



# The “Naked Coral” Hypothesis Revisited – Evidence for and Against Scleractinian Monophyly

Marcelo V. Kitahara<sup>1,2,9</sup>, Mei-Fang Lin<sup>3,4,5,9</sup>, Sylvain Forêt<sup>5,6</sup>, Gavin Huttley<sup>7</sup>, David J. Miller<sup>3,5,\*</sup>, Chaolun Allen Chen<sup>4,8,9\*</sup>

**1** Departamento de Ciências do Mar, Universidade Federal de São Paulo, Santos, São Paulo, Brazil, **2** Centro de Biologia Marinha (CEBIMar), Universidade de São Paulo, São Sebastião, São Paulo, Brazil, **3** School of Pharmacy and Molecular Sciences, James Cook University, Townsville, Queensland, Australia, **4** Biodiversity Research Centre, Academia Sinica, Taipei, Taiwan, **5** ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland, Australia, **6** Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia, **7** John Curtin School of Medical Research, Australian National University, Canberra, Australian Capital Territory, Australia, **8** Institute of Oceanography, National Taiwan University, Taipei, Taiwan, **9** Taiwan International Graduate Program (TIGP)-Biodiversity, Academia Sinica, Taipei, Taiwan

## Abstract

The relationship between Scleractinia and Corallimorpharia, Orders within Anthozoa distinguished by the presence of an aragonite skeleton in the former, is controversial. Although classically considered distinct groups, some phylogenetic analyses have placed the Corallimorpharia within a larger Scleractinia/Corallimorpharia clade, leading to the suggestion that the Corallimorpharia are “naked corals” that arose via skeleton loss during the Cretaceous from a Scleractinian ancestor. Scleractinian paraphyly is, however, contradicted by a number of recent phylogenetic studies based on mt nucleotide (nt) sequence data. Whereas the “naked coral” hypothesis was based on analysis of the sequences of proteins encoded by a relatively small number of mt genomes, here a much-expanded dataset was used to reinvestigate hexacorallian phylogeny. The initial observation was that, whereas analyses based on nt data support scleractinian monophyly, those based on amino acid (aa) data support the “naked coral” hypothesis, irrespective of the method and with very strong support. To better understand the bases of these contrasting results, the effects of systematic errors were examined. Compared to other hexacorallians, the mt genomes of “Robust” corals have a higher (A+T) content, codon usage is far more constrained, and the proteins that they encode have a markedly higher phenylalanine content, leading us to suggest that mt DNA repair may be impaired in this lineage. Thus the “naked coral” topology could be caused by high levels of saturation in these mitochondrial sequences, long-branch effects or model violations. The equivocal results of these extensive analyses highlight the fundamental problems of basing coral phylogeny on mitochondrial sequence data.

**Citation:** Kitahara MV, Lin M-F, Forêt S, Huttley G, Miller DJ, et al. (2014) The “Naked Coral” Hypothesis Revisited – Evidence for and Against Scleractinian Monophyly. *PLoS ONE* 9(4): e94774. doi:10.1371/journal.pone.0094774

**Editor:** John Murray Roberts, Heriot-Watt University, United Kingdom

**Received:** October 31, 2013; **Accepted:** March 20, 2014; **Published:** April 16, 2014

**Copyright:** © 2014 Kitahara et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by grants from the National Science Council (NSC) and Academia Sinica (Thematic Grants 2005-2010) to CAC, from the Australian Research Council to DJM and SF, and from São Paulo Funding Agency (FAPESP grant 2012/21583-1) to MVK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: cac@gate.sinica.edu.tw (CAC); david.miller@jcu.edu.au (DJM)

† These authors contributed equally to this work.

## Introduction

The order Scleractinia, comprising the anthozoan cnidarians that produce a continuous external aragonitic skeleton [1], are not only the architects of some of the most complex habitats (i.e. coral reefs) but are also near ubiquitous in distribution. Despite their global significance [2–7], several key aspects of scleractinian evolution are as yet poorly understood. Most coral families are first identifiable in the Triassic, by which time much of the extant morphological diversity is represented. Molecular data implies a deep split of extant corals between two large clades, the “Complex” and “Robust” [8–17], but many families defined by morphology are not monophyletic by molecular criteria and some are split between “Complex” and “Robust” clades [9–11,14,17–20]. One hypothesis to explain the sudden appearance of a highly diverse Middle Triassic coral fauna is that skeletonisation has been an ephemeral trait during the evolution of the Scleractinia [21]. Under this scenario, scleractinian lineages may have undergone

skeleton loss in the face of global environmental instability [21], which would severely compromise fossil preservation. Consistent with this idea, some corals have been shown to undergo complete (but reversible) skeleton loss under acid conditions [22], whereas other species are apparently much less susceptible to skeleton dissolution [23].

The “naked coral” hypothesis [24] is a topical extension of the idea of skeleton ephemerality in corals. Corallimorpharians, anthozoans that lack skeletons, have a close but unclear relationship to the Scleractinia. Corallimorpharians and scleractinians are very similar both in terms of anatomy and histology (see [92]), and these characteristics have in the past been used to argue for merging the orders [21,26]. Medina et al. [24] conducted a phylogenetic analysis based on the proteins encoded by 17 complete mitochondrial (mt) genomes, which suggested that scleractinians are paraphyletic, corallimorpharians being more closely related to “Complex” than are “Robust” corals (Figure 1).

The authors hypothesize that the Corallimorpharia ("naked corals") may have arisen during the Cretaceous (110~132 Mya) from a scleractinian ancestor that had undergone skeleton loss as a consequence of ocean acidification. A recent study [27] using complete mitochondrial genomes from a broad range of representative cnidarians also failed to unambiguously reject the "naked coral" hypothesis.

Here we applied a wide variety of analytical methods to a more comprehensive dataset of complete mitochondrial genome sequences (50 representative anthozoans) to better understand hexacorallian phylogeny. Whilst phylogenetic analyses based on amino acid (aa) data were for the most part consistent with the "naked corals" scenario (i.e. paraphyletic Scleractinia), it remains possible that the "naked corals" topology is an artefact caused by saturation, compositional biases or other violations of model assumptions. With the possible exception of cerianthiids [28], anthozoan cnidarians are thought to carry out mt DNA repair and thus differ from bilaterians *sensu stricto* in this respect. The main lines of evidence for repair are the extremely slow rate of evolution of the mt genomes of Anthozoa [29,30] and the presence of a *MutS*-like gene in the octocoral mt genome [25,31]. We speculate that reduced efficiency of mt DNA repair in the "Robust" coral lineage could account for the observed anomalies in codon use and composition.

## Material and Methods

### DNA extraction and Polymerase Chain Reaction

Total genomic DNA was extracted using a Qiagen Qiamap or DNeasy Tissue Kit (QIAGEN). DNA concentrations were determined using a Nanodrop 1000 (Thermo Scientific) and an aliquot of each species total genomic DNA was diluted or concentrated to a final concentration of 40 ng/μl. Previously described primer sequences [32] were used to amplify the complete mt genomes of the following scleractinians: *Alveopora* sp.; *Astreopora explanata*; *A. myriophthalma*; *Isopora palifera*; and *I. togianensis*.

Two to three fragments (3~9 kb) covering the entire mt genome of all but *G. hawaiiensis* were amplified by long Polymerase Chain Reaction (long-PCR) technique [33] from partial fragments of *ms*, *ml* and *cox1* genes obtained from coral-specific primers and invertebrate universal primer [34,35,36]. Long-PCR were carried out using Takara La Taq using a slightly modified conditions from those recommended by the manufacturer as follows: 94°C for

1 min, then 30 cycles of 10 s at 98°C, 45 s at 62~63°C, 14.25 min at 68°C for, and 10 min at 72°C. For *G. hawaiiensis* portions of *ml*, *ms* and *cox1* were first amplified with the universal primers mentioned above, followed by the scleractinian universal primers CS-1 to CS-21 [37] that covered the entire mt genome. To obtain sequence from regions that did not yield product using these primers, nineteen specific primers were developed based on the sequences retrieved from *G. hawaiiensis* (Table S1). Polymerase Chain Reaction (PCR) were carried out using the Advantage2 polymerase kit (Clontech) under the conditions recommended by the manufacturer. PCR conditions were: 95°C for 5 min, then 30 cycles of 30 s at 94°C, 60 to 90 s at 54~60°C, 90 s at 72°C, followed by 5 min at 72°C.

