocilloporid corals encoded 3,843~3,888 codons (data not shown), and no significant differences in codon usage were detected among them ($p>0.99$, Chi-squared test). Leucine was the most frequently used amino acid, followed by phenylalanine; while arginine was the least utilized. UUU (phenylalanine) was the most frequently used codon (12.4~13%), and UUA (leucine) was the second most common codon (7.8~8.4%). In addition, 10 or 11 protein-coding genes of pocilloporid corals have AUG as the start codon, while *nad2* and *cob*, respectively, have AUU and GUG (Table 2). Ten to 11 of their protein-coding genes have UAA as the stop codon, while *cox2* and *nad4* have UAG as the stop codon (Table 2).

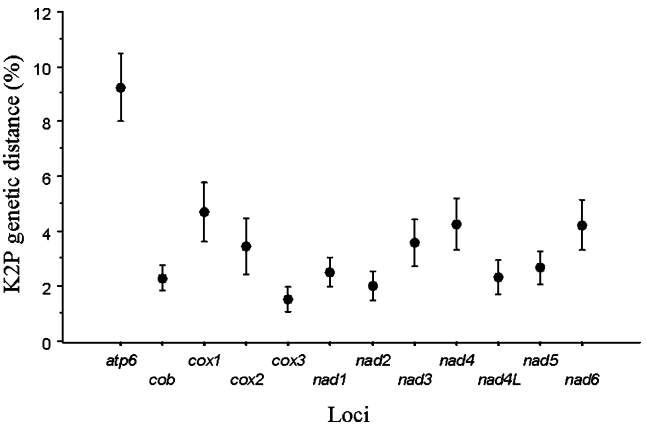

Fig. 2 Average genetic distances (Kimura 2-parameter) among corresponding protein-coding genes (except *atp8*) of pocilloporid corals. *Error bar*: ±SE

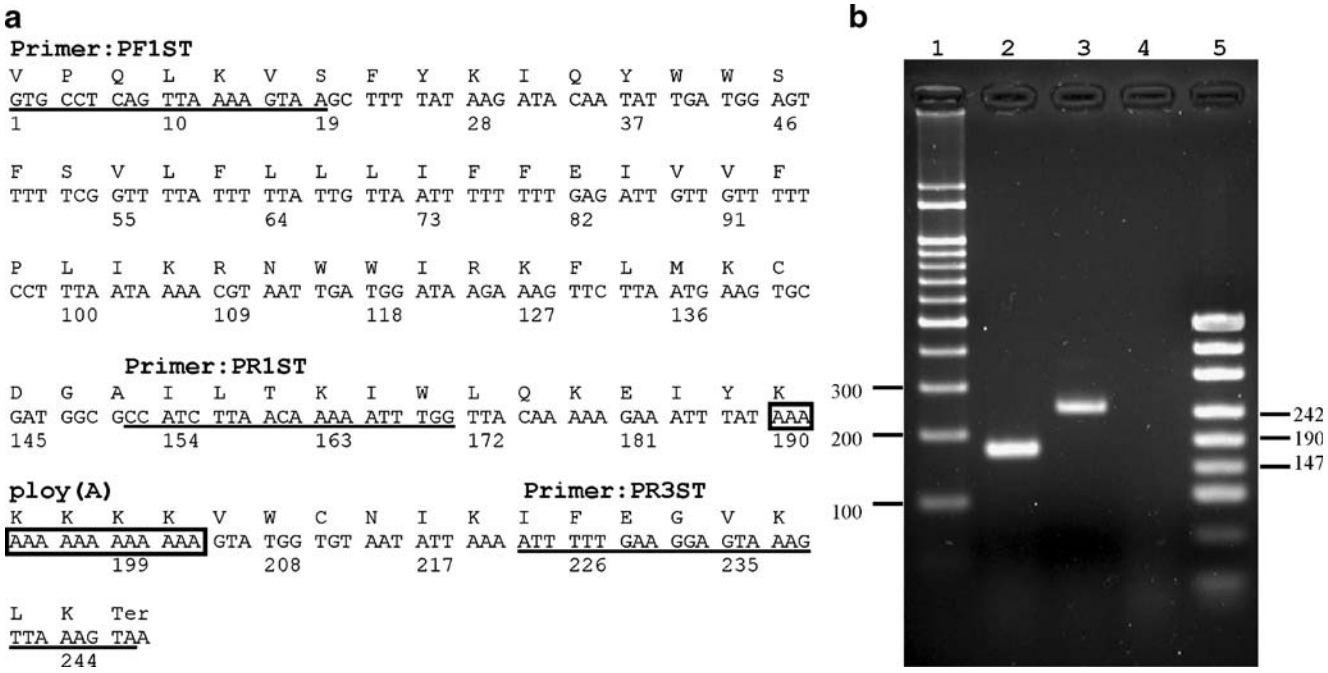

Fig. 3 Identification of the *atp8* by RT-PCR analysis. **a** Nucleotide and predicted amino acid sequences of the *atp8* of *St. pistillata*. The positions of primers used in the RT-PCR experiment and the largest poly(A) region are *underlined* and *indicated in blocks*, respectively. *Numbering* refers to nucleotide positions in the particular *atp8* coding sequence. **b** Agarose gel electrophoresis analysis of the RT-PCR product. *Lanes 1 and 5* Low-molecular-weight standards (Promega100-bp Ladder DNA marker and Fermentas pUC19 MspI marker), *lanes 2 and 3* RT-PCR products and their sizes estimated by comparison with the standards corresponded well with those predicted by the sequences (171 and 248 bp), *lane 4* negative control of RT-PCR by the primers spanning *nad5(5')* and *atp8*

Transfer RNA Genes

In addition to the two tRNAs, *trnM* and *trnW*, reported for scleractinians, an additional *trnW*, namely, *trnW'*, was identified in the region close to the *cox1* in the mt genomes of *Se. hystrix and St. pistillata*, but not in those of *Ma. mirabilis* and *Poc. damicornis* (Fig. 1). Although the anticodons of *trnM* and *trnW* were the same among the corresponding tRNAs of pocilloporid corals, explicit differences in nucleotide sequences were detected among them.

In total, 7 nucleotide substitutions were observed in comparisons of *trnM* among pocilloporid corals, and 21 nucleotide substitutions were observed in comparisons of the *trnW* gene and *trnW'* of *Se. hystrix* (Appendix 1). Sequence similarities among corresponding *trnM* genes of pocilloporid corals were 90.1~100% and were significantly higher than those of *trnW* genes of pocilloporid corals ($p<0.05$ Mann–Whitney U test; sequence similarity of *trnW* and *trnW'*: 74.2~100%). These substitutions, which occurred in stems of corresponding tRNAs, were either compensated for by a substitution in the complementary strand or involved changes such as G-C, G-U, or A-U pairings.

