The complete mitochondrial genomes of needle corals, *Seriatopora* spp. (Scleractinia: Pocilloporidae): An idiosyncratic *atp8*, duplicated *trnW* gene, and hypervariable regions used to determine species phylogenies and recently diverged populations

Chienhsun Chen a,b, Chang-Feng Dai a, Sakanan Plathong c, Chih-Yung Chiou b, Chaolun Allen Chen a,b,*

a Institute of Oceanography, National Taiwan University, Taipei 106, Taiwan
b Research Center for Biodiversity, Academia Sinica, Nankang, Taipei 11529, Taiwan
c Department of Biology, Prince of Songkla University, Hat Yai, Songkla, Thailand

Received 29 November 2006; revised 11 September 2007; accepted 16 September 2007
Available online 21 September 2007

Abstract

Complete DNA sequences were determined for the mitochondrial (mt) genomes of the needle corals, *Seriatopora caliendrum* (17,011 bp) and *S. hystrix* (17,060 bp). Gene arrangement of the *Seriatopora* mt genomes is similar to the 14 currently published scleractinian mitogenomes with three unusual features, including an idiosyncratic *atp8*, a duplicated *trnW* (*tRNA* TRP), and a putative control region located between *atp6* and *nad4*. *Atp8*, located between duplicate *trnW* genes, showed relatively low amino acid similarity (25.6–34.6%) with those of published scleractinian corals. A reverse-transcription polymerase chain reaction confirmed the transcription of this novel *atp8* gene in *Seriatopora*. A duplicated *trnW* was detected in the region close to the *cox1* gene and shares the highly conserved primary and secondary structure of its original counterpart. The intergenic spacer between *atp6* and *nad4*, which contains several distinct repeated elements, is being designated as the putative control region in the *Seriatopora* mt genomes. Evaluation of the molecular evolution of several protein-coding genes and intergenic spacers showed 3- to 4-fold higher divergence rates among populations or between species than those published for scleractinian mt genomes. This study not only successfully revealed the phylogenies of *S. hystrix* and *S. caliendrum* from the West Pacific Ocean by mtDNA, but also highlighted the potential utilities of mt hypervariable regions in phylogenetic construction below the species level for *Seriatopora*.

© 2007 Published by Elsevier Inc.

Keywords: Complete mitochondrial genomes; *Seriatopora*; Idiosyncratic ATP8; Control region; tRNA duplication; Scleractinia

1. Introduction

In the last decade, attempts to sequence complete mitochondrial (mt) genomes from diverse taxa have challenged the concept of mitogenomic evolution of several lower invertebrates (Boore, 1999; Burger et al., 2003; Dellaporta et al., 2006). Increasing evidence has shown that the mt genomes of many invertebrates do not agree with the standard knowledge of mitochondrial characterization (reviewed in Wolstenholme, 1992). For example, mt genomes of scleractinian corals include extensive intergenic spacers between genes, only two of the typical 22 *tRNA* genes, and contain a group I intron that bisects the NADH dehydrogenase subunit 5 gene (*nad5*) (Fukami and Knowlton, 2005; Medina et al., 2006; Tseng et al., 2005; van Oppen et al., 2002). A smaller subset of Scleractinia also contain a group I intron in the cytochrome oxidase I gene
(coxI), which may contain signals indicative of a homing endonuclease gene (Fukami et al., 2007; Goddard et al., 2006). Despite these unusual features, evolutionary rates in corals or anthozoans in general are considered to be relatively slow in comparison to those of higher animals (Hellberg, 2006; Shearer et al., 2002). A phylogenetic analysis showed that the evolutionary rate of the Acropora cytochrome b gene (cob) is 10–20 times slower than the standard molecular clock rate of 2% per million years (Ma−1) (van Oppen et al., 1999). Among sibling species of the Montastraea annularis complex, the evolutionary rate of their mitogenomes is only 0.03% Ma−1 (Fukami and Knowlton, 2005). This value corresponds to only 25 variable sites across 16,134 bp, 16 of which occurred in one of the three species analysed. Fossil record data places the divergence of the Montastraea annularis complex at approximately 3–4 million years ago (Fukami and Knowlton, 2005; Pandolfi et al., 2002). Most surprisingly, Tseng et al. (2005) noted invariance in DNA sequences between corresponding mt protein-coding genes of Anancorpora matthai and Montipora cactus (0–0.91% differences in p-distances). The most divergent mt region (the 3rd intergenic spacer) between these two genera of the Acroporidae was only 1.12% in pairwise distances for over 30 million years of divergence time (Frost, 1977; Gregory and Trench, 1916; Tseng et al., 2005).

The observed slow evolution may limit the applications of mtDNA sequences to the elucidation of species phylogenies and population genetics of scleractinian corals. Previous studies of population genetics in coding regions of mtDNA revealed little to no variation among conspecifics; even geographically distant or potentially isolated populations are invariant (Hellberg, 2006; Medina et al., 1999; Snell, 2000; Snell et al., 1998). For example, there was no variation within coxI among populations over 3000 km apart in the range of the coral, Ballanophyllia elegans (Hellberg, 2006). Watanabe et al. (2005) examined the polymorphisms of the cob-nad2 intergenic spacer in Galaxea fascicularis in Ryukyu Archipelago, and observed significant differences in the haplotype frequencies among three sampling sites. Their result, however, may be explained by sympatric occurrence of two cryptic species rather than genetic differentiation among the three populations. Whether this conclusion is applicable to all scleractinian coral remains uncertain, since patterns of evolutionary rates can vary greatly among genes or regions across different coral lineages. Further evaluation of complete scleractinian mt genomes is needed before we can determine which mtDNA sequences will be useful for species and/or population-level phylogenies.