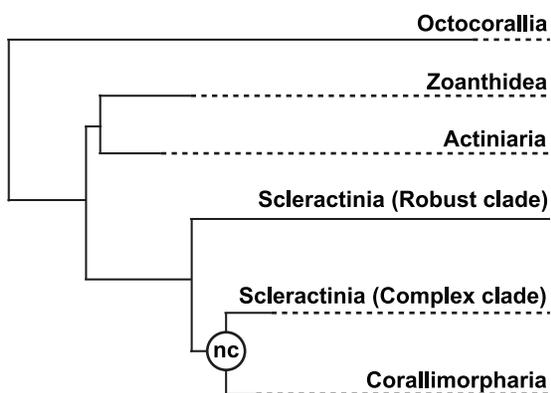
### Phylogenetic analysis

Resulting sequences were verified and assembled using Sequencher 4.8 (Gene Codes Corporation) and then analyzed in Vector NTI 9.0 (InforMax, Invitrogen life science software). Examination of open reading frames (ORFs) and codon usage, as well as other DNA statistics were performed using Dual Organelle Genome Annotator [38], Sequence Manipulation Suite v.2 [39], and MEGA5 [40]. In addition to the 6 new complete scleractinian mt genome sequences determined for this study, those of 25 other scleractinians, 12 corallimorpharians, 2 sea anemones, single antipatharian and zoanthid species, and 3 octocorals were obtained from public databases (Table 1).

In order to make the analyses based on nucleotide and amino acid sequence data strictly comparable, the rRNAs, IGS, and tRNAs coding sequences were excluded from consideration. Therefore, for each species included in the present analysis, the data set included all protein-coding genes. The predicted amino acid sequences encoded by each of the 50 mt genomes were aligned using MAFFT v.5 [41]. These alignments were reverse translated to generate nucleotide sequence alignments, and phylogenetic inferences carried out on the concatenated amino acid and nucleotide alignments, removing all positions containing more than 50% gaps. The most appropriate model of nucleotide substitution was determined for the final alignment (totaling 11,802 bp) by the hierarchical likelihood ratio test implemented in MEGA5 as GTR+I+G (lnL -133020.1). Maximum Likelihood fits of 48 different amino acid substitution models using only positions that did not contain any gaps or missing data were calculated in MEGA5 [40]. There were a total of 3,666 positions (from the 3,934 aa) in this final dataset and JTT+G+I+F (lnL -51687.5) was chosen as the best evolutionary model.

Phylogenetic analyses were performed using PhyML [42] for Maximum Likelihood (ML) and MrBayes version 3.1.2 [43] for Bayesian Inference (BI). ML analyses were performed under the GTR model for nt alignments and JTT for aa alignments. For the BI, 2 runs of 4 chains were calculated for 10 million generations with topologies saved every 1,000 generations. One million generations were discarded as burn-in to ensure that the likelihood had plateaued and that the two runs had converged to less than 0.002 average standard deviation of split frequencies.

Given concerns for the influence of the long branch of the "Robust" scleractinian clade, ML phylogenetic analyses were repeated under several different scenarios for the nt data matrices as follow: i) different substitution model categories following Bayesian Information Criterion [BIC] and ML [lnL] recommendations; ii) systematically codon exclusion (1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup>); iii) different nucleotide divergence rates across frames; iv) coding nucleotide data as purines and pyrimidies (RY-coding) (see [44,45]) RY-coding excluding the 3<sup>rd</sup> codon. Likewise, additional ML phylogenetic analyses of the aa final alignment included: i)



**Figure 1. Phylogenetic relationships within the cnidarian Class Anthozoa according to the "naked corals" (nc) hypothesis (modified from [24]).**

doi:10.1371/journal.pone.0094774.g001

**Table 1.** Mitochondrial genome sequence data included in the present analyses

Order	Species	size (bp)	GenBank accession #	Reference
<b>Actiniaria</b>				
	<i>Metridium senile</i>	17,443	NC000933	[78]
	<i>Nematostella</i> sp.	16,389	DQ643835	[24]
<b>Alcyonacea</b>				
	<i>Briareum asbestinum</i>	18,632	DQ640649	[24]
	<i>Pseudopterogorgia bipinnata</i>	18,733	DQ640646	[24]
	<i>Sarcophyton glaucum</i>	Incomplete	AF064823	[86]
<b>Antipatharia</b>				
	<i>Chrysopathes formosa</i>	18,398	NC008411	[87]
<b>Corallimorpharia</b>				
	<i>Actinodiscus nummiformis</i>	20,922		Lin et al. (submitted)
	<i>Amplexidiscus fenestrafer</i>	20,188		Lin et al. (submitted)
	<i>Corallimorphus profundus</i>	20,488		Lin et al. (submitted)
	<i>Corynactis californica</i>	20,632		Lin et al. (submitted)
	<i>Discosoma</i> sp.1	20,908	DQ643965	[24]
	<i>Discosoma</i> sp.2	20,912	DQ643966	[24]
	<i>Pseudocorynactis</i> sp.	21,239		Lin et al. (submitted)
	<i>Rhodactis indosinesis</i>	20,092		Lin et al. (submitted)
	<i>Rhodactis mussoides</i>	20,826		Lin et al. (submitted)
	<i>Rhodactis</i> sp.	20,093	DQ640647	[24]
	<i>Ricordea florida</i>	21,376	DQ640648	[24]
	<i>Ricordea yuma</i>	22,015		Lin et al. (submitted)
<b>Scleractinia</b>				
	<i>Acropora tenuis</i>	18,338	AF338425	[64]
	<i>Agaricia humilis</i>	18,735	DQ643831	[24]
	<i>Anacropora matthai</i>	17,888	AY903295	[32]
	<i>Alveopora</i> sp.	18,146		KJ634271
	<i>Astrangia</i> sp.	14,853*	DQ643832	[24]
	<i>Astreopora explanata</i>	18,106		KJ634269
	<i>Astreopora myriophthalma</i>	18,106		KJ634272
	<i>Colpophyllia natans</i>	16,906*	DQ643833	[24]
	<i>Euphyllia ancora</i>	18,875	JF825139	[37]
	<i>Fungiacyathus stephanus</i>	19,381	JF825138	[37]
	<i>Gardineria hawaiiensis</i>	19,429		Lin et al. (submitted)
	<i>Goniopora columna</i>	18,766	JF825141	[37]
	<i>Isopora palifera</i>	18,725		KJ634270
	<i>Isopora togianensis</i>	18,637		KJ634268
	<i>Madracis mirabilis</i>	16,951*	EU400212	[61]
	<i>Madrepora oculata</i>	15,839*		[65]
	<i>Montastraea annularis</i>	16,138*	AP008974	[62]
	<i>Montastraea faveolata</i>	16,138*	AP008978	[62]
	<i>Montastraea franksi</i>	16,137*	AP008976	[62]
	<i>Montipora cactus</i>	17,887*	AY903296	[32]
	<i>Mussa angulosa</i>	17,245*	DQ643834	[24]
	<i>Pavona clavus</i>	18,315	DQ643836	[24]
	<i>Pocillopora damicornis</i>	17,425*	EU400213	[61]
	<i>Pocillopora eydouxi</i>	17,422*	EF526303	[88]
	<i>Polycyathus</i> sp.	15,357*	JF825140	[37]
	<i>Porites okinawensis</i>	18,647	JF825142	[37]

Table 1. Cont.

Order	Species	size (bp)	GenBank accession #	Reference
	<i>Porites porites</i>	18,648	DQ643837	[24]
	<i>Seriatopora caliendrum</i>	17,010*	NC010245	[59]
	<i>Seriatopora hystrix</i>	17,059*	EF633600	[59]
	<i>Siderastrea radians</i>	19,387	DQ643838	[24]
	<i>Stylophora pistillata</i>	17,177*	EU400214	[61]
Zoanthidea				
	<i>Savalia savaglia</i>	20,764	NC008827	[89]

doi:10.1371/journal.pone.0094774.t001

coding aa using the common six groups that usually replace one another [46,47], where MVIL were recoded as 1, FYW as 2, ASTGP as 3, DNEQ as 4, and RKH as 5, and C as 6; ii) to allow general-time-reversible (GTR) matrix to be used, the aa dataset was recoded to four categories instead of six. In this case, following [47] the aromatic (FYW) and hydrophobic (MVIL) amino acids were combined and the rare cysteine was considered as missing data. The four amino acid categories were named A, T, G, and C, respectively; iii) exclusion of all Phe, Ala, Thr and Tyr from the alignment, once the percentage of occurrence of these aa, especially of Phe and Ala, in the "Robust" scleractinian clade are significantly different once compared to all other hexacorallians included in the present analysis; iv) phylogenetic reconstructions using different evolutionary models as retrieved from results of the Bayesian Information Criterion [BIC] and ML [lnL] recommendations (JTT+G+I+F, cpREV+G+I+F, WAG+G+I+F, and Dayhoff+G+I+F); v) exclusion of Octocorallia sequences as outgroups; and vi) systematically exclusion of fast evolving sites. To find such sites, the mean (relative) evolutionary rate was estimated under the JTT+G+F in MEGA5, and a discrete Gamma (+G) distribution with 5 categories was used to model evolutionary rate differences among sites. Subsequently, 3 minimum evolution ML phylogenies were reconstructed systematically excluding all fast evolving sites that had means of >2.99, >1.99, and >1.49 respectively.

Trees with non-stationary, non-homogeneous models were computed using nhPhyML [48] with 5 categories of (G+C) content. Quartet puzzling with the Barry and Hartigan model was implemented using the PyCogent library [49], as was the Goldman [50] test.