All the three tRNAs could be folded into typical secondary structures except for *trnW'* of *St. pistillata* (Appendix 1). These typical structures were 70 or 71

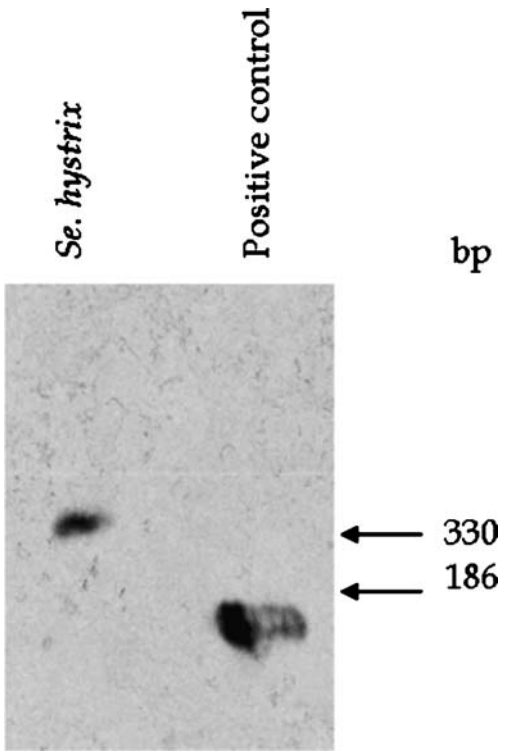

Fig. 4 Northern blot analysis of the *atp8* of *Se. hystrix*. *Lanes 1 and 2* represent a detected mRNA of *Se. hystrix* and a probe of the positive control, respectively. *Arrows and numbers*, respectively, indicate the position and size of the probes

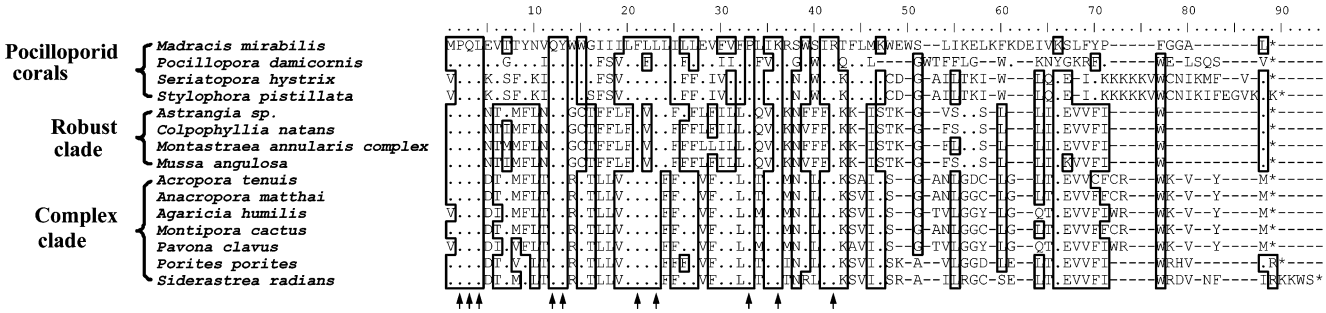

Fig. 5 Comparisons of amino acids of pocilloporid corals' mt presumptive *atp8* with the corresponding amino acid sequences of 13 other scleractinian corals, identified with their database accession numbers as follows: *Ag. humilis* NC008160, *Astrangia* sp. NC008161, *Co. natans* NC008162, *Mu. angulosa* NC008163, *Pa. clavus* NC008165, *Por. porites* NC008166, *Si. radians* NC008167 (Medina et al. 2006); *Monta. annularis* complex NC007224, NC007225, and NC007226 (Fukami and Knowlton 2005); *An. matthai* NC006898, *Monti. cactus* NC006902 (Tseng et al. 2005); and *Ac. tenuis* NC003522 (van Oppen et al. 2002). In the alignment, *dots* indicate identical amino acids compared to the first sequence, and *dashes* indicate indels. Under the sequences, an *arrow* indicates an amino acid conserved in all sequences

nucleotides (nt) in length and had an aminoacyl stem of 7 bp, a dihydrouridine (DHU) stem of 3 bp (4 bp in *trnW*), a D-loop of 6 or 9 nt, an anticodon stem of 5 bp, an anticodon loop of 7 nt, a TψC stem of 4 bp, and a TψC loop of 7 nt. In contrast with the standard secondary structures, a different format of the secondary structure was predicted for *trnW*' of *St. pistillata* (Appendix 1C). In addition to the deficiencies of complementary pairings at its