1.1. Study organism

The genus Seriatopora is a member of the Pocilloporidae, and known for wide variations in colony morphologies (Veron and Pichon, 1976). Two major species groups, or six valid species, are recognized according to differences in colony morphology and branch diameter, respectively (Veron, 2000; Veron and Pichon, 1976). In this study, the complete mt genome sequences of S. hystrix and S. caliendrum were characterized and compared to those of Acropora tenuis and the Montastraea annularis complex. Different mtDNA regions were evaluated by analysing variations and divergences within and between populations of the same species and by comparisons between two sister Seriatopora species.

2. Materials and methods

2.1. Sample collection and DNA extraction

Both Seriatopora caliendrum and S. hystrix were collected from Tiaoshi (21°57’27”N; 120°45’56”E), southern Taiwan. Apical fragments (1–2 cm long) of corals were narcotized in calcium-free seawater, and the attached epifauna were removed by fine tweezers under a dissecting microscope within 2 h to prevent further coral tissue dissociation (Domart-Coulon et al., 2004, 2001), before preserving in 70% ethanol.

Total genomic DNA of the preserved sample was extracted with an optimized protocol by including CTAB (Cetyl trimethylammonium bromide) to remove mucus and by modification of the standard phenol-chloroform extraction techniques (Sambrook and Russell, 2001). The terminal branch (0.5–1 cm) of the preserved sample was ground into powder under liquid nitrogen with a mortar and pestle. The powder was lysed with DNA isolation buffer, containing 400 mM NaCl, 1% (w/v) sodium dodecyl sulphate SDS and 50 mM EDTA, pH 8, for 1 h at 65°C, and digested with proteinase K (500 μg/ml) at 55°C for 6–8 h. RNase was added to a final concentration of 20 μg/ml and the lysate was incubated at 37°C for 30 min in order to eliminate any RNA molecules. Subsequently, the lysate was treated with 800 mM NaCl and 1% (w/v) CTAB and incubated at 65°C for 30 min. After incubation, solution was extracted once by adding an equal volume of chloroform and its supernatant was extracted once by adding an equal volume of chloroform and its supernatant was precipitated by adding an equal volume of phenol–chloroform–isoamyl alcohol (25:24:1). The DNA of the supernatant from the 2nd 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-distilled water.

2.2. Molecular protocols

The complete mt genome was amplified by means of two long overlapping PCR reactions (Cheng et al., 1994) from the total DNA of both S. caliendrum and S. hystrix. To accomplish this, partial sequences of rns and rnl were initially obtained by a standard polymerase chain reaction (PCR) using coral-specific primers (Chen and Yu, 2000;
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 CGA ACA GCC AAG TC-3' (R21), and 5'-TAC CCT GGG GAT AAC AGC GCA ATA ACG-3' (F02) and 5'-AAAG GCCCAA TAA CCT TCC ATT GCA TCC GGT AGC-3' (R06), respectively. Long PCRs were performed using the Long PCR Enzyme Mix (Fermentas) under the conditions recommended by the manufacturer, in a PTC-200 Thermocycler (MJ Research). Long PCR reactions were setup in a volume of 50 μl: 1× PCR buffer, 1.5 mM MgCl₂, 0.2–0.4 mM of each dNTP, 2.5 mM MgCl₂, 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, 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.

2.3. Sequence analysis

The DNA sequences were assembled using Sequencher 4.2 (Gene Codes Corporation), then analysed in Vector NTI 6.0 (InforMax, Invitrogen life science software). Open reading frames (ORFs) of considerable length (>50 amino acids) in the sequences were initially translated using cnidarian mt genetic codes (NCBI translation Table 4), then compared with the databases using the BLASTX program (Beaglely et al., 1998; Gish and States, 1993). The identified protein coding and tRNA genes were aligned with the corresponding genes of previously published scleractinian mitogenomes for final recognition (Fukami and Knowlton, 2005; Tseng et al., 2005; van Oppen et al., 2002), using MEGA v3.1 (Kumar et al., 2004) with the weighted matrix of ClustalW (Thompson et al., 1994) and Gonnet (Gonnet et al., 1992). In addition to comparisons of the nucleotide and amino acid sequences, the hydrophathy profiles of the unassigned ORFs were also compared with those of the identified atp8 of scleractinian corals (Kyte and Doolittle, 1982). In order to assess the stability of the potential secondary structures of the intergenic spacers, free energies of the predicted secondary structures and their random sequences were examined using the DNA Mfold server (http://bioinfo.math.rpi.edu/~mfold/dna/; Zuker, 2000). Intergenic spacers were also examined for the presence of canonical tRNAs with tRNAscan-SE search server v1.21 (Lowe and Eddy, 1997), 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. The nucleotide composition was calculated using BioEdit 7.01 (Hall, 1999), while tandem repeat sections in the intergenic spacer were identified using the Tandem Repeat Finder 4.0 (Benson, 1999).

2.4. Utility of mtDNA for species phylogeny and population genetics

For evaluating the phylogenetic utility of the mtDNA regions of Seriatopora, an additional total 29 coral colonies were sampled from Taiwan, Taiping Island (Itu Aba Island) in the Spratly Islands (10°43′20″N; 115°49′32″E),
South China Sea, and Similan Island (8°39′17″N; 97°39′85″E), Andaman Sea, Thailand (Fig. 1, Table 1).