Codon-based ML trees were inferred using CodonPhyml [51]. The results presented here used the Yap et al. model [52], but similar results were observed with other models. Codon-based trees were also inferred using MrBayes with a GTR substitution model and three categories of non-synonymous/synonymous ratios (M3 model).

For amino acids, phylogenies based on the CAT-GTR, CAT-Poisson and GTR models were inferred using PhyloBayes [53]. For each inference, the program was run until the effective size was greater than 300 and until the largest discrepancy across bipartition between runs was less than 0.1. Majority rule posterior consensus trees were built after deleting 1000 burn-in samples and taking every 10 generations.

Comparisons of topologies were carried out using the Approximately Unbiased, Kishino-Hasegawa and Shimodaira-Hasegawa tests implemented in the program Consel [54]. These tests compare the significance of the difference in likelihood of two competing topologies under the same model.

### Base Frequencies Distance Trees

The homogeneity of base frequencies among taxa is a major assumption of many molecular phylogenetic methods [55]. Therefore, changes in base composition between lineages can lead to errors in phylogenetic inference, particularly in the case of mt genome data (see [56]). In an attempt to assess the potential for compositional bias affecting the anthozoan phylogenetic inference, minimum evolution Base Frequencies distance trees (BF) were estimated using MEGA5 from matrices of pairwise BF distances.

Following Phillips et al. [45], BF distance was calculated for each taxon pair for each nucleotide category (i.e.  $BF\ distance = ((\{A_i - A_j\} + \{T_i - T_j\} + \{C_i - C_j\} + \{G_i - G_j\})/2)$ , where  $i$  and  $j$  are the frequencies of each corresponding nucleotide for the  $i$ th and  $j$ th taxa, respectively.

### tRNA and rRNA trees

For each of the 50 anthozoan species studied, the four mitochondrial genes encoding stable RNAs (i.e., 12S rRNA, 16S rRNA, *tmM*, and *tmW*) were retrieved, but the octocoral and actinarian data were excluded from phylogenetic analyses because of the difficulty of generating unambiguous alignments when they were included. For these analyses, the scleractinian and corallimorpharian sequences were aligned and the antipatharian *Chrysopathes formosa* used as the outgroup. Each stable RNA sequence was aligned using essentially the same approach as for the protein-coding genes, the final alignments being 1,039 bp for 12S rDNA, 1,866 bp for 16S rDNA, 72 bp for *tmM*, and 70 bp for the *tmW*. Phylogenetic inferences were based on concatenated alignments and the most appropriate model of nucleotide substitution as determined by the hierarchical likelihood ratio test was GTR+G. ML (SH-like and 100 bootstrap) were performed using PhyML [42], and BI using MrBayes version 3.1.2 [43]. BI and ML analyses were performed using the GTR model as described above.

### Results and Discussion

Whereas previous analyses were based on a limited range of scleractinian and corallimorpharian mitochondrial genomes [24,27], taxon sampling was increased in the present study to a total of 50 mt genomes, which included 12 corallimorpharians and 31 scleractinians (Table 1). Two taxa included in the present study are of particular evolutionary significance: *Corallimorphus profundus* and *Gardinieria hawaiiensis*. Both anatomical [57] and molecular (this study) data suggest that *C. profundus* represents a deep-diverging corallimorpharian clade. *G. hawaiiensis* represents a lineage of scleractinians that is thought to have diverged prior to the "Complex"/"Robust" split [11,17,58].

**Table 2.** Compositional biases of the mitochondrial genomes of the anthozoan Orders included in the present analysis.

Sub-class	Order	Group	Nucleotide		Protein			
			G+C (%)	A+T (%)	FYMINK (%)	GARP (%)	FYMINK/GARP	
Octocorallia	Alcyonacea		35.33	64.66	30.54	21.96	1.39	
Hexacorallia	Antipatharia		38	62	29.78	22.05	1.35	
	Actiniaria		37.95	62.05	29.92	22.47	1.33	
	Zoanthidea		43.9	56.1	29.4	23.09	1.27	
	Corallimorpharia		37.95	62.05	28.73	22.15	1.29	
	Scleractinia		Basal*	38.8	61.2	28.51	22.19	1.28
			Complex*	37.59	62.41	29.42	21.81	1.34
		Robust*	31.2	68.8	33.71	19.36	1.74	

The figures shown are averages across the range of species included. For proteins, the comparisons are made between the (A+T)-rich amino acids FYMINK (Phe, Tyr, Met, Ile, Asn, and Lys), and the (G+C)-rich amino acids GARP (Gly, Ala, Arg, and Pro). Asterisks indicate groupings based on molecular data but whose taxonomic validity remains to be established.

doi:10.1371/journal.pone.0094774.t002

### General characteristics of the mt genomes of hexacorallians

All of the hexacorallian mt genomes sequenced to date contain 13 protein-coding genes (*atp6* and *8*, *cox1-3*, *cob*, *nad1-6*, and *nad4L*), 2 genes encoding ribosomal RNAs (*rns* and *ml*), and 2 encoding tRNAs (*trnW* and *trnM*), although members of the scleractinian genus *Seriatopora* have a duplication of *trnW* and thus have a total of 3 tRNA genes [59]. Whilst Hexacorallia in general display little variation in size of the mt genome, members of the Scleractinia are exceptional in having mt genomes ranging from >19.4 Kb in the "Basal" coral *Gardineria*, to <15 Kb in some "Robust" corals, those of "Complex" corals being intermediate in size (~18.9–19.4 Kb) [37]. The size of each mt gene is also relatively stable across the range of hexacorallians, exceptions being *ml* and *rns*, which vary by almost 500 and 700 bp respectively. As in Octocorallia [60], intergenic regions and introns (*cox1* and *nad5*) account for most of the observed variation in mt genome size. Some differences were apparent across hexacorallian orders, but gene organisation was remarkably uniform across the full range of Scleractinia, the only deviations from the canonical gene map [24,32,59,61–64] being two azooxanthellate corals, *Lophelia pertusa* [90] and *Madrepora* spp. [65]. Likewise the majority (10 of 12) of corallimorpharian mt genomes conform to a distinct gene order, the exceptions being the azooxanthellate species *Corallimorphus profundus* and *Corynactis californica*.

The nucleotide composition of the mt protein-coding genes of hexacorallians has a clear (A+T)-bias, ranging from around 56% in the zoanthid *Savalia* to an average of 69% in "Robust" corals (Table 2 and Figure 2). The coding sequences of "Robust" corals have a high thymine and low cytosine content compared to other scleractinians (Figure 2). Surprisingly, this T-enrichment over other scleractinians is not restricted to silent codon positions, but is also clear at the first (5%) and second (3%) codon positions (Figure 2), resulting in over 400 aa substitutions in "Robust" corals relative to other hexacorallians (see also Table 2).

### Contradictory phylogenies based on nucleotide or amino acid sequence data

Based on the final nt alignment (11,298 bp) consisting of all 13 protein-coding genes from 50 anthozoan mt genomes (Table 1), ML and BI recovered identical topologies and indicated that all anthozoan orders included in the analysis are monophyletic

(Figure 3a). The topology shown was strongly supported, with only few nodes not having 100% support in both ML and BI. Scleractinians and corallimorpharians appear as sister groups. Whilst these results based on nt data flatly contradict the "naked coral" hypothesis [24], the application of the same phylogenetic methods to the corresponding amino acid sequence data (3,934 aa residues) consistently placed corallimorpharians as the sister group to the "Complex" Scleractinia, within the scleractinian clade (Figure 3b). In addition, the protein-based phylogenies differ in the positions of Actiniaria, Zoanthidea, and Antipatharia, and also in placing *G. hawaiiensis* as a member of the "Complex" corals instead of forming a basal scleractinian lineage [11,17]. Codon-based phylogenies also strongly support the grouping of corallimorpharians with "Complex" corals (Figure 4). The significance of the difference in likelihood between the competing topologies is shown on table (Table 3). The preference for the "naked coral" topology is highly significant for the trees based on amino acids and codons, whereas the significance of the difference is weaker for nucleotide-based trees.

To better understand the basis of these contrasting results, we examined the potential for artifacts to arise as a result of the analytical methods or biases in the datasets.

### Use of different substitution models and removal of rapidly evolving sites

In the case of both nt and aa analyses, changing the outgroup had no effect on topology, and neither did the use of different substitution model categories. For nucleotides, the models validated included: (i) parametric GTR [93] with gamma distribution of rates among sites; (ii) TN93 [66] with gamma distribution and invariable sites; (iii) and HKY [67]. Using amino acid data, the JTT+G+I+F [68], cpREV+G+I+F [91], WAG+G+I+F [69], and Dayhoff+G+I+F [70]. Furthermore, in order to verify differences in evolutionary models selection, the same analyses were extended with the exclusion of all sequences from "Robust" corals from the dataset. However, these new analyses recovered similar results as described above.