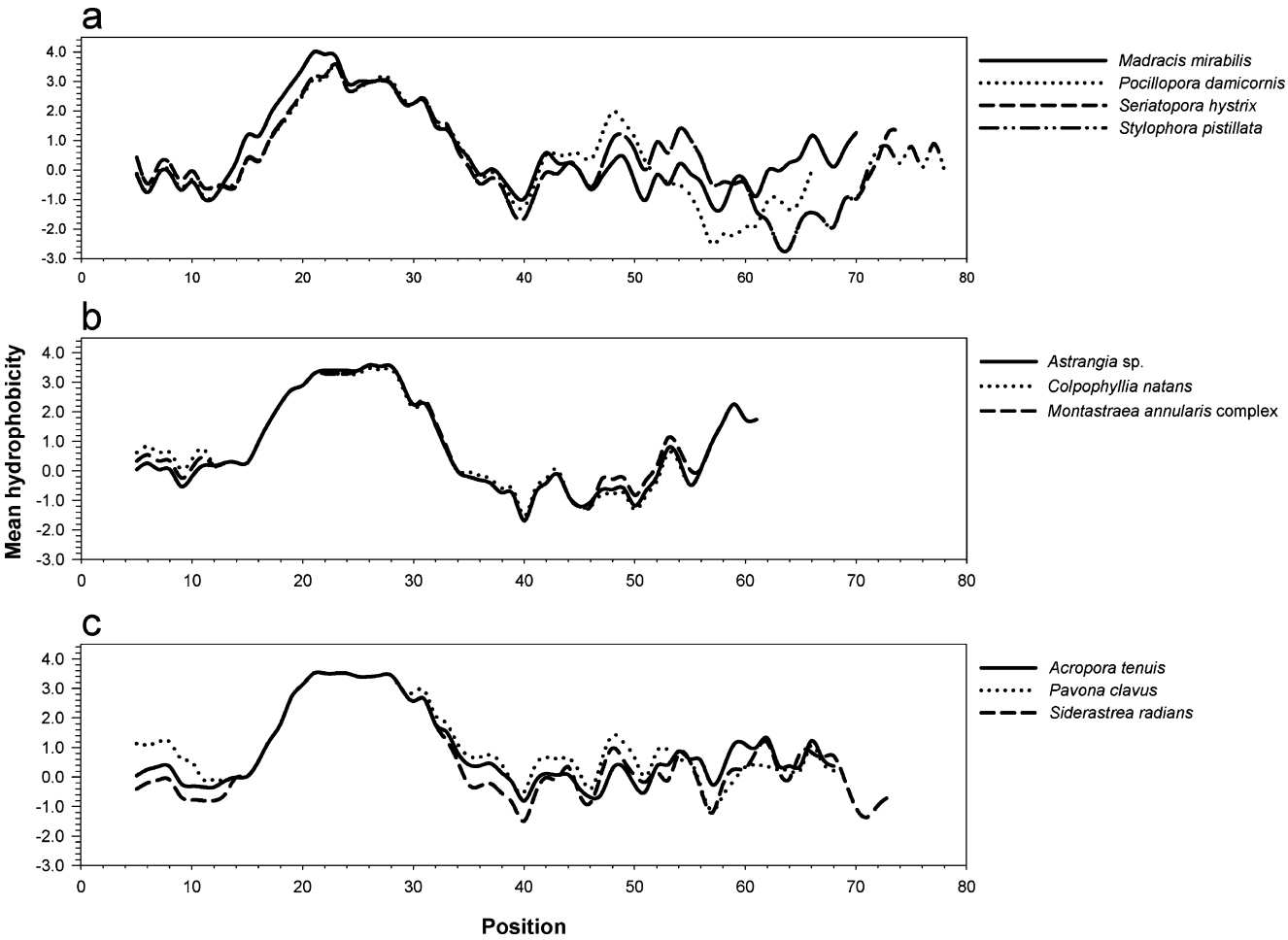

Fig. 6 Comparisons of hydropathy profiles among **a** pocilloporid, **b** robust-clade, and **c** complex-clade scleractinians, respectively. Window size=9

fifth and sixth positions, the *trnW'* gene of *St. pistillata* possessed a TψC stem of 3 bp, but did not have a TψC loop.

Intergenic Spacers

Various intergenic spacers (IGSs) were detected between most mt genes of pocilloporid corals (Table 2). Eight genes overlapped with each other in *Ma. mirabilis*: *nad6* with *atp6* by 1 bp, *nad4* with *rns* by 1 bp, *cox3* with *cox2* by 25 bp, *cox2* with n*ad4L* by 19 bp, and *nad4L* with *nad3* with 1 bp. Six protein-coding genes overlapped with each other in *Poc. damicornis* and *Se. hystrix*: *nad6* with *atp6* by 1 bp, *cox3* with *cox2* by 25 bp, *cox2* with *nad4L* by 19 bp, and *nad4L* with *nad3* by 1 bp. Eight protein-coding genes overlapped with each other in *St. pistillata*: *nad6* with *atp6* by 1 bp, *cox3* with *cox2* by 23 bp, *cox2* with *nad4L* by 19 bp and *nad4L* with *nad3* by 1 bp, and the *atp8* with *trnW'* by 20 bp.

IGSs, excluding the putative control region (see below), were 1,086~1,163 bp in pocilloporid corals. These regions varied in length from 1 to 851 bp (Table 2). Over 50% of the IGSs were attributed to the IGS of *atp8-cox1* for *Ma. mirabilis* and *Poc. damicornis*, and *trnW'-cox1* of *Se. hystrix* and *St. pistillata*. No repetitive sequence was detected in these IGSs of pocilloporid corals. High levels of nucleotide sequence similarities were also detected among the corresponding IGSs of pocilloporid corals, except for the longest IGS before *cox1*.

The Intergenic Spacer Between *atp6* and *nad4* Contains Features of the Mitochondrial Control Region

Striking length variations were observed in the *atp6-nad4* IGS among different pocilloporid corals (Table 2). The shortest one was 857 bp in *Ma. mirabilis*, the longest one was 1,390 bp in *Poc. damicornis*, and those in *Se. hystrix* and *St. pistillata* were about 1,000~1,100 bp long. This IGS, containing distinct features, tandem repeated elements, conserved sequence blocks (CSBs), conserved sequence domains (CSDs), and presumably functional structures, was identified as a putative control region for pocilloporid corals.

Tandem repeats were detected in *Poc. damicornis*, *Se. hystrix*, and *St. pistillata*, and simple sequence repeats were found in *Ma. mirabilis* (Appendix 2). For *Poc. damicornis*, three copies of 92-bp repeated fragments were recognized (consensus unit: 5′-GAG WAA AAA AAA ACT CTT YMR RAG TTT ASA AWA AAA GTG GGR TRT TTT TYT WGG AGT TTW TTT KTA TGS GTC YGG GRG GAT TWT AAK GAG AG-3′). The location of a tandem repeat was close to *nad4* beginning from a position 991 bp downstream from the 3′ end of *atp6*. The searching strategy of the Tandem Repeats Finder detected five and three copies of tandem repeats in *Se. hystrix* and *St. pistillata*, respectively. The repeated sequences began at the 290- and 460-bp positions downstream from the 3′-end of *atp6* in *Se. hystrix* and *St. pistillata*, respectively. The core element of the tandem repeat of these two genera was 51 bp long, and the nucleotide sequences were highly similar to each other (consensus unit: 5′-BYA GAA AGT AKW GRB RRY TTR AGR GAR DGT GWM RYT ARS GBA TTA MGT SAK-3′). In contrast with the above detections, three copies of incomplete simple repeats were identified in *Ma. mirabilis*. These repeated sequences began at a position 631 bp downstream from the 3′-end of *atp6* with a consensus unit of 5′-YWT YTT G-3′.

Seventeen CSBs, which totaled 556 bp, were recognized and patchily distributed in the aligned non-repeating regions (Appendix 3). Overall CSB sequences were highly conserved among *Poc. damicornis*, *Se. hystrix*, and *St. pistillata* (76.4~84.5% nucleotide similarity), but low (about 55%) between *Ma. mirabilis* and the other three genera. These CSBs ranged 10~137 bp long, and over 50% of them were shorter than 15 bp. Long CSBs were distributed at the 5′-end of this IGS, and most of the short CSBs were observed at the 3′-end and were separated by little disjunctions.