Specimens sampled from different oceans were arbitrarily considered to be different populations, due to the differences in oceanography, geography, and far distances among locations (2000–4000 km, Fig. 1). Amongst them, Taiwan is located in the middle of Ryukyu-Philippine Archipelago; the oceanography of its east side is dominated by the northward Kuroshio Current. The South China Sea is a marginal sea of the Pacific Ocean, and is semi-enclosed between the West Pacific and the Indian Oceans. Its surface currents are more complex than the nearby oceans and have dramatic seasonal variability driven by the monsoon (Morton and Blackmore, 2001). In winter, an anticlockwise gyre is created by the influence of the Northeast Monsoon; but, a clockwise gyre and a northeasterly current are established by the Southwest Monsoon during the summer (Chern and Wang, 2003; Chu and Guihua, 2003). The Andaman Sea is part of the Indian Ocean and located to the west of Malay Peninsula. Studies of the physical oceanography indicated that the current system of Andaman Sea is also driven by seasonal monsoons; it develops a clockwise and an anticlockwise gyre during the summer and winter, respectively (Varkey et al., 1996). Hence, our sampling represented 1 population of *S. caliendrum* and 3 populations *S. hystrix* corals (Table 1).

These coral samples were preserved in the field in 70% EtOH, and total DNA extractions were completed following the methods described previously. The partial or complete nucleotide sequences of *atp6*, *cox1*, the 9th intergenic spacer (IGS9, an intergenic spacer spanning the region between *trnW* and *cox1*), and the putative control region were chosen for evaluation of their phylogenetic utilities, after genetic variations of each region was evaluated between mt genomes of *Seriatopora* spp. Table 2 illustrates the PCR primers and the optimized annealing temperatures of these PCR reactions. The PCR was setup in a volume of 50 μl consisting of 1x PCR buffer, 2.0–2.5 mM MgCl2, 0.075 mM of each dNTP, 0.17 μm of each primer, 1% DMSO, 1.0 U of HP HotStart Taq (Protech Technology), and 0.5 μg of genomic DNA. The PCR profiles consisted of 1 cycle of 95 °C for 11 min; 35 cycles of 93 °C for 1 min, an optimized annealing temperature for 40 s, and 72 °C for 1–2 min; then 1 cycle of 72 °C for 2–3 min. The PCR products were electrophoresed and inspected in 1.2% agarose gels, and their nucleotide sequences were determined for both strands with the same primers in the PCR amplification or with additional internal primers on an ABI 377 automated DNA sequencer.

The DNA sequences were assembled in Sequencher 4.2 (Gene Codes Corp.), aligned in ClustalW 1.6 (Thompson et al., 1994) then manually edited in MEGA v3.1 (Kumar et al., 2004). The alignment parameters of IGS9 were the same as those of the putative control region. Their gap opening and extension penalties were set to the values of 15 and 6, respectively. Changing the penalty values does

![Fig. 1. Map of the Indo-West Pacific region. Starts and dot lines with arrows indicate sampling locations of *Seriatopora* populations and the directions of major ocean currents in the region, respectively.](image)
not affect the aligning result. The repetitive regions within the putative control region of the mtDNA were excluded from the analysis, because the processes by which these repeats evolve are not well understood for scleractinian corals. Pairwise genetic distances (uncorrected p-distances) of each mtDNA region were calculated between species, and between and within populations of the same species using MEGA v3.1 (Kumar et al., 2004). The Mann–Whitney U-test or Kruskal–Wallis test was applied to examine the differences in p-distances among the three hierarchical categories using Stateview 5.0 (SAS Institute).

Phylogenetic relationships among sequences were constructed on the basis of neighbor-joining (NJ), maximum-parsimony (MP), and maximum-likelihood (ML) methods with the use of PAUP*4.0b10 (Swofford, 1999), and Bayesian (BAY) analysis in the program MrBayes 3.12 (Huelsenbeck and Ronquist, 2001). Similar tree topologies were generated using methods that included and excluded indels. Thus, analyses were run with indels treated as missing data, according to suggestions in Nei and Kumar (2000). The NJ analysis was performed with the Kimura 2-parameter model of nucleotide substitution (Kimura, 1980). In the MP and ML analyses, heuristic searches with TBR branch swapping and 10 random additions of sequences were performed. For the ML and BAY analyses, the best-fitting model of DNA substitution and parameter estimates were performed by the Akaike information criterion (AIC) in PAUP*4.0b10 (Swofford, 1999) and Modeltest 3.7 (Posada and Crandall, 1998). The best-fitting evolutionary models were TrN for atp6 and cox1, K81uf for IGS9 and GTR+G for the putative control region. Five-hundred bootstrapping values were used to evaluate support for the NJ, MP and ML trees. In the BAY analysis, the Markov Chain Monte Carlo search was run with 4 chains for 1·10⁶ generations, with trees sampled every 100 generations. The first 2500 trees were discarded as the “burn-in”, after which the likelihood scores had stabilized.