Next, standard rate effects were examined. Potential saturation effects in the nt data were examined by systematically excluding the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> codon positions from analyses, but the ML topology retrieved and statistical support for nodes did not differ significantly from those shown in Figure 3a. In fact, exclusion of the 3<sup>rd</sup> codon position improved support for some nodes (Figure

S1). Removing the most rapidly evolving sites in the aa alignment [47] also had no effect on the topology. This approach involved estimation of the mean (relative) evolutionary rate (ER) for each site under the JTT+G+F [68] model and then excluding those sites with ERs of  $\geq 2.99$ ,  $\geq 1.99$ , or  $\geq 1.49$  (note that sites displaying ER  $> 1$  are evolving faster). The percentages of sites excluded in these cases were 18.3%, 23.1%, and 32.4% respectively; 1,275 of the 3,934 sites could therefore be excluded from the analyses without influencing the overall topology (Figure S2).

### Compositional bias effects: nucleotides

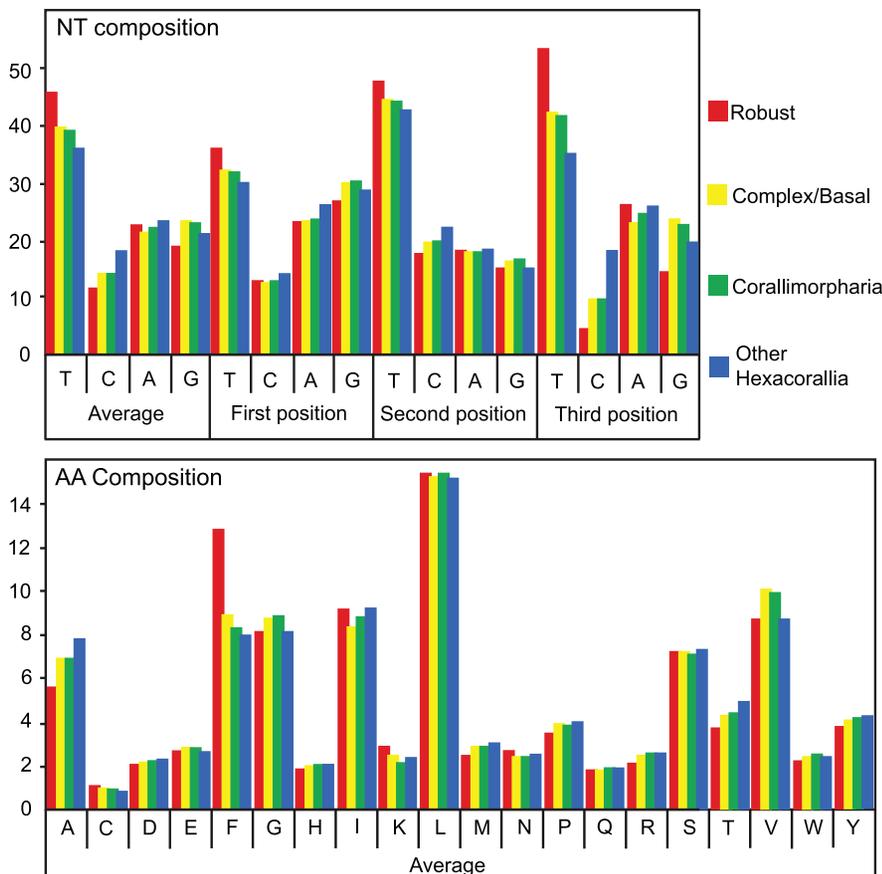
Having investigated potential artifacts arising from standard rate effects, the effects of compositional heterogeneity in the nucleotide and amino acid data were examined. In the case of the nt alignment, this involved RY coding [44,45], with or without exclusion of the third codon position from the resulting alignment (Table S2), which also resulted in scleractinian monophyly (Figure S3).

As noted above, the mt genomes of "Robust" corals do differ significantly from those of all other hexacorallians in terms of nucleotide composition, and this has consequences for both codon use and amino acid composition in the proteins that it encodes. Figure 2, Table 2 and Figure S4 show the overall base composition of mt protein-coding genes of the anthozoans included in the present analysis, and also the percentage of each base occurring in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> codon positions.

Whereas most hexacorallians have (A+T) contents around 62% (hence (A+T)/(G+C) of around 1.63), "Robust" corals have a significantly higher (A+T) content ((A+T)/(G+C) = 2.20). Consequently, the (A+T)-skew is  $>6\%$  higher in "Robust" corals than in all other hexacorallians (Figure S4). This bias is asymmetrically distributed, such that in "Robust" corals the coding strand is anomalously high in thymine and low in cytosine. Such heterogeneities in base composition are a potential source of error in phylogenetic analyses [71].

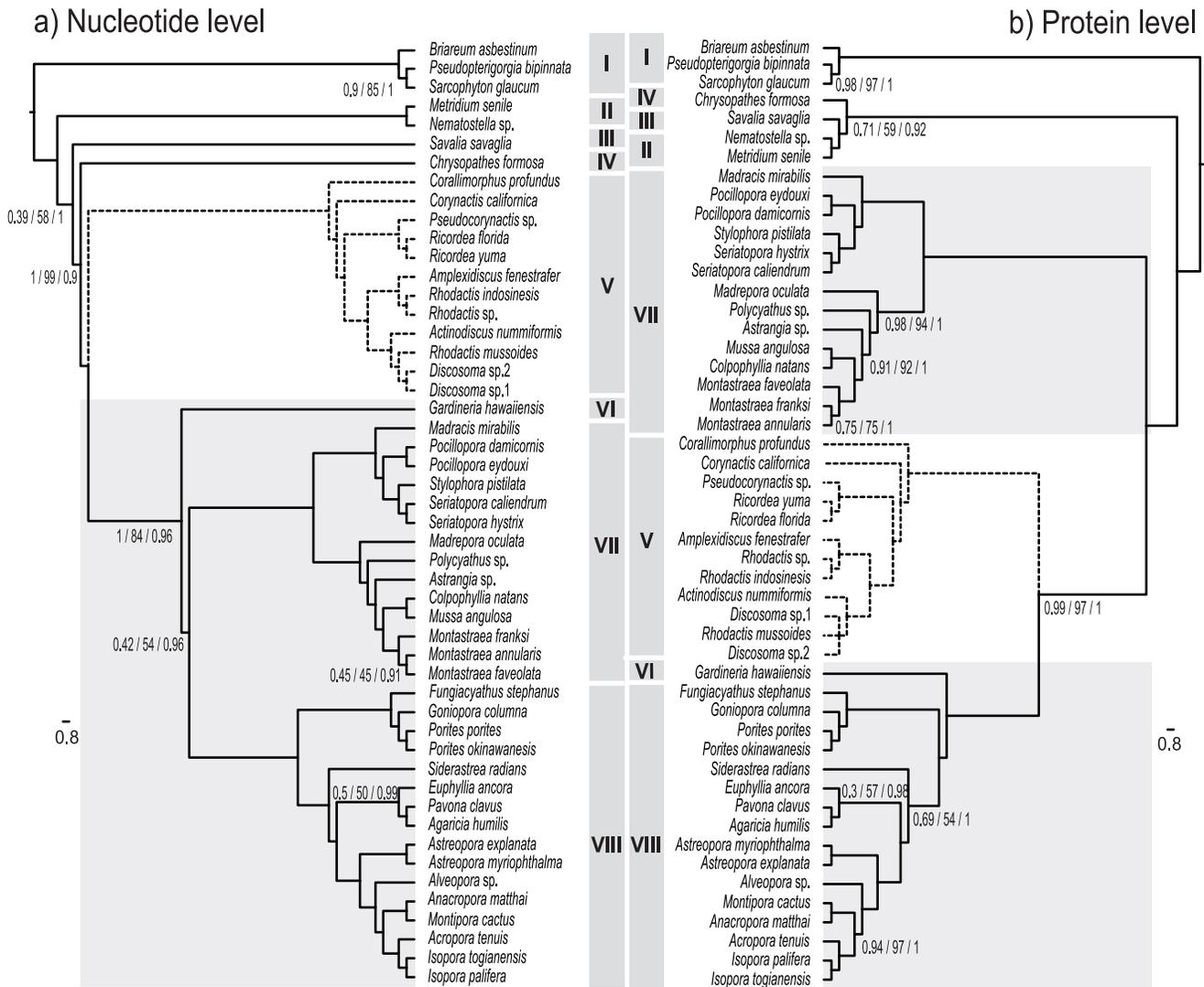
In order to take into account this variability in nucleotide composition, we used the GG98 non-stationary, non-homogeneous model [72] implemented in the nhPhyML software [48]. In this approach the "naked coral" topology has the highest likelihood, but the difference in likelihood of the two competing topologies is not statistically significant (Table 3).