Owing to their patchy distributions, it seemed likely that these CSBs could be arbitrarily classified into three domains (CSD1~3), according to their relative locations (Appendix 3). CSD1, containing two long CSBs, was located at the 5′-end and spanned the positions of 20 to 220 bp of the aligned sequence. CSD2 contained the two longest CSBs and spanned the positions of 716 to 952 bp of the aligned sequence. CSD3 was located at the 3′-end of the aligned sequence and contained six short CSBs.

These CSBs and CSDs were supported by several potential secondary structures. Among them, a high level of topological similarity was also observed among corresponding sequences, especially for those estimated in *Poc. damicornis*, *Se. hystrix*, and *St. pistillata* (Appendices 4~6). For example, similar D-loop structures of a corresponding region which was near the 3′-end of CSD1 were detected in *Se. hystrix* and *St. pistillata*. However, a significant structure was not developed in this region in *Ma. mirabilis* or *Poc. damicornis*. On the contrary, a complex D-loop structure was detected spanning the positions 54~153 and 83~136 bp in *Ma. mirabilis* and *Poc. damicornis*, respectively (Appendix 4). For CSD2, the topology of its simulated structure was more conservative than that of CSD1 (Appendix 5). The conserved structure contained a short hairpin and a D-loop structure which were located at positions between 837 and 934 bp in *Poc. damicornis*, *Se. hystrix*, and *St. pistillata*. The analysis predicted a complex D-loop structure corresponding to positions 797~897 in the alignment of *Ma. mirabilis* (Appendix 5D).

Although CSD3 contained short CSBs, significant structures were developed in *Poc. damicornis*, *Se. hystrix*, and *St. pistillata* (Appendix 6). The analysis predicted an AT-rich hairpin, which corresponded to the position between 1,116 and 1,158 in *Se. hystrix* and *St. pistillata*, and between 1,128 and 1,170 in *Poc. damicornis*. There were five estimated hairpin structures in *Ma. mirabilis*; however, most of the stems were less than 5 bp long and probably had evolved no biological function (data not shown). Among these simulated structures, those at CSD2 were energetically most stable. The free energies of these structures ranged from −29.7 to −35.9, which were significantly lower than those estimated for CSD1 and CSD3 ($p < 0.05$, Kruskal–Wallis test). The most stable structures were specifically associated with the two longest CSBs at CSD2, and the correlation of sequence conservation with the ability to form stable secondary structure strongly implies a biological function.

Discussion

The mt genomes of pocilloporid corals contain three unique features that differ from the published mt genomes of the other scleractinian corals: (1) a novel *atp8*, (2) the putative control region being located between *atp6* and *nad4*, (3) and a duplicated *trnW* in *Seriatopora* and *Stylophora*.

Novel but Diverse *atp8* Gene

The peculiar *atp8* are presumably functional in mt genomes of pocilloporid corals, although both the nucleotide and amino acid sequences are highly different from other scleractinian corals (Fig. 5). The transcribed ORF contained a conserved N-terminal domain and a central positive hydrophobic domain (Fig. 6a). They were more variable at the C-terminal region. Despite their low similarities among other *atp8* in scleractinians, the ORF analyzed here can be unambiguously regarded as a novel *atp8* with a longer C-terminal domain in pocilloporid corals compared to those of other scleractinian corals.

Functionally, *atp8* is an intrinsic membrane protein composed of three domains (Gray et al. 1998). The N-terminal domain, beginning with a conserved Met-Pro-Gln-Leu motif, is located in the intermembrane space and plays a role in ATPase activity (Devenish et al. 1992; Gray et al. 1998). The central domain is composed of a membrane-spanning hydrophobic domain (Devenish et al. 1992). The C-terminal domain is exposed in the matrix space with positively charged amino acids which are involved in both the assembly and function of the F0 sector (Papakonstantinou et al. 1993, 1996). The presence of a unique *atp8* was also discovered in a nematode, *Trichinella spiralis*, and an ascidian, *Ciona intestinalis*, while an *atp8* deficiency was previously considered to be an evolutionary constraint of these organisms (Yokobori et al. 1999; Le et al. 2000; Hu et al. 2003; He et al. 2005). Analyses of sequence and hydropathic similarities support these organisms' three-domain compilation and their similarities to *atp8* in the limulus and humans (Lavrov and Brown 2001; Gissi et al. 2004).

Putative Control Region Located Between *atp6* and *nad4*

Mitochondrial control region is not well-defined and is highly variable among scleractinian corals in location and sequence divergence. Scleractinians' mt control region was suggested to be featured with repetitive sequences, conserved sequence blocks (CSBs), and the capability to form a hairpin structure (van Oppen et al. 2002). The mt putative control region is suggested to be located between *rns* and *cox3* in the Acroporidae (van Oppen et al. 2002), and between *cob* and *nad2* in *Siderastraea* spp. (Chuang 2006). Using these criteria, a putative control region could also be identified at the location between *cox1* and *trnM* for the *Montastraea anularis* complex, between *cox3 and cox2* for *Mussa angulosa*, between *cox3* and *rns* for *Agaricia humilis*, but could not confidently be identified in *Astrangia* sp., *Colpophyllia natans*, *Pavona clavus*, or *Porites porites* (Table 3).

Based on the same criteria, the IGS between the *atp6* and *nad4* genes (IGS of *atp6–nad4*) could be identified as a potential candidate for the mt control region in pocilloporid corals, which is also unique among scleractinian mt genomes (Table 3). Nevertheless, our prediction is discordant with the proposed control regions in two recently published *Pocillopora* mt genomes (Flot and Tillier 2007). In contrast, the region we referred to as the putative control region was annotated as a novel ORF with unknown function by Flot and Tillier (2007). This contradictory prediction should be reexamined and clarified. First, an announcement of an unknown ORF should be viewed with caution, as it indicates the expression of an unknown protein. In fact, several unidentified ORFs were also mentioned in the mt genome of *Ac. tenuis*; the longest one was even located in its putative control region (van Oppen et al. 2002). However, the announcement of an unknown ORF with additional expression confirm, by using assays of RT-PCR or Northern blot analysis, was suggested safer [e.g., Beagley et al. (1996), and van Oppen et al. (2002) conducted for the intron-spliced protein-coding genes, and we conducted for the *atp8*] rather than merely predicting it by a computer program and translating it into amino acids (Majoros 2007). Second, in addition to the conserved ORF announced by Flot and Tillier (2007), unidentified ORFs were also found in the recognized