### 3. Results

#### 3.1. Characterization of the Seriatopora mt genomes

The entire mitochondrial (mt) genomes of Seriatopora caliendrum and S. hystrix were 17,011 and 17,060 bp in length, respectively. These are smaller than those of acroporids (17,887–18,338 bp; Tseng et al., 2005; van Oppen et al., 2002), but larger than those of the Montastraea ann-
The *Seriatopora* mt genomes contained 12 identified open reading frames (ORFs), the small (rns) and large (rnl) subunits of ribosomal RNA genes, and 3 tRNA genes, all of which were transcribed on the same strand (Fig. 2). Except for *atp8*, these identified 12 ORFs included 7 NADH dehydrogenases (*nad*), 3 cytochrome oxidases (*cox*), ATPase (*atp6*), and cytochrome b (*cob*) (Table 3). The arrangements of the protein-coding genes of *Seriatopora* species were similar to those of other scleractinian corals (Fukami and Knowlton, 2005; Medina et al., 2006; Tseng et al., 2005; van Oppen et al., 2002). The *nad5* of both *Seriatopora* species was also interrupted by an intron, which contains 10 protein-coding genes, *rns*, and the putative control region (Fig. 2). The A + T base composition varied among different regions of both *Seriatopora* mitogenomes, ranging from 58% to 86%, with the highest value for the *nad5* intron (86%) and the lowest for the 9th intergenic spacer (IGS9) and the putative control region (Table 3). The mean A + T content of mt genomes of both *Seriatopora* species was significantly higher than those of other scleractinian corals (*p* < 0.01, Chi-square test; 60% and 66% for acroporids and the *M. annularis* complex, respectively).

### 3.2. Codon usage

Thirty-eight hundred and twenty-three amino acids were encoded for the 12 protein-coding genes, and no significant differences in codon usage were detected between these 2 *Seriatopora* species (*p* > 0.05, Chi-square test). Leucine was the most frequent amino acid, followed by phenylalanine; arginine was the least frequent. UUU (phenylalanine) was the most frequently used codon (57.5%), and UUA (leucine) was the second most common codon (38.0%). Although leucine and the UUU codon are also the most frequently used in acroporids and *Montastraea* corals, codon usages significantly differed among *Seriatopora*, acroporids, and the *M. annularis* complex (*p* < 0.01, Chi-square test for each comparison).

Translation initiation and termination codons of the *Seriatopora* mt protein-coding genes were similar to those of the *M. annularis* complex (*p* = 0.07, paired Sign test), and no differences existed between *S. hystrix* and *S. caliendrum* (Table 3; except for the termination of *nad4*). For *Seriatopora* corals, 10 of the 12 protein-coding genes used methionine (AUG) as the start codon, while *nad2* and *cob* use isoleucine (AUU) and valine (GUG), respectively. Most of the protein-coding genes were terminated by a UAA codon, except for *cox2* and *nad4* genes (UAU). Start/stop codon usage is in contrast with the majority of start/stop codons of acroporids (*p* < 0.01, paired Sign test), where start codons (AUG and GUG) and stop codons (UAA and UAG) were in equal use (Tseng et al., 2005; van Oppen et al., 2002).

### 3.3. Idiosyncratic *atp8* gene

ORFs corresponding to 12 of 13 typical protein-coding genes were detected in the mt genomes for *Seriatopora*, except for *atp8*. The search strategies of BLASTN and BLASTX detected no *atp8* coding sequences or *atp8*-like motifs. However, an unidentified ORF found in both mt genomes may correspond to the *atp8* gene of other scleractinian corals. This ORF, located between the 2 *trnW* genes, had low amino acid similarity to other scleractinian coral *atp8* genes (25.6–34.6%), but was significantly larger than its counterparts in other scleractinian corals (79 sense codons in *Seriatopora*, 66 in the *M. annularis* complex, 73 in acroporids, and 66–78 in other scleractinian corals). The transcription of the idiosyncratic *atp8* was confirmed by the fact that a RT-PCR product could be amplified in the predicted partial *atp8* using the total RNA prepared from *S. hystrix* as the template (Fig. 3). The identity of the RT-PCR product was also confirmed by DNA sequencing (data not shown).

In 9 of 14 released scleractinian mt genomes, the start of *atp8* was inferred by the presence of a methionine (Fig. 4). In *Seriatopora*, *atp8* begins with a valine, which is the same as in *Agaricia humilis* and *Pavona clavus* (Fig. 4). High similarities were observed at the beginning of the aligned sequences, and this resulted in a well-conserved N-terminal motif (Fig. 4). In addition to sequence similarities, the hydropathy profile of the ORF-encoded protein was similar to those of *atp8* in other scleractinian corals (Fig. 5). However, the hydropathy varied at the C-terminals, which became negatively charged at the 58th amino-acid position.
Table 3
Mitogenomic organizations of *Seriatopora caliendrum* and *S. hystrix*