We further explored the effect of compositional heterogeneity using the Barry and Hartigan model [4] implemented in the PyCogent library [49]. The Barry and Hartigan model is the most general (makes the fewest assumptions) substitution model for nucleotides. It allows for non-reversible and non-stationary processes on every branch of a phylogeny and does not assume the process is time-homogeneous within or between branches. The complexity of this model precluded tree inference; instead, 1,000 quartets, each comprising a "Robust" coral, a "Complex" coral, a corallimorpharian and an outgroup, were sampled. The majority



**Figure 2. Nucleotide (upper) and amino acid (lower) content of the protein coding genes in the mitochondrial genomes of hexacorallians.** The data shown are averages across the "Robust" corals (n = 14), basal and "Complex" corals (n = 17), corallimorpharians (n = 12) and other hexacorallians (n = 4).

doi:10.1371/journal.pone.0094774.g002



**Figure 3. Phylogenetic analyses based on (a) the nucleotide sequences of the mitochondrial genes encoding proteins and (b) the amino acid sequences of the proteins encoded by the mitochondrial genomes.** Values on the nodes indicate the non-parametric SH test and bootstrap replicates (ML), and posterior probability (BI) support respectively. Where no values are shown on a node, that edge was fully supported under all analyses. Dashed lines indicate the corallimorpharian clade. Light grey blocks identify the scleractinian clades. (I) Octocorallia used as outgroup; (II) Actiniaria; (III) Zoantheida; (IV) Antipatharia; (V) Corallimorpharia; (VI) "Basal" Scleractinia; (VII) "Robust" Scleractinia; (VIII) "Complex" Scleractinia.

doi:10.1371/journal.pone.0094774.g003

(94%) of these quartets grouped complex corals with corallimorphs and "Robust" corals with the outgroup.

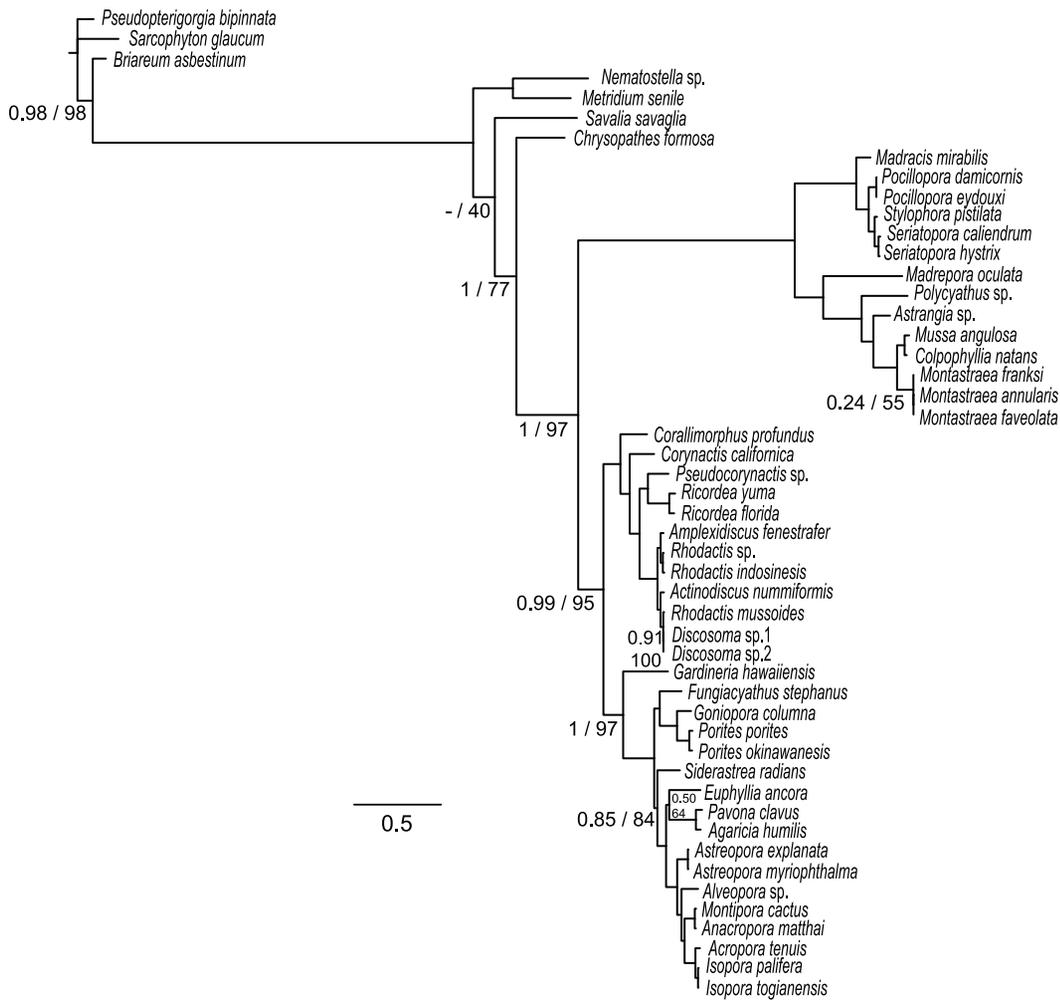
Taken together, these results based on models that do not assume compositional homogeneity or time reversibility suggest that the strong support of nt-based phylogenies for scleractinian monophyly might be an artifact of sequence composition. However, phylogenetic analyses carried out on a concatenated rRNA and tRNA alignment recovered a monophyletic Scleractinia clade with high statistical support irrespective of the method of analysis (Figure 5 and Table 3). Using this alignment, quartet puzzling with a Barry and Hartigan model also favored scleractinian monophyly in 99.9% of cases.

**Compositional bias effects: codons and amino acids**

Clear biases in codon usage are seen throughout the Hexacorallia, but in "Robust" corals, this bias is more extreme, as

evidenced by consistently lower effective number of codon (NC) scores and higher codon adaptation indices (calculated using CodonW [73]) than other scleractinians or corallimorpharians (Figure 6). For some amino acids, codon usage in "Robust" corals differed markedly from that in the other hexacorallians for which data are available (Figure 2). This pattern was also seen in the AT skew analyses (Figure S4).

In both aa and codon based phylogenies (Figures 3b and 4 respectively), the branch leading to the "Robust" coral clade is disproportionately long due to the presence of a large number of aa substitutions that are shared across most or all members of this clade but are not seen in other hexacorallians. Phenylalanine is the second most abundant aa in the mt-encoded proteins of "Robust" corals, and is approximately 1.5-fold more abundant in members of this clade, compared to other hexacorallians (Figure 2). The observed increase in abundance of (TTT-encoded) Phe residues in



**Figure 4. Codon-based phylogenetic analyses inferred using CodonPhyml – Yap et al. model [52] for Maximum Likelihood – and MrBayes – GTR substitution model and three categories of non-synonymous/synonymous ratios (M3 model).** Values on the nodes indicate the non-parametric SH test and bootstrap replicates (ML), and posterior probability (BI) support respectively. Where no values are shown on a node, that edge was fully supported under all analyses. doi:10.1371/journal.pone.0094774.g004

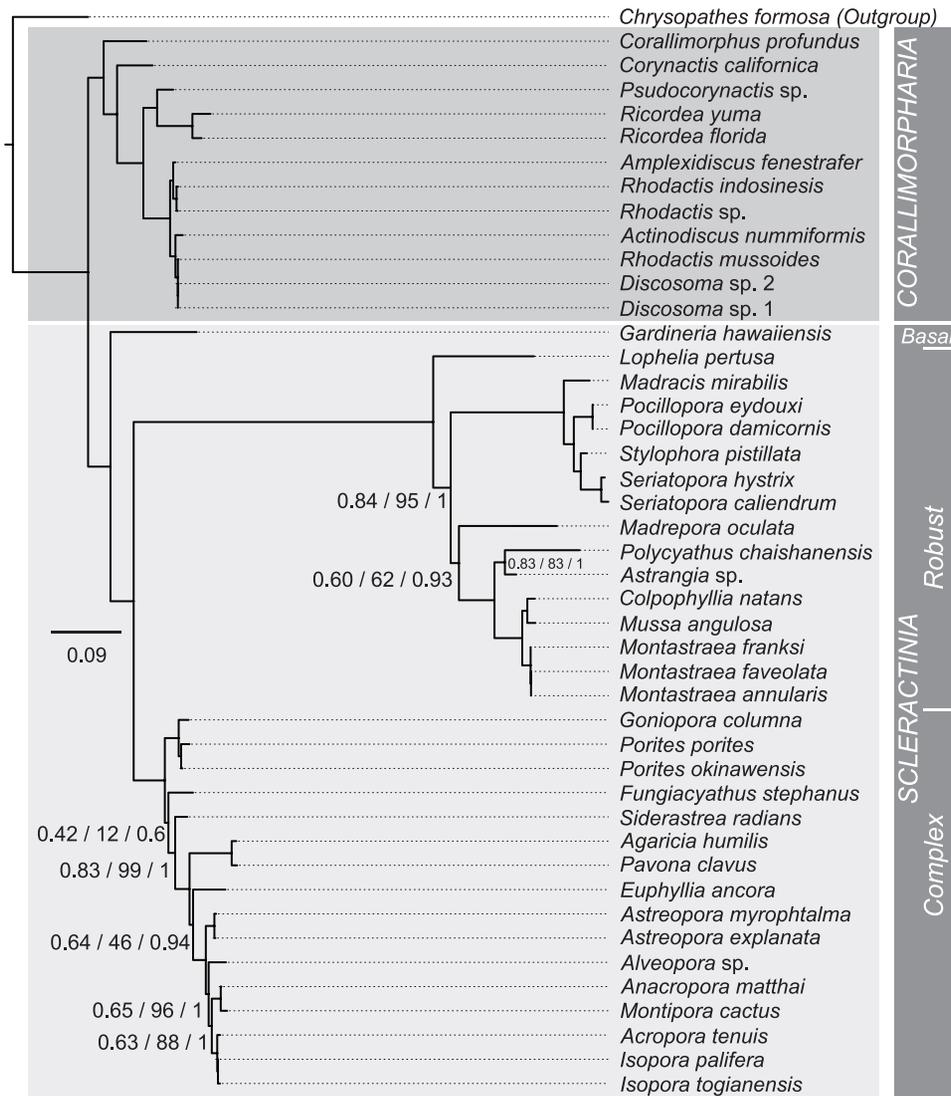
**Table 3.** Comparison of the two competing topologies (scleractinian monophyly, SM, and “naked coral”, NC) using the Approximately Unbiased (AU), Kishino-Hasegawa (KH) and Shimodaira-Hasegawa (SH) tests for a variety of tree reconstruction methods.