Table 3 Mitochondrial genomes of scleractinian corals

Species	Conventional grouping[a]	Genome size (bp)	A+T (%)	Protein no.	tRNA no.	rRNA no.	Putative control region (criteria)[b]	Tandem repeat (motif bp x no.)	GenBank No. (reference)
Pocilloporid									
Madracis mirabilis	Astrocoeniina	16,951	68.4	13	2	2	*atp6-nad4* (F, L)	8×3	EU400212 (F)
Pocillopora damicornis	Astrocoeniina	17,425	69.8	13	2	2	*atp6-nad4* (F, L, T)	92×3	EU400213 (F)
Seriatopora hystrix	Astrocoeniina	17,059	69.9	13	3	2	*atp6-nad4* (F, L, T)	51×5	EF633600 (E)
Stylophora pistillata	Astrocoeniina	17,177	70.2	13	3	2	*atp6-nad4* (F, L, T)	51×3	EU400214 (F)
Robust clade									
Astrangia sp.	Faviina	14,853	68.1	13	2	2	*cob-nad2* (L)	N.A.	NC008161 (D)
Colpophyllia natans	Faviina	16,906	66.4	13	2	2	*cox3-cox2* (L)	N.A.	NC008162 (D)
Montastraea annularis complex[c]	Faviina	16,138	66.4	13	2	2	*cox1-trnM* (L, T)	40×2	NC007224-26 (B)
Mussa angulosa	Faviina	17,245	66.3	13	2	2	*cox3-cox2* (T)	66×2	NC008163 (D)
Complex clade									
Acropora tenuis	Astrocoeniina	18,338	62.0	13	2	2	*rns-cox3* (F, L, T)	107×2	NC003522 (A)
Agaricia humilis	Fungiina	18,735	59.6	13	2	2	*rns-cox3* (T)	357×2	NC008160 (D)
Anacropora matthai	Astrocoeniina	17,888	61.6	13	2	2	*rns-cox3* (L, T)	22×2	NC006898 (C)
Montipora cactus	Astrocoeniina	17,887	61.6	13	2	2	*rns-cox3* (L, T)	22×2	NC006902 (C)
Pavona clavus	Fungiina	18,315	59.5	13	2	2	*cox1-trnM* (L)	N.A.	NC008165 (D)
Porites porites	Fungiina	18,648	63.7	13	2	2	*cob-nad2* (L)	N.A.	NC008166 (D)
Siderastrea radians	Fungiina	19,387	63.1	13	2	2	*cob-nad2* (L, T)	31×2	NC008167 (D)

Genome size, numbers of protein-coding, tRNA, and rRNA genes, and GenBank accession numbers
N.A. Not available, Reference: *A* van Oppen et al. (2002), *B* Fukami and Knowlton (2005), *C* Tseng et al. (2005), *D* Medina et al. (2006), *E* Chen et al. (2008), *F* present study
[a] According to the taxonomic classification of Veron (1995) and Veron and Pichon (1976)
[b] Definitive criteria of putative control region: *F* Functional secondary structure, *L* largest intergenic spacer, *T* presence of tandem repeats
[c] Species complex including three sibling species of *Montastraea annularis*, *Monta. franksi*, and *Monta. faveolata*

primary control region of the *Pocillopora* mt genome (IGS of *atp8-cox1*), and in the corresponding regions of the *Seriatopora* and *Stylophora* mt genomes (Chen et al. 2006). But, it is not clear why these ORFs were not mentioned by Flot and Tillier (2007) and how they affected the annotation of the primary control region. Third, due to the function as the origin of replication and transcription, the mt control region mainly features secondary structures, especially those shaped by complementary tandem repeats (Clayton 1992; Casane et al. 1994; Wilkinson et al. 1997; Saito et al. 2005). However, the length of the IGS was a major concern for Flot and Tillier (2007) when defining the primary putative control region of *Pocillopora* corals, even though tandem repeats and secondary structures were present elsewhere (i.e., the IGS of *atp6-nad4*).

On the contrary, our prediction using tandem repeats, conserved sequence blocks, conserved sequence domains, and functional structures as criteria clearly demonstrated all the necessary characteristics for determining a putative control region, not only in the *Pocillopora*, but this was also applicable to the other genera in the family Pocilloporidae. In addition, the highest intraspecific genetic divergence was detected at the mt IGS of *atp6-nad4* of *Poc. damicornis* in our further analyses (data not shown). The comparisons revealed eight variable nucleotide sites at the IGS of *atp6-nad4*, and only 0~2 variable nucleotide sites at the IGS of *atp8-cox1* and the other protein-coding genes between Taiwan and New Caledonia *Poc. damicornis* mt genomes. The results also implied a putative mt control region at the IGS of *atp6-nad4* for pocilloporid corals and highlighted its application in investigating recently divergences for pocilloporid corals (e.g., Chen et al. 2008). Otherwise, the discordance between our estimate and that of Flot and Tillier (2007) implies the possibility of different mt-replication and -transcription mechanisms for anthozoans and closely related taxa from those of higher animals (van Oppen et al. 2002).

The observed differences in the mt control regions of diverse scleractinians may reflect their multiple origins. If the control region had been rearranged after the origin of scleractinian corals, they would be similar to each other in both nucleotide sequences and structural compositions. However, high similarities in nucleotide sequences and in molecular organization were only detected among congeners or genera of the same family (e.g., within the Acroporidae or Pocillloporidae), but these did not hold true among families (Table 3). In addition, no rearrangement or recombination of genes has been reported among the published scleractinian mt genomes. Hence, the most parsimonious explanation for disparities in the control

region is that it evolved independently among different lineages of scleractinians.

Duplication of *trnW* in *Seriatopora* and *Stylophora*

Duplication of an additional tRNA gene has been reported in other metazoan mt genomes. Two isoacceptors of *trnM* have been mentioned in the mt genomes of the mussels, *Mytilus edulis* and *My. californianus* (Hoffmann et al. 1992; Beagley et al. 1999), and the ascidian, *Ciona intestinalis* (Gissi et al. 2004). In *My. edulis*, similarities at the 5′ end of the two *trnM* genes suggest that they arose by gene duplication, and the additional *trnM* is thought to originate from its isoacceptor and functionally differs from it (Hoffmann et al. 1992). It has also been suggested that *trnM*(AUG) and *trnM*(AUA) were respectively used as an initiator and elongator, although there is no experimental evidence to support this hypothesis (Beagley et al. 1999).