<table>
<thead>
<tr>
<th>Region</th>
<th>Position</th>
<th><em>S. caliendrum</em></th>
<th><em>S. hystrix</em></th>
<th>Length (bp)</th>
<th>AT %</th>
<th>Start/stop codon</th>
<th>Intergenic nucleotides&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>trnM</td>
<td>1–71</td>
<td>1–71</td>
<td>71</td>
<td>59.2</td>
<td></td>
<td>ATG</td>
<td>0</td>
</tr>
<tr>
<td>rnl</td>
<td>72–1973</td>
<td>72–1975</td>
<td>1902 (1904)</td>
<td>72.5 (72.8)</td>
<td>0</td>
<td>ATG/TAA</td>
<td>19, IGS1</td>
</tr>
<tr>
<td>nad5&lt;sup&gt;5&lt;/sup&gt;</td>
<td>1974–2684</td>
<td>1976–2686</td>
<td>711</td>
<td>71.6 (71.7)</td>
<td>ATG/TAA</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Group I intron&lt;sup&gt;5&lt;/sup&gt;</td>
<td>2685–2808</td>
<td>2687–2810</td>
<td>124</td>
<td>85.5</td>
<td></td>
<td>ATG/TAA</td>
<td>1, IGS3</td>
</tr>
<tr>
<td>nad1</td>
<td>2809–3786</td>
<td>2811–3788</td>
<td>1092</td>
<td>74.7</td>
<td></td>
<td>ATT/TAA</td>
<td>1</td>
</tr>
<tr>
<td>nad6</td>
<td>6259–6822</td>
<td>6261–6824</td>
<td>564</td>
<td>72.7 (72.9)</td>
<td>ATG/TAA</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>atp6</td>
<td>6822–7499</td>
<td>6824–7501</td>
<td>678</td>
<td>71.4 (71.2)</td>
<td></td>
<td>ATG/TAA</td>
<td>0</td>
</tr>
<tr>
<td>Putative control region</td>
<td>7500–8494</td>
<td>7502–8524</td>
<td>995 (1023)</td>
<td>62.8 (62.0)</td>
<td></td>
<td>ATG/TAG (TAA)</td>
<td>12, IGS4</td>
</tr>
<tr>
<td>nad4</td>
<td>8495–9940</td>
<td>8525–9970</td>
<td>1446</td>
<td>68.7 (69.0)</td>
<td></td>
<td>ATG/TAG (TAA)</td>
<td>12, IGS4</td>
</tr>
<tr>
<td>rns</td>
<td>9953–10,868</td>
<td>9983–10,898</td>
<td>916</td>
<td>65.8 (66.6)</td>
<td></td>
<td>ATG/TAG (TAA)</td>
<td>12, IGS4</td>
</tr>
<tr>
<td>cox3</td>
<td>10,869–11,648</td>
<td>10,899–11,678</td>
<td>780</td>
<td>66.4</td>
<td></td>
<td>ATG/TAA</td>
<td>36, IGS5</td>
</tr>
<tr>
<td>cox2</td>
<td>11,685–12,443</td>
<td>11,715–12,473</td>
<td>759</td>
<td>69.7 (69.6)</td>
<td></td>
<td>ATG/TAA</td>
<td>–19</td>
</tr>
<tr>
<td>nad4L</td>
<td>12,425–12,724</td>
<td>12,455–12,754</td>
<td>300</td>
<td>77.3 (76.7)</td>
<td></td>
<td>ATG/TAA</td>
<td>–1</td>
</tr>
<tr>
<td>nad3</td>
<td>12,724–13,068</td>
<td>12,754–13,098</td>
<td>345</td>
<td>70.8 (71.0)</td>
<td></td>
<td>ATG/TAA</td>
<td>0</td>
</tr>
<tr>
<td>Group I intron&lt;sup&gt;3&lt;/sup&gt;</td>
<td>13,069–13,122</td>
<td>13,099–13,152</td>
<td>54</td>
<td>81.5</td>
<td></td>
<td>ATG/TAA</td>
<td>0</td>
</tr>
<tr>
<td>nad5&lt;sup&gt;3&lt;/sup&gt;</td>
<td>13,123–14,250</td>
<td>13,153–14,280</td>
<td>1128</td>
<td>74.0 (73.5)</td>
<td>TAA</td>
<td>12, IGS6</td>
<td></td>
</tr>
<tr>
<td>trnW</td>
<td>14,263–14,332</td>
<td>14,293–14,362</td>
<td>70</td>
<td>62.9 (64.3)</td>
<td></td>
<td>ATG/TAG (TAA)</td>
<td>1</td>
</tr>
<tr>
<td>Putative atp8</td>
<td>14,334–14,570</td>
<td>14,364–14,600</td>
<td>237</td>
<td>75.5</td>
<td></td>
<td>GTG/TAA</td>
<td>128 (91), IGS8</td>
</tr>
<tr>
<td>trnW</td>
<td>14,699–14,768</td>
<td>14,692–14,761</td>
<td>70</td>
<td>64.3 (62.9)</td>
<td></td>
<td>ATG/TAA</td>
<td>657 (713), IGS9</td>
</tr>
<tr>
<td>cox1</td>
<td>15,426–16,973</td>
<td>15,475–17,022</td>
<td>1548</td>
<td>64.7 (64.9)</td>
<td></td>
<td>ATG/TAA</td>
<td>38, IGS10</td>
</tr>
</tbody>
</table>

Different values and codons of *S. hystrix* are given in parentheses.

<sup>a</sup> Intergenic nucleotides refer to intergenic bases between the gene on the same line and the gene on the line underneath, with a negative number indicating an overlap of that length. The names of the intergenic spacers (IGS1-10) are designated after the intergenic nucleotides.

of the presumptive atp8, then gradually positively charged in both *Seriatopora* species (Fig. 5). This negatively charged region at the C-terminal is shorter in other scleractinian corals; it was about 4 codons in length, starting 8 codons upstream from the end of the atp8. Considering the results of the analysis of amino acid similarity and hydropathy together suggested that the presumptive atp8 had a greatly extended 3' end.

### 3.4. Duplicated trnW (tRNA<sup>TRP</sup>)

To date, only two tRNAs have been found in scleractinian mt genomes (*trnM* and *trnW*). However, an additional tRNA, namely *trnW<sup>′</sup>* was identified in the region close to the *coxI* gene in both *Seriatopora* genomes (Fig. 2). All three tRNA genes could be folded into typical secondary structures (Fig. 6). These structures were 70 or 71 nucleotides 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 anticyodon loop of 7 nt, an ΨC stem of 4 bp, and a ΨC loop of 7 nt. The sequence similarity was high between *trnW* and *trnW<sup>′</sup>* of *Seriatopora* corals with only 1 nucleotide substitution observed at position 1 for *S. hystrix* and position 9 for *S. caliendrum*, suggesting that *trnW<sup>′</sup>* is a duplicated gene of its counterpart (Fig. 6).