	Best ML topology	AU	KH	SH
Nucleotides (GTR+G+I+F)	SM	0.10	0.11	0.11
Nucleotides (GG98)	NC	0.33	0.32	0.32
Codons (YAP+W+K+F)	NC	0.03	0.03	0.03
Amino Acids (JTT+G+I+F)	NC	0.007	0.008	0.008
Nucleotides (rRNA, tRNA, GTR+G+I+F)	SM	0.009	0.013	0.013
Nucleotides (rRNA, tRNA, nhPhyML)	SM	0.086	0.094	0.094

The p-values denote the probability that the best ML topology is equivalent to the alternative topologies. Unless otherwise indicated, the trees were based on the alignment of protein coding genes. doi:10.1371/journal.pone.0094774.t003

“Robust” corals suggests that shifts in nt abundance may have forced the large number (>18%) of changes at the aa level that are unique to and shared within the “Robust” clade. (A+T)-rich amino acids (FYMICK) are more abundant, and (G+C)-rich amino acids (GARP) less abundant in the proteins encoded by the mt genomes of “Robust” corals than in other hexacorallians (Table 2).

To further investigate the possibility of long-branch artifacts on the amino-acid-based phylogenies, the PhyloBayes program was employed to infer trees using the CAT-Poisson, CAT-GTR and GTR models [74]. The CAT model allows different positions to evolve using a distinct substitution process and to have a different equilibrium frequency. It has been shown that this type of model is less sensitive to saturation and can reduce long-branch artifacts [75]. Table 4 summarizes the topologies and posterior probabilities obtained with this approach. The CAT-Poisson and CAT-GTR models tend to support scleractinian monophyly, while the GTR model is consistent with the naked coral hypothesis. These results are consistent with a recent study [27], where the monophyly of scleractinians received a modest support from the



**Figure 5. Phylogenetic analyses of the mitochondrial sequences encoding stable RNAs (12S rRNA, 16S rRNA, *trnM*, and *trnW*).** Values on the nodes indicate the non-parametric SH test and bootstrap replicates (ML), and posterior probability (BI) support respectively. Where no values are shown on a node, that edge was fully supported under all analyses. Large boxes indicate the Corallimorpharia (dark-gray) and Scleractinia (light-gray) clades. Note that deep-water azooxanthellate species (*Corallimorphus profundus* and *Gardineria hawaiiensis*) represent the earliest diverging branches for Corallimorpharia and Scleractinia respectively. doi:10.1371/journal.pone.0094774.g005

CAT-GTR model, but was not supported by the GTR model or by any nucleotide-based phylogenies.

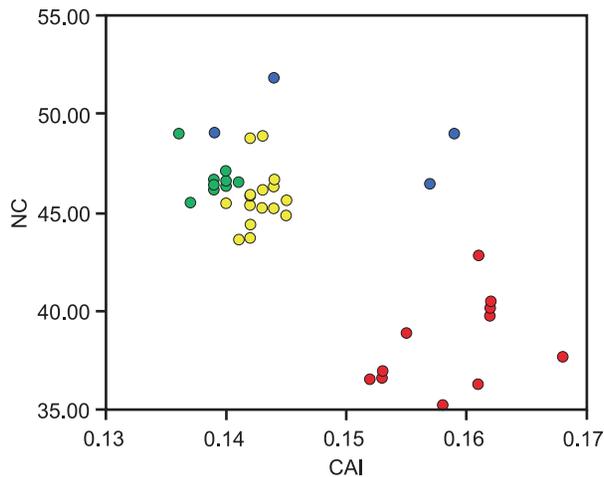
#### Hypothesis: did impaired mt DNA repair and constraints on tRNA uptake result in the altered amino acid usage in “Robust” corals?

If mt DNA repair is an ancestral property within Anthozoa [76], then the faster rate of mt genome evolution and differences in base composition that characterize “Robust” corals may reflect decreased efficiency of the repair process in this clade (see also [77]); if the resulting mismatches were not repaired, spontaneous deamination of cytosine to uracil could account for the atypical base composition. A consequence of the atypical base composition in the mt genomes of the “Robust” corals (higher T and lower C when compared to other hexacorallians) is a shift in the amino acid

composition of mt proteins towards those encoded by T-rich codons.

The mt genomes of Anthozoa differ from those of the Bilateria in encoding only two tRNAs – tRNA<sup>Met</sup> and tRNA<sup>Trp</sup> [78], hence in anthozoans most of the tRNAs required for mt translation must be imported. The mt tRNA uptake systems of anthozoans clearly have specificity, as codon use differs between mitochondrial and nuclear genes in *Acropora* despite similar overall base composition; for example, TTT being the strongly favoured Phe codon in mt-genes but bias being much less apparent in the case of nuclear genes [79,80].

“Robust” corals consistently display higher %(A+T) (around 5 to 6% higher than in “Complex” corals, for example) than either corallimorpharians or “Complex” corals, the most obvious difference being an increased frequency of thymine at third codon positions on the coding strand. In organisms that must import



**Figure 6. Codon usage in hexacorallian anthozoans.** CodonW [73] was used to estimate codon usage biases; default settings were used in calculation of the codon adaptation index. NC: effective codon number. CAI: codon adaptation index. Colour coding and taxon choice is as shown in Figure 2.  
doi:10.1371/journal.pone.0094774.g006

most tRNAs into mitochondria, changes in the base composition of the mt genome may lead to changes at the amino acid level in the proteins that they encode; the higher % (A+T) in the mt genomes of “Robust” corals not only drives protein coding sequences towards (A+T)-rich codons but may also force non-silent changes towards amino acid residues that are encoded by (A+T)-rich codons [81]. Such a mechanism could account for the higher abundance of phenylalanine residues in proteins encoded by the mt genomes of “Robust” corals, due to the increased frequency of TTT codons. Hence many of the amino acid substitutions unique to the proteins encoded by the mt genomes of “Robust” corals likely reflect the compound effects of base composition changes and the constraints under which tRNA uptake operates. We hypothesize that these amino acid substitutions bias phylogenetic analyses based on mitochondrial amino acid sequence data, obscuring relationships amongst the major scleractinian clades and corallimorpharians.

Consistent with compositional biases affecting analyses based on amino acid sequence data, phylogenetic analyses based on the mitochondrial rRNA and tRNA sequences consistently resulted in monophyletic Scleractinia (Figure 5). Furthermore, BF distance topologies inferred to assess the potential for compositional bias to affect phylogenetic inference suggested that overall, the aa data slightly favors the “naked coral” hypothesis, whereas nt based BF topology appears to be more homogeneous (Figure S5). Although the issue remains equivocal, molecular support for the “naked coral” hypothesis may therefore be an artifact resulting from compositional bias and saturation between the two major scleractinian clades. Note that these results do not challenge the issue of skeleton ephemerality sensu Stanley and Fautin [21] in Scleractinia, but imply that corallimorpharians are not descendants of a scleractinian that had undergone skeletal loss.

Changes in DNA repair mechanisms in some clades would result, in an evolutionary history, in violation of the assumptions of the models used for inference. With the exception of that of Barry and Hartigan, all models of substitution assume time-homogeneity both within and between branches [82]. All models of recorded sequences, including the aa substitution models, are non-Markovian, which results in a non-linear relationship between the true

**Table 4.** Posterior probability of the topologies recovered by different models using Phylobayes (SM: scleractinian monophyly, NC: naked coral).