Being intermediate to the high similarities between the two *trnW* genes respectively identified in *St. pistillata* and *Se. hystrix* suggests their origin by a gene duplication event in their common ancestor. By folding *trnW*', the difference of the nucleotides does not change the secondary structure of the loop region, which indicates that *trnW*' genes are functional in *Se. hystrix* (Chen et al. 2008). The loss of the TψC loop (T) from *trnW*' of *St. pistillata* and the mispairings in its aminoacyl stem may result from a deletion and/or the naturally high substitution of a tRNA gene. Some animal mt tRNAs have lost the DHU (D-) or T-stems, forming reduced secondary structures with replacement loops instead of stems (Wolstenholme 1992). All known examples of T-stem shifting to replacement loops are in nematodes (Wolstenholme et al. 1987). Lack of the D-stem have been reported for diverse organisms, such as nematodes (Wolstenholme et al. 1987), mollusks (Boore and Brown 1994), annelids (Boore and Brown 1995), marsupials (Janke et al. 1994), frogs (Wolstenholme et al. 1987), and spiders (Masta 2000). Due to the required function of the aminoacyl and anticodon stems, losses of the D-and T stems do not inhibit their functions. Thus, the *trnW* duplication is supposed to occur at the recently common ancestor of *Seriatopora* and *Stylophora*, but deletion or generally reduced in sequences and structure in *Stylophora*. The additional *trnW* is supported by a postulated model of tRNA molecule replication which has been hypothesized that their origin is simply by direct duplication of a molecule housing double-hairpin structures (Di Giulio 1992, 2004). However, these *trnW*' in *Se. hystrix* and *St. pistillata* are coded by the same 5′-UCA-3′ anticodon, and no supplemental function can be deduced. In addition, tRNA duplication was probably a single derived event in the common ancestor of *Se. hystrix and St. pistillata* in the family Pocilloporidae, as it was not found in *Pocillopora*, *Madracis*, or other scleractinians (Table 3).

General Comparisons Among mt Genomes of Scleractinians

In addition to these unique features described previously, the nucleotide composition and codon usages of the mt genomes of pocilloporids significantly differ from those of complex-clade scleractinians and somewhat differ from those of robust-clade scleractinians (Table 3).

For example, the nucleotide compositions of scleractinian mt genomes are A+T rich biased and range 59.5~70.2% (Table 3). Values of the A+T composition significantly differ among pocilloporids, complex-clade and robust-clade corals ($p<0.01$, Kruskal–Wallis test). An A+T rich bias is highest for pocillporid corals, which is significantly higher than those of complex- and robust-clade clades ($p<0.01$, = 0.02, Mann–Whitney *U* test for each comparison).

The mt genomes of scleractinian corals range from 14.5 to 19.4 kb in length, and there are significant differences among complex-clade, robust-clade, and pocilloporid corals ($p<0.01$, Kruskal–Wallis test; Table 3). The mt genomes of pocilloporid corals are significantly smaller than those of complex-clade corals ($p<0.01$, Mann–Whitney *U* test), but do not differ from those of robust-clade corals ($p=0.15$, Mann–Whitney *U* test).

Translation initiation and termination for mt protein-coding genes of pocilloporids are similar to those of robust-clade corals (data not shown). Minor differences between pocilloporid and robust-clade corals are observed at the initiation of *cob*, and termini of *cox2*, *nad1*, and *nad5*. However, the consequence of star/stop codon usage of pocilloporid corals contrasts with the majority start/stop codons of complex-clade corals, where start codons (AUG and GUG) and stop codons (UAA and UAG) are equally in use.

In conclusion, a novel *atp8* gene, the nucleotide and amino acid sequences of which are diverse in comparison to published counterparts in scleractinian mitochondrial genomes, was identified and confirmed in the family Pocilloporidae using RT-PCR and Northern blot analyses. The *atp6-nad4* IGS, containing distinct repeated elements, conserved sequence blocks and domains, and functional structures, possesses the typical characteristics of a putative control region of the family Pocilloporidae. Duplicated *trnW*, detected in the region close to the *cox1* gene and which shares the highly conserved primary and secondary structures of its original counterpart, was discovered in both *Seriatopora* and *Stylophora*. These molecular characteristics are unique and phylogenetically informative for future evaluations of the status of the family Pocilloporidae in the evolutionary history of scleractinian corals.

Acknowledgements We thank Drs. Hironobu Fukami and Nancy Knowlton for initiating this study and their encouragement, members of the Evolutionary and Ecology Genetics of Coral Reef Discussion Group, Research Center for Biodiversity, Academia Sinica and two anonymous referees for constructive comments. This work was supported by Academia Sinica Thematic Grants (2002–2004), a grant from the Genomic Research Center, Academia Sinica (2006–2007), and a National Science Council (NSC) grant (NSC94-2621-B-001-005), Taiwan to C.A.C., and an NSC grant (NCS94-EPA-Z002-004) to C.-F. Dai. This is the Evolutionary and Ecology Genetics of Coral Reef Group, RCBAS contribution no. 44.

References

Azuma N, Kunihiro Y, Sasaki J, Mihara E, Mihara Y, Yasunaga T, Jin D-H, Abe S (2008) Genetic variation and population structure of hair crab (*Erimacrus isenbeckii*) in Japan inferred from mitochondrial DNA sequence analysis. Mar Biotechnol 10:39–48

Beagley CT, Okada NA, Wolstenholme DR (1996) Two mitochondiral group 1 introns in a metazoan, the sea anemone *Metridium senile*: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. Proc Natl Acad Sci USA 93:5619–5623

Beagley CT, Okimoto R, Wolstenholme DR (1998) The mitochondrial genome of the sea anemone *Metridium senile* (Cnidaria): introns, a paucity of tRNA genes, a near-standard genetic code. Genetics 148:1091–1108

Beagley CT, Okimoto R, Wolstenholme DR (1999) *Mytilus* mitochondrial DNA contains a functional gene for a $tRNA^{Ser}$(UCN) with a dihydrouridine arm-replacement loop and a pseudo-$tRNA^{Ser}$(UCN) gene. Genetics 152:573–652

Beaton MJ, Roger AJ, Cavalier-Smith T (1998) Sequence analysis of the mitochondrial genome of *Sarcophyton glaucum*: conserved gene order among octocorals. J Mol Evol 47:697–708

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580

Blair D, Waycott M, Byrne L, Dunshea G, Smith-Keune C, Neil KM (2006) Molecular discrimination of *Perna* (Mollusca: Bivalvia) species using the polymerase chain reaction and species-specific mitochondrial primers. Mar Biotechnol 8:380–385

Boore JL (1999) Animal mitochondrial genomes. Nucleic Acids Res 27:1767–1780

Boore JL, Brown WM (1994) Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. Genetics 138:423–443