### 3.5. Intergenic spacers and the putative control region

As in other scleractinian mt genomes, intergenic spacers also existed between most mt genes of *Seriatopora*. Only 5 protein-coding genes overlapped with each other: *nad6* with *atp6* by 1 bp, *cox2* with *nad4L* by 19 bp, and *nad4L* with *nad3* by 1 bp (Table 3). Intergenic spacers, excluding the putative control region (see below), totalled 1124 and 1143 bp in *S. caliendrum* and *S. hystrix*, respectively. These regions varied in length from 1 to 713 bp. Approximately 48% and 52% of the intergenic spacers were attributed to the IGS of *trnW<sup>−</sup>−cox1* (IGS9), and another 704 and 667 bp were located in the other nine IGSs. Total lengths of *Seriatopora* coral IGSs were smaller than those of acroporids (2508 bp) and the *M. annularis* complex (about 1201 bp).

The *atp6*-*nad4* intergenic spacer, containing distinct repeated elements, was identified as a putative control region for *Seriatopora* (Fig. 7). The tandem repeats in both species began forming at the same position, 290-bp downstream from the 3'-end of *atp6*. In the aligned sequences, 4 and 5 copies of the 51-bp repeated fragments (consensus unit: 5'-BYA GAA AGT AKH GVV RAY TTR AGR GAG DGT GWM RHT ARS GYA WTA MGT SAG-3') were, respectively, recognized for *S. caliendrum* and *S. hystrix*. This predicted control region was also supported by several potential hairpin structures (Fig. 8). The analysis predicted 2 stable hairpins with a long stem corresponding to positions 614–692 and 764–825 in the alignment of both *Seriatopora* species. The stabilities of the structures predicted from the real data were compared with those predicted from 10 randomized sequences. All of the randomized sequences had the potential to fold into hairpin structures, but their predicted structures were shorter than those of the actual
sequences (data not shown). In addition, while the free energies of the secondary structures of these two hairpins were $-12.2$ and $-13.2$ kcal/mol, respectively, the free energies of the potential secondary structures obtained by the randomized sequences were $-1.9$ to $-6.1$ and $-0.6$ to $-5.4$ kcal/mol. These predicted secondary structures for the randomized sequences were significantly less stable than those predicted from the actual data ($p < 0.01$, Sign test), implying the biological functions of these two hairpin structures.

### 3.6. Evaluating the utilities of mtDNA for species-level phylogeny and population genetics

Overall, there were 78 variable nucleotides between the corresponding protein-coding and rRNA genes, and 13
variable amino acids between the corresponding protein-coding genes of these 2 *Seriatopora* species (Table 4). Among protein-coding genes, *atp6* was the most divergent gene, followed by *nad5(3’)*, while *cob* and *nad2* were identical in terms of their nucleotide and amino acid sequence similarity. No amino acid differences were detected between the corresponding *nad1*, *nad2*, *nad5(5’)*, *nad6*, *nad4L*, and *cob* of these two species (Table 4). Based on these comparisons, we chose mt *atp6*, *cox1*, IGS9, and the putative control region to evaluate the phylogenetic utilities of the *Seriatopora* mt genomes, specifically for species phylogeny and population genetics.

For *atp6*, 11 variable sites were observed among the analysed 334-bp sequence for 21 individuals (Table 1). Genetic distances of the partial *atp6* sequences ranged from 0 to 0.041 [mean = 0.014 ± 0.014 (SD), n = 231], and the differences among these three hierarchical categories were significant (p < 0.01, Mann–Whitney U-test for each comparison). The highest divergences were in the interspecific category with a p-distance of 0.029 ± 0.004, which was
about 7.5- and 30-times higher than those of the interpopulational and intrapopulational categories (Fig. 9). For cox1, 9 variable sites were observed among the aligned 775 nt sequences. The pairwise genetic distances (p-distances) ranged from 0 to 0.01 (mean = 0.002 ± 0.002, n = 153). The genetic divergence of cox1 of the interspecific category was 0.004 ± 0.002 and was significantly higher than those of the intraspecific categories (p < 0.01, Mann–Whitney U-test for each comparison, Fig. 9).

Among the 713-bp of IGS9 sequences from 23 individuals, 25 substitutions and 3 indels were observed. The p-distances of the IGR9 sequences ranged from 0 to 0.038 (mean = 0.020 ± 0.019, n = 253). The genetic distance of the interspecific category was 0.038 and was significantly higher than that of the intraspecific categories (p < 0.01, Mann–Whitney U-test for each comparison, Fig. 9). For the non-repeating regions of the control region, alignment consisted of 792 positions, of which 736 bp was constant and 37 sites were parsimoniously informative. The pairwise genetic distances of the non-repeating regions ranged from 0 to 0.044 (mean = 0.021 ± 0.016, n = 231), and the differences among the 3 hierarchic categories were significant (p < 0.01, Kruskal–Wallis test). The highest divergences were in the interspecific category with a p-distance of 0.039 ± 0.003, which was 3.3- and 10-times higher than those of the interpopulational and intrapopulational categories (Fig. 9).

3.7. Phylogenetic implications

The significance of the observed hierarchical genetic divergences was also revealed in the phylogenetic analysis (Fig. 10). All of the phylogenetic trees suggested that these Seriatopora corals were divided into 2 major lineages, supported by bootstrap values exceeding 80%.