Model	Topology	Posterior probability
CAT GTR	SM	0.56
CAT Poisson	SM	0.94
GTR	NC	1

doi:10.1371/journal.pone.0094774.t004

substitution dynamics operating on the nucleotide sequences and what is inferred using these models [83]. Thus, the ambiguous results outlined above could be a consequence of a poor fit between the models (despite these being selected as fitting the data best from the collection of models tested) and the evolutionary process. We evaluated how well the models fit compared to the best-possible likelihood, as proposed by Goldman [50] and implemented in the PyCogent library [49]. In brief, this test compares the difference in likelihood between the fitted model and the best-possible likelihood (calculated without assuming any phylogenetic relationship between the sequences) to the distribution of difference between these two likelihoods that one would expect if the data were generated according the fitted model (Figure S6). These tests can only be carried out in the maximum likelihood framework, thus for nt data the GTR and Barry-Hartigan models were tested and for aa data the JTT model. P-values were computed based on 200 Monte Carlo simulations, and for all models the fitted likelihoods were vastly inferior compared to the best possible, confirming a poor agreement between the data and the models used, even the most general one (with the fewest assumptions).

## Conclusion

The hypothesis outlined here – that, for hexacorallians, analyses based on mitochondrial sequences may be intrinsically biased - can and should be tested when appropriate nuclear sequence data are available for a wide range of corals and corallimorpharians. Molecular phylogenetics has led to radical revisions in thinking about coral evolution, but such analyses have largely been based on mt sequence data. Similar problems with mitochondrial sequences have been highlighted for a number of other animal groups including mammals [84] and beetles [85]. Given the above concerns, it is important that the bias towards mt data is redressed, and coral phylogenetics more broadly be based on a wide range of nuclear loci.

## Supporting Information

**Figure S1 Phylogenetic analyses based on the nucleotide sequences of the mitochondrial genes encoding proteins with the exclusion of the 3<sup>rd</sup> codon position.** Values on the nodes indicate the posterior probability (BI) support. Where no values are shown on a node, that edge was fully supported under all analyses.  
(EPS)

**Figure S2 Phylogenetic analyses based on amino acid sequences of the mitochondrial genes encoding proteins removing the most rapidly evolving sites based on the mean (relative) evolutionary rate (ER) for each site under the JTT+G+F model.** A, B, and C indicate the topologies recovered excluding all those sites with ERs of  $\geq 2.99$ .

$\geq 1.99$ , or  $\geq 1.49$  respectively. For each ER reconstruction the topology with the highest log likelihood is shown. Values on the nodes indicate the ML bootstrap (100 replicates) support. (EPS)

**Figure S3 Phylogenetic analyses based on the nucleotide sequences of the mitochondrial genes encoding proteins re-coded as purines and pyrimidies (RY-coding see [44,45]) with the exclusion of the third codon position from the resulting alignment.** Values on the nodes indicate the posterior probability (BI) support. Where no values are shown on a node, that edge was fully supported under all analyses. (EPS)

**Figure S4 Graphical representation of (G+C)- (red line) and (A+T)- (blue line) skew calculated on the whole mitochondrial genome of all species included in the present analysis.** The (A+T)-skew is  $>6\%$  higher in "Robust" corals than in all other hexacorallians included in the present analyses (highlighted in yellow). (EPS)

**Figure S5 Minimum evolution tree on BF distances from the complete mt protein coding DNA sequences.** Topologies are based on nucleotide BF distances (left topology) and aa BF distances (right topology). AA compositional bias slightly favors the "naked coral" hypothesis (yellow box) whereas nt based BF topology appears to be more homogeneous. (EPS)

**Figure S6 Empirical distribution of the difference between the likelihood of the fitted model and the best possible likelihood (the product of column pattern frequencies).** In each case the arrow indicates the observed value of that difference. The distributions are shown for the nucleotide alignment of the protein coding sequences with the GTR model (A) and the Barry and Hartigan model (B), for the

amino acid alignment with the JTT model (C), and for the RNA alignment with the GTR model (D). (EPS)

**Table S1 Primer names and sequences used for the amplification/sequence of the mitochondrial genome of *Gardineria hawaiiensis*.** The position and amplicons length of primers designed in the present study or the reference for previously published primers are provided. (DOC)

**Table S2 Alignment of the nucleotide sequences from the mitochondrial genes encoding proteins re-coded as purines and pyrimidines (RY-coding see [44,45]).** (TXT)

## Acknowledgments

The authors thank members of the Coral Reef Evolutionary Ecology and Genetics Laboratory (CREEG), Biodiversity Research Center, Academia Sinica (BRCAS) for assistance with sampling and field logistics. MVK acknowledges the support of the São Paulo Research Foundation (FAPESP) and São Paulo University Marine Biology Centre (CEBIMar). MFL gratefully acknowledges receipt of a James Cook University Postgraduate Research Scholarship (JCUPRS). This research was undertaken with the assistance of resources provided to SF at the NCI National Facility systems at the Australian National University through the National Computational Merit Allocation Scheme supported by the Australian Government. Coral samples were collected under appropriate permits. This is the CREEG, RCBAS contribution no. 95.

## Author Contributions

Conceived and designed the experiments: MK MFL DJM CAC. Performed the experiments: MK MFL. Analyzed the data: MK MFL SF GH. Contributed reagents/materials/analysis tools: DJM CAC. Wrote the paper: MK MFL SF DJM CAC.

## References

- Roberts JM, Wheeler AJ, Freiwald A, Cairns SD (2009) Cold-water corals: the biology and geology of deep-sea coral habitats. Cambridge: Cambridge University Press. xvii+334 p.
- Dower JF, Perry RI (2001) High abundance of larval rockfish over Cobb Seamount, an isolated seamount in the Northeast Pacific. *Fish Oceanogr* 10: 268–374.
- Fischer E, Jones G (2012) Atmospheric dimethylsulphide production from corals in the Great Barrier Reef and links to solar radiation, climate and coral bleaching. *Biogeochemistry* 110: 31–46.
- Kleypas JA, Danabasoglu G, Laugh JM (2008) Potential role of the ocean thermostat in determining regional differences in coral reef bleaching events. *Geophys Res Lett* 35: L03613.
- Reaka-Kudla ML (1997) Biodiversity II: understanding and protecting our biological resources. In: Reaka-Kudla ML, Wilson DE, Wilson EO editors. Washington DC: Joseph Henry Press. pp. 83–108.
- Reed JK (2002) Deep-water *Oculina* coral reefs of Florida: biology, impacts, and management. *Hydrobiologia* 471: 43–55.
- Rogers AD (1999) The biology of *Lophelia pertusa* (Linnaeus, 1758) and other deep-water reef-forming corals and impacts from human activities. *Hydrobiologia* 84: 315–406.
- Barbeitos M, Romano S, Lasker HR (2010) Repeated loss of coloniality and symbiosis in scleractinian corals. *Proc Natl Acad Sci USA* 26: 11877–11882.
- Fukami H, Chen CA, Budd AF, Collins A, Wallace C, et al. (2008) Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS One* 3: e3222.
- Kitahara MV, Cairns SD, Miller DJ (2010a) Monophyletic origin of the *Caryophyllia* (Scleractinia; Caryophylliidae), with description of six new species. *Syst Biodivers* 8: 91–118.
- Kitahara MV, Cairns SD, Stolarski J, Blair D, Miller DJ (2010b) A comprehensive phylogenetic analysis of the Scleractinia (Cnidaria, Anthozoa) based on mitochondrial CO1 sequence data. *PLoS One* 5: e11490.
- Kitahara MV, Cairns SD, Stolarski J, Miller DJ (2012a) Deltocyathiidae, an early-diverging family of Robust corals (Anthozoa, Scleractinia). *Zool Scr* 42: 201–212.
- Kitahara MV, Stolarski J, Cairns SD, Benzoni F, Stake JL, et al. (2012b) The first modern solitary Agariciidae (Anthozoa, Scleractinia) revealed by molecular and microstructural analysis. *Invertebr Syst* 26: 303–305.
- Le Goff-Vitry MC, Rogers AD, Baglow D (2004) A deep-sea slant on the molecular phylogeny of the Scleractinia. *Mol Phylogenet Evol* 30: 167–177.
- Romano SL, Cairns SD (2000) Molecular phylogenetic hypotheses for the evolution of scleractinian corals. *Bull Mar Sci* 67: 1043–1068.
- Romano SL, Palumbi SR (1996) Evolution of scleractinian corals inferred from molecular systematics. *Science* 271: 640–642.
- Stolarski J, Kitahara MV, Miller DJ, Cairns SD, Mazur M, et al. (2011) The ancient evolutionary origins of Scleractinia revealed by azooxanthellate corals. *BMC Evol Biol* 11: 2–15.
- Fukami H, Budd AF, Paulay G, Sole-Cava A, Chen CA, et al. (2004) Conventional taxonomy obscures deep divergence between Pacific and Atlantic corals. *Nature* 427: 832–835.
- Huang D, Licuanan WY, Baird AH, Fukami H (2011) Cleaning up the "Bigmessidae": molecular phylogeny of scleractinian corals from Faviidae, Merulinidae, Pectiniidae and Trachyphylliidae. *BMC Evol Biol* 11: 37.
- Huang D (2012) Threatened reef corals of the world. *PLoS ONE* 7: e34459.
- Stanley GD Jr, Fautin DG (2001) The origin of modern corals. *Science* 291: 1913–1914.
- Fine M, Tchernov D (2007) Scleractinian coral species survive and recover from decalcification. *Science* 315: 1811.
- Rodolfo-Metalpa R, Martin S, Ferrier-Pages C, Gattuso JP (2011) Response of the temperate coral *Cladocora caespitosa* to mid- and long-term exposure to pCO<sub>2</sub> sub and temperature levels projected for the year 2100 AD. *Biogeosciences* 7: 289–300.
- Medina M, Collins AG, Takaoka TL, Kuehl JV, Boore JL (2006) Naked corals: skeleton loss in Scleractinia. *Proc Natl Acad Sci USA* 103: 96–100.
- Culligan KM, Meyer-Gauen G, Lyons-Weiler J, Hays JB (2000) Evolutionary origin, diversification and specialization of eukaryotic MutS homolog mismatch repair proteins. *Nucl Acids Res* 28: 463–471.
- Stephenson TA (1921) On the classification of Actiniaria. Part II. Consideration of the whole group and its relationships, with special reference to forms not treated in part I. *Quart J Micr Sci* 65: 493–576.