Boore JL, Brown WM (1995) Complete sequence of the mitochondrial DNA of the annelid worm, *Lumbricus terrestris*. Genetics 141:305–319

Bridge D, Cunningham CW, Schierwater B, DeSalle R, Buss LW (1992) Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. Proc Natl Acad Sci USA 89:8750–8753

Brugler MR, France SC (2007) The complete mitochondrial genome of the black coral *Chrysopathes formosa* (Cnidaria:Anthozoa: Antipatharia) supports classification of antipatharians within the subclass Hexacorallia. Mol Phylogenet Evol 42:776–788

Burger G, Gray MW, Lang BF (2003) Mitochondrial genomes: anything goes. Trends Ecol Evol 19:709–716

Burridge CP, Versace VL (2007) Population genetic structuring in *Acanthopagrus butcheri* (Pisces: Sparidae): does low gene flow among estuaries apply to both sexes? Mar Biotechnol 9:33–44

Casane D, Dennebouy N, De-Rochambeau H, Mounolou JC, Monnerot M (1994) Genetic analysis of systematic mitochondrial heterplasmy in rabbits. Genetics 138:471–480

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540–552

Chen CA, Yu J-K (2000) Universal primers for amplification of mitochondrial small subunit ribosomal RNA-encoding gene in scleractinian corals. Mar Biotechnol 2:146–153

Chen CA, Wallace CC, Wolstenholme J (2002) Analysis of the mitochondrial 12S rRNA gene supports a two-clade hypothesis of the evolutionary history of scleractinian corals. Mol Phylogenet Evol 23:137–149

Chen C, Dai C-F, Fukami H, Knowlton N, Chen CA (2006) Evolutionary mitogenomics of pocilloporid corals: ATP8 gene loss, a novel open reading frame (ORF) and phylogenetic utility. First Asia Pacific Coral Reef Symposium, Hong Kong, p M39

Chen C, Dai C-F, Plathong S, Chiou C-Y, Chen CA (2008) The complete mitochondrial genomes of needle corals, *Seriatopora* spp. (Scleractinian; Pocilloporidae): an idiosyncratic *atp8*, duplicated *trnW* gene, and hypervariable regions used to determine species phylogenies and recently diverged populations. Mol Phylogenet Evol 46:19–33

Cho E-S, Jung C-G, Sohn S-G, Kim C-W, Han S-J (2007) Population genetic structure of the ark shell *Scapharca broughtonii* Schrenck from Korea, China, and Russia based on COI gene sequences. Mar Biotechnol 9:203–216

Chuang Y-Y (2006) Mitogenomics and molecular evolution of the group I intron in the cytochrome oxidase I gene of *Siderastrea* (Cnidaria; Scleractinia; Siderastreidae). Institute of Oceanography. National Taiwan University, Taipei, Taiwan

Clayton DA (1991) Replication and transcription of animal mitochondrial DNA. Annu Rev Cell Biol 7:453–478

Clayton DA (1992) Transcription and replication of animal mitochondrial DNAs. Int Rev Cytol 141:217–232

Devenish RJ, Papakonstantinou T, Galanis M, Law RH, Linnane AW, Nagley P (1992) Structure/function analysis of yeast mitochondrial ATP synthase subunit 8. Ann NY Acad Sci 671:403–414

Di Giulio M (1992) On the origin of the transfer RNA molecule. J Theor Biol 159:199–214

Di Giulio M (2004) The origin of the tRNA molecule: implications for the origin of protein synthesis. J Theor Biol 226:89–93

Flot J-F, Tillier S (2007) The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: the putative D-loop and a novel ORF of unknown function. Gene 401:80–87

Fukami H, Knowlton N (2005) Analysis of complete mitochondrial DNA sequences of three members of the *Montastraea annularis* coral species complex (Cnidaria, Anthozoa, Scleractinia). Coral Reefs 24:410–417

Fukami H, Budd AF, Levitan DR, Jara J, Kersanach R, Knowlton N (2004) Geographic differences in species boundaries among members of the *Montastraea annularis* complex based on molecular and morphological markers. Evolution 58:324–337

Gish W, States DJ (1993) Identification of protein coding regions by database similarity search. Nat Genet 3:266–272

Gissi C, Pesole G (2003) Transcript mapping and genome annotation of ascidian mtDNA using EST data. Genome Res 13:2203–2212

Gissi C, Iannelli F, Pesole G (2004) Complete mtDNA of *Ciona intestinalis* reveals extensive gene rearrangement and the presence of an *atp8* and an extra *trnM* gene in ascidians. J Mol Biol 58:376–389

Glynn PW (1976) Some physical and biological determinants of coral community structure in the eastern Pacific. Ecol Monogr 46:431–456

Gonnet GH, Cohen MA, Benner SA (1992) Exhaustive matching of the entire protein sequence database. Science 256:1443–1445

Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint-Louis D, Zhu Y, Burger G (1998) Genome

structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 26:865–878

Grigg RW, Maragos JE (1974) Recolonization of hermatypic corals on submerged larva flows in Hawaii. Ecology 55:387–395

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

He Y, Jones J, Armstrong M, Lamberti F, Moens M (2005) The mitochondrial genome of *Xiphinema americanum sensu stricto* (Nematoda: Enoplea): considerable economization in the length and structural features of encoded genes. J Mol Biol 61:819–883

Hellberg ME (2006) No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. BMC Evol Biol 6:24

Hoffmann RJ, Boore JL, Brown WM (1992) A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. Genetics 131:397–412

Hu M, Chilton NB, Gasser RB (2003) The mitochondrial genome of *Strongyloides stercoralis* (Nematoda)-idiosyncratic gene order and evolutionary implications. Int J Parasitol 33:1393–1408

Ikeguchi K, Ineno T, Itoi S, Kondo H, Kinoshita S, Watabe S (2006) Increased levels of mitochondrial gene transcripts in the thermally selected rainbow trout (*Oncorhynchus mykiss*) strain during embryonic development. Mar Biotechnol 8:178–188

Janke A, Feldmaier-Fuchs G, Thomas WK, von Haeseler A, Paabo S (1994) The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137:243–256

Kayal E, Lavrov DV (2008) The mitochondrial genome of *Hydra oligactis* (Cnidaria, Hydrozoa) sheds new light on animal mtDNA evolution and cnidarian phylogeny. Gene 410:177–186

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. J Mol Biol 157:105–142

Lavrov DV, Brown WM (2001) *Trichinella spiralis* mtDNA: a nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAs and has a gene arrangement relatable to those of coelomate metazoans. Genetics 157:621–637