Phylogenetic analysis of atp6 and the control region provided greater resolution among populations of S. hystrix.
than did those of cox1 and IGS9 (Fig. 10). The ML tree for the 22 control region sequences and 21 atp6 sequences, with the homologous region of Stylophora pistillata as an outgroup, showed two distinct lineages, each with high support (>80% bootstrap support; Fig. 10a and b). One clade was comprised of all individuals of S. caliendrum from Taiwan. All populations of S. hystrix belonged to two separate subclades, and these subclades were joined as a sister lineage to the S. caliendrum clade. The S. hystrix-Andaman Sea subclade was monophyletic, containing only individuals from the Andaman Sea. Another subclade represented S. hystrix of the Pacific Ocean, which contained individuals from Taiwan and the South China Sea. The tree topology showed that the most recent diversification was the split between the Andaman Sea and West Pacific S. hystrix. The previous split was between the

<table>
<thead>
<tr>
<th>Loci</th>
<th>NT</th>
<th>AA</th>
<th>Vn</th>
<th>Identity</th>
<th>Vn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein-coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nad5</td>
<td>99.3</td>
<td>12</td>
<td>98.8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>(5’) nad1</td>
<td>99.8</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(3’) nad2</td>
<td>99.0</td>
<td>11</td>
<td>98.1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>cox1</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>nad6</td>
<td>99.6</td>
<td>2</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>atp6</td>
<td>98.8</td>
<td>8</td>
<td>99.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>nad4</td>
<td>99.5</td>
<td>7</td>
<td>99.1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>cox3</td>
<td>99.3</td>
<td>5</td>
<td>99.6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>cox2</td>
<td>99.6</td>
<td>3</td>
<td>99.2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>nad4L</td>
<td>99.3</td>
<td>2</td>
<td>100.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>nad3</td>
<td>99.7</td>
<td>1</td>
<td>99.1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>cox1</td>
<td>99.8</td>
<td>3</td>
<td>99.8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>rRNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rnl</td>
<td>98.4</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rns</td>
<td>98.2</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 9. Interspecific (inter-spp) and intraspecific (inter-pop and intra-pop) p-distances of different mtDNA loci of Seriatopora caliendrum and S. hystrix. Error bar represents ±SD; (○) intra-pop comparisons; (□) inter-pop comparison; (●) inter-spp comparison.

Fig. 8. Putative secondary structures in the Seriatopora candidate control region. (a) Hairpin structure at position 614–692. (b) Hairpin structure at position 764–825.

S. caliendrum clade and the ancestor of S. hystrix. Nearly identical topologies were generated by the NJ, MP, ML and BAY analyses, except for moderate bootstrap support at the node between the Andaman Sea and West Pacific S. hystrix for atp6 (Fig. 10a). Relationships between species
were resolved by the phylogenetic analysis of both \textit{cox1} and IGS9 datasets with \textit{S. caliendrum} and \textit{S. hystrix} clustering into 2 distinct clades, but polytomically among different populations of \textit{S. hystrix} (Fig. 10c and d).

4. Discussion

4.1. Molecular evolution of the Seriatopora mitochondrial genomes

The \textit{Seriatopora} mt genomes contain three unique features that differ from the published mt genomes of scleractinian corals: (1) a novel \textit{atp8}, (2) a duplicated \textit{trnW}, and (3) the location of the putative control region between \textit{atp6} and \textit{nad4}.

The peculiar \textit{atp8} is presumably a functional gene in \textit{Seriatopora} mt genomes, although it highly differs at both the nucleotide and amino acid levels from those of other scleractinian corals. The \textit{atp8} is quite heterogeneous in length among diverse organisms, e.g., 52 amino acids (aa) in \textit{Drosophila} and 68 aa in human. This protein is more conserved at the level of its secondary structure and the chemical character of the identified amino acids (Papakonstantinou et al., 1996b). Comparisons of the amino acid sequences from diverse organisms indicate that \textit{atp8} is an intrinsic membrane protein composed of 3 domains (Papakonstantinou et al., 1993, 1996a). The N-terminal domain is located in the intermembrane space and shows a conserved Met-Pro-Gln-Leu motif, a central membrane-spanning hydrophobic domain, and a positively charged C-terminal region exposed in the matrix space (Devenish et al., 1992).

The presence of a unique \textit{atp8} was also discovered in a nematode, \textit{Trichinella spiralis}, and an ascidian, \textit{Ciona intestinalis}, while an \textit{atp8} deficiency was previously considered an evolutionary constraint of these organisms (He et al.,...
Analyses of sequence and hydrophatic similarities support their 3-domain compilation and their likeness to atp8 in Limulus and humans (Gissi et al., 2004; Lavrov and Brown, 2001). The transcribed ORF between the 2 trnW genes in the Seriatopora mitogenomes also bears similar chemical compilations to atp8 of other scleractinian corals. All of them contain the conserved N-terminal domain and a central hydrophobic domain, but are more variable at the C-terminal positively charged region (Fig. 5). Despite its low global similarity to other atp8 genes, the ORF analysed here can be unambiguously regarded as a novel atp8 form with a longer C-terminal domain when compared to other scleractinian corals.

Duplication of an additional tRNA gene has been reported in other metazoan mt genomes. For example, 2 isoacceptors of trnM have been mentioned in the mt genomes of the mussels, Mytilus edulis and M. californianus (Beagley et al., 1999; Hoffmann et al., 1992), and the ascidian, Ciona intestinalis (Gissi et al., 2004). The additional trnM is thought to originate from its isoacceptor and functionally differs from it. In M. edulis, similarities at the S’ end of the 2 trnM genes suggest that they arose by gene duplication (Hoffmann et al., 1992). It has also been suggested that trnM(AUG) and trnM(AUA) are used as an initiator and elongator, respectively, although there is no experimental evidence to support this hypothesis (Beagley et al., 1999).