Le TH, Blair D, Agatsuma T, Humair P-F, Campbell NJH, Iwagami M, Littlewood DTJ, Peacock B, Johnson DA, Bartley J, Rollinson D, Herniou EA, Zarlenga DS, McManus DP (2000) Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. Mol Biol Evol 17:1123–1125

Lin G, Lo LC, Zhu ZY, Feng F, Chou R, Yue GH (2006) The complete mitochondrial genome sequence and characterization of single-nucleotide polymorphisms in the control region of the Asian seabass (*Lates calcarifer*). Mar Biotechnol 8:71–79

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Loya Y (1976) The Red Sea coral *Stylophora pistillata* is an *r* strategist. Nature 259:478–480

Mahidol C, Na-Nakorn U, Sukmanomon S, Taniguchi N, Nguyen TTT (2007) Mitochondrial DNA diversity of the Asian moon scallop, *Amusium pleuronectes* (Pectinidae), in Thailand. Mar Biotechnol 9:352–359

Majoros WH (2007) Methods for computational gene prediction. Cambridge Univ. Press, Cambridge

Masta SE (2000) Mitochondrial sequence evolution in spiders: intraspecific variation in tRNAs lacking the TYC arm. Mol Biol Evol 17:1091–1100

Medina M, Collins AG, Takaoka TL, Kuehl JV, Boore JL (2006) Naked corals: skeleton loss in scleractinia. Proc Natl Acad Sci USA 103:9096–9100

Milbury CA, Gaffiney PM (2005) Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*. Mar Biotechnol 7:697–612

Moriya S, Sato S, Azumaya T, Suzuki O, Urawa S, Urano A, Abe S (2007) Genetic stock identification of chum salmon in the Bering Sea and North Pacific Ocean using mitochondrial DNA microarray. Mar Biotechnol 9:179–191

Papakonstantinou T, Galanis M, Nagley P, Devenish RJ (1993) Each of three positively-charged amino acids in the C-terminal region of yeast mitochondrial ATP synthase subunit 8 is required for assembly. Biochim Biophys Acta 1144:22–32

Papakonstantinou T, Law RH, Nagley P, Devenish RJ (1996) Non-functional variants of yeast mitochondrial ATP synthase subunit 8 that assemble into the complex. Biochem Mol Biol Int 39:253–260

Pont-Kingdon GA, Okada NA, Macfarlane JL, Beagley CT, Wolstenholme DR (1995) A coral mitochondrial *mutS* gene. Nature 375:109–111

Pont-Kingdon G, Okada NA, Macfarlane JL, Beagley CT, Watkins-Sims CD, Cavalier-Smith T, Clark-Walker GD, Wolstenholme DR (1998) Mitochondrial DNA of the coral *Sarcophyton glaucum* contains a gene for a homologue of bacterial MutS: a possible case of gene transfer from the nucleus to the mitochondrion. J Mol Evol 46:419–431

Romano SL, Palumbi SR (1996) Evolution of scleractinian corals inferred from molecular systematics. Science 271:640–642

Romano SL, Palumbi SR (1997) Molecular evolution of a portion of the mitochondrial 16S ribosomal gene region in scleractinian corals. J Mol Evol 45:397–411

Saito S, Tamura K, Aotsuka T (2005) Replication origin of mitochondrial DNA in insects. Genetics 171:1695–1705

Shadel GS, Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. Annu Rev Biochem 66:409–435

Shao Z, Graf S, Chaga OY, Lavrov DV (2006) Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. Gene 381:92–101

Shearer TL, van Oppen MJH, Romano SL, Worheide G (2002) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). Mol Ecol 11:2475–2487

Sinniger F, Chevaldonné P, Pawlowski J (2007) Mitochondrial genome of *Savalia savaglia* (Cnidaria, Hexacorallia) and early metazoan phylogeny. J Mol Evol 64:196–203

Stolarski J, Roniewicz E (2001) Towards a new synthesis of evolutionary relationships and classification of scleractinia. J Paleontol 75:1090–1108

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tseng C-C, Wallace CC, Chen CA (2005) Mitogenomic analysis of *Montipora cactus* and *Anacropora matthai* (cnidaria; scleractinia; acroporidae) indicates an unequal rate of mitochondrial evolution among Acroporidae corals. Coral Reefs 2005:502–508

van Oppen MJH, Willis BL, Miller DJ (1999) Atypical low rate of cytochrome b evolution in the scleractinian coral genus *Acropora*. Proc R Soc Lond B Biol Sci 266:179–183

van Oppen MJH, Catmull J, McDonald BJ, Hislop NR, Hagerman PJ, Miller DJ (2002) The mitochondrial genome of *Acropora tenuis* (Cnidaria; Scleractinia) contains a large group 1 intron and a candidate control region. J Mol Evol 55:1–13

Vaughan TW, Wells JW (1943) Revision of the suborders, families, and genera of the Scleractinia. Geol Soc Am Spec Pap 44:1–363

Veron JEN (1995) Corals in space and time: The biogeography and evolution of the scleractinian. University of new South Wales Press, Sydney

Veron JEN (2000) Corals of the world. Australian Institute of Marine Science, Queensland, Australia

Veron JEN, Pichon M (1976) Scleractinia of Eastern Australia. Part I. Thamnasteriidae, Astrocoeniidae and Pocilloporidae. Aust Inst Mar Sci Monogr 1:1–86

Wallace CC (1985) Seasonal peaks and annual fluctuations in recruitment of juvenile scleractinian corals. Mar Ecol Prog Ser 21:289–298

Wells JW (1956) Scleractinia. In: Moore RC (ed) Treatise on invertebrate paleontolohy, coelenterata. Geological Society of America and University of Kansas Press, pp 328–440

Wilkinson GS, Mayer F, Kerth G, Petri B (1997) Evolution of repeated sequence arrays in the D-loop region of bat mitochondrial DNA. Genetics 146:1035–1048

Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution. Int Rev Cytol 141:173–216

Wolstenholme DR, MacFarlane JL, Okimoto R, Clary DO, Wahleithner JA (1987) Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. Proc Natl Acad Sci USA 84:1324–1328

Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252–3253

Yokobori S-i, Ueda T, Feldmaier-Fuchs G, Pääbo S, Ueshima R, Kondow A, Nishikawa K, Watanabe K (1999) Complete DNA sequence of the mitochondrial genome of the ascidian *Halocynthia roretzi* (Chordata, Urochordata). Genetics 153:1851–1862

Zuker M (2000) Calculating nucleic acid secondary structure. Curr Opin Struct Biol 10:303–310