In Seriatopora, high similarities between the 2 trnW genes of the same species also suggest their origin by recent gene duplication of a common ancestor. By folding the trnW’, the difference of nucleotide does not change the secondary structure of the loop region, which indicates that trnW’ genes are functional in both Seriatopora species. These 2 trnW genes in Seriatopora are coded by the same 5’-UCA-3’ anticodon, and no supplemental function can be deduced. The additional trnW is supported by a postulated model of tRNA molecule replication which hypothesizes their origin simply by direct duplication of a molecule housing double-hairpin structures (Di Giulio, 1992, 2004). Nevertheless, it is not clear why tRNA was duplicated in Seriatopora mtDNA but not in other scleractinian corals. Further comparative data of other species and related genera are needed to elucidate the evolutionary process for this duplication.

Mitochondrial control regions are not well-defined and may be highly variable among scleractinian corals. The presence of tandem repeats and functional secondary structures indicate that the control region is located between atp6 and nad4 in Seriatopora and other members of pocilloporid corals (Chen et al., 2006). However, they are located between rns and cox3 in the Acroporidae (Tseng et al., 2005; van Oppen et al., 2002), and between cob and nad2 in Siderastraea (Chuang, 2006), but not yet recognized for other published scleractinian mt genomes (Fukami and Knowlton, 2005; Medina et al., 2006).

The observed differences in the mt control region of diverse Scleractinia may reflect their independent origins. If the control region had been rearranged soon after the origin of scleractinian corals, they would be close to each other in nucleotide sequences and structural composition. However, there are relatively high similarities in nucleotide sequences and in the molecular organization among congeners or genera of the same family, but the same does not hold true among families. In addition, no rearrangement or recombination of genes has been reported among the 14 published scleractinian mt genomes. Hence, the most parsimonious explanation for disparities in the control region is that it evolved independently among the different scleractinian lineages.

Taken together with the divergent atp8 and duplicated trnW, these distinctive features of Seriatopora mt genomes reveal profound implications for the evolution of the family Pocilloporidae in scleractinial phylogeny. Similar structures were also observed in the congeneric genera, Stylophora and Pocillopora (Chen et al., 2006). The Pocilloporidae has been regarded as a monophyletic group of the robust clade in scleractinian phylogeny, as inferred by mitochondrial ribosomal DNA sequences (Chen et al., 2002; Romano and Palumbi, 1996, 1997), but could be regarded as a distinct lineage and apart from the robust clade based on our study. This hypothesis is also supported by a new phylogenetic analysis based on the nucleotide sequences of mitochondrial cob and cox1 and nuclear β-tubulin and ribosomal genes (Fukami et al., 2006).

4.2. Applications of the hypervariable regions to the phylogeny of coral species and population genetics

Our study provides evidence that several protein-coding genes and intergenic spacers in scleractinian corals, at least in pocilloporids, are potentially useful for species phylogenetic construction or population genetic analyses. This contradicts the general impression that anhzoan mt genomes are slowly evolving and have little potential for phylogenetic utility below the species level (revealed in Shearer et al., 2002). The genetic distance (p-distance) between these 2 populations (Andaman Sea and Pacific Ocean) of S. hystrix ranged 0.3–0.9% and 0.9–1.7% for atp6 and the putative control region, respectively. These variations provide enough information to successfully recover the reciprocal monophyly between the Andaman Sea and Pacific Ocean S. hystrix populations (Fig. 10a and b), which is concordant with the biogeographic patterns observed in several marine organisms distributed in both the Indian and Pacific Oceans (reviewed in Benzie, 1998). Benzie (1999) suggested that although the Indo-Pacific has a complicated geographic and climatic history, recent repeated periods of lowered sea levels during the Pleistocene are the suggested major vicariance events initiating the genetic differences of separated populations for most of the last 3 Ma. If we calibrate the divergent time in isolating these 2 populations based on this scenario, the maximum diver-
gence rate would be 0.1–0.3% and 0.3–0.57% Ma⁻¹ for atp6 and the putative control region, respectively. These divergences are about 2- to 3-times faster than that of the cob divergence rate of Acropora congener (0.1–0.18% Ma⁻¹). However, a lower resolution was observed for cox1 and IG9, which were successfully used to recover the species phylogeny between S. caliandrum and S. hystrix (Fig. 10c and d). Our findings indicate that the evolutionary rate of the mt genomes could be highly variable among different genes and intergenic spacers, and in different lineages of scleractinian corals. Thus, the generalization of slow evolution and limitations of phylogenetic utility in coral mt genomes should be treated with caution before the molecular evolution of each coral lineage is evaluated.

Acknowledgments

We thank the Minister of Domestic Affairs and Minister of National Defence for providing field trips to the Taiping Island (Itu Aba Island) in the Spratly Islands; Drs. Hiro-nobu Fukami and Nancy Knowlton for the initiation and encouragement of this study; members of the Evolutionary and Ecology Genetics of Coral Reef Discussion Group, Research Center for Biodiversity, Academia Sinica (RCBAS) and three anonymous referees for constructive comments. This work was supported by the Academia Sinica Thematic Grants (2002–2004), a Grant from the Genomic Research Center, Academia Sinica (2006–2007), and the National Science Council (NSC) Grant (NSC-94-2621-B-001-005), Taiwan to C.A.C., and a NSC Grant of the National Science Council (NSC) Grant (NSC-94-EPA-Z002-004), Taiwan to C.-F. Dai. This is the Evolutionary and Ecology Genetics of Coral Reef Group, RCBAS contribution no. 42.

References


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