27. Kayal E, Roure B, Philippe H, Collins AG, Lavrov DV (2013) Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evol Biol* 13: 5.
28. Stampar SN, Maronna MM, Kitahara MV, Reimer JD, Morandini AC (2014) Fast-evolving mitochondrial DNA in Ceriantharia: a reflection of Hexacorallia paraphyly? *PLoS One* 9: e86612.
29. Huang D, Meier R, Todd PA, Chou LM (2008) Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *J Mol Evol* 66: 167–174.
30. Shearer TL, Coffroth MA (2008) DNA BARCODING: Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Mol Ecol Resour* 8: 247–255.
31. Pont-Kingdon GA, Okada NA, Macfarlane JL, Beagley CT, Wolstenholme DR, et al. (1995) A coral mitochondrial MutS gene. *Nature* 375: 109–111.
32. Tseng CC, 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* 24: 502–508.
33. Cheng S, Higuchi R, Stoneking M (1994) Complete mitochondrial genome amplification. *Nature Genet* 7: 350–351.
34. Chen CA, Yu JK (2000) Universal primers for amplification of mitochondrial small subunit ribosomal RNA-encoding gene in scleractinian corals. *Mar Biotechnol* 2: 146–503.
35. Folmer O, Black M, Hoch W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3: 294–299.
36. 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.
37. Lin MF, Luzon KS, Licuanan WY, Ablan-Lagman CM, Chen AC (2011) Seventy-four universal primers for characterizing the complete mitochondrial genomes of scleractinian corals (Cnidaria; Anthozoa). *Zool Stud* 50: 513–524.
38. Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20: 3252–3255.
39. Stothard P (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 28: 1102–1104.
40. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
41. Katoh K, Kuma K, Toh H, Miyata T (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33: 511–518.
42. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704.
43. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics Applications Note* 17: 754–755.
44. Harrison GL, McLenachan PA, Phillips MJ, Slack KE, Cooper A, et al. (2004) Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. *Mol Biol Evol* 21: 974–983.
45. Phillips MJ, Gibb GC, Crimp EA, Penny D (2010) Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. *Syst Biol* 59: 90–107.
46. Hrdy I, Hirt RP, Dolezal P, Bardonova L, Foster PG, et al. (2004). *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* 432: 618–622.
47. Rodríguez-Ezpeleta N, Brinkmann H, Roure B, Lartillot N, Lang F, et al. (2007) Detecting and overcoming systematic errors in genome-scale phylogenies. *Syst Biol* 56: 389–399.
48. Boussau B, Gouy M (2006) Efficient Likelihood Computations with Non-Reversible Models of Evolution. *Syst Biol* 55: 756–68.
49. Knight R, Maxwell P, Birmingham A, Carnes J, Caporaso JG, et al. (2007) PyCogent: a toolkit for making sense from sequence. *Genome Biol* 8: R171.
50. Goldman N (1993) Statistical tests of models of DNA substitution. *J Mol Evol* 36: 182–198.
51. Gil M, Zanetti MS, Zoller S, Anisimova M (2013) CodonPhyML: fast maximum likelihood phylogeny estimation under codon substitution models. *Mol Biol Evol* 30: 1270–1280.
52. Yap VB, Lindsay H, Eastale S, Huttley G (2010) Estimates of the effect of natural selection on protein-coding content. *Mol Biol Evol* 27: 726–734.
53. Lartillot N, Lepage T, Blanquart S (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25: 2286–2288.
54. Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246–1247.
55. Rosenberg MS, Kumar S (2003) Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference. *Mol Biol Evol* 20: 610–621.
56. Phillips MJ, Penny D (2003) The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol Phylogenet Evol* 28: 171–185.
57. Riemann-Zürneck K, Iken K (2003) *Corallimorphus profundus* in shallow Antarctic habitats: bionomics, histology, and systematics (Cnidaria: Hexacorallia). *Zool Verh Leiden* 345: 367–386.
58. Stolarski J (1995) Ontogenetic development of the thecal structures in Caryophylliinae scleractinian corals. *Acta Palaeontol Pol* 40: 19–44.
59. Chen C, Dai CF, Plathong S, Chiou CY, Chen CA (2008a) The complete mitochondrial genomes of needle corals, *Seriatopora* spp. (Scleractinia: Pocilloporidae): an idiosyncratic *atp8*, duplicated *tmw* gene, and hypervariable regions used to determine species phylogenies and recently diverged populations. *Mol Phylogenet Evol* 46: 19–33.
60. Park E, Song J, Won Y (2011) The complete mitochondrial genome of *Calicogorgia granulosa* (Anthozoa: Octocorallia): potential gene novelty in unidentified ORFs formed by repeat expansion and segmental duplication. *Gene* 486: 81–87.
61. Chen C, Chiou CY, Dai CF, Chen CA (2008b) Unique mitogenomic features in the scleractinian family Pocilloporidae (Scleractinia: Astrocoeniina). *Mar Biotechnol* 10: 538–553.
62. Fukami H, Knowlton N (2005) Analysis of complete mitochondrial DNA sequences of three members of the *Montastraea annularis* species complex (Cnidaria, Anthozoa, Scleractinia). *Coral Reefs* 24: 410–417.
63. Park E, Hwang DS, Lee JS, Song JI, Seo TK, et al. (2012) Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Mol Phylogenet Evol* 62: 329–345.
64. van Oppen MJ, Catmull J, McDonald BJ, Hislop NR, Hagerman PJ, et al. (2000) The mitochondrial genome of *Acropora tenuis* (Cnidaria; Scleractinia) contains a large group I intron and a candidate control region. *J Mol Evol* 55: 1–13.
65. Lin MF, Kitahara MV, Tachikawa H, Fukami H, Miller DJ, et al. (2012) Novel organization of the mitochondrial genome in the deep-sea coral, *Madrepora oculata* (Hexacorallia, Scleractinia, Oculinidae) and its taxonomic implications. *Mol Phylogenet Evol* 65: 323–328.
66. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512–526.
67. Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22: 160–174.
68. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8: 275–282.
69. Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* 18: 691–699.
70. Dayhoff M, Schwartz R, Orcutt B (1978) A model of evolutionary change in protein. In: Dayhoff MO editor. *Atlas of protein sequences and structure*. Washington, DC: National Biomedical Research Foundation. pp. 345–351.
71. van Den Bussche RA, Baker RJ, Huelsenbeck JP, Hillis DM (1998) Base compositional bias and phylogenetic analyses: a test of the “flying DNA” hypothesis. *Mol Phylogenet Evol* 13: 408–416.
72. Galtier N, Gouy M (1998) Inferring pattern and process: Maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol Biol Evol* 15: 871–879.
73. Peden J (2005) CodonW version 1.4.2. Available: <http://codonw.sourceforge.net/culong.html#License>
74. Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol Biol Evol* 21: 1095–1109.
75. Lartillot N, Brinkmann H, Philippe H (2007) Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evol Biol* (Suppl 1): S4.
76. Shearer TL, van Oppen MJH, Romano SL, Worheide G (2002) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol Ecol* 11: 2475–2487.
77. Aranda M, DeSalvo MK, Bayer T, Medina M, Voolstra C (2012) Evolutionary insights into scleractinian corals using comparative genomic hybridizations. *BMC Genomics* 13: 501.
78. 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.
79. Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, et al. (2011) Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476: 320–323.
80. van Oppen MJH, Hislop NR, Hagerman PJ, Miller DJ (1999) Gene content and organization in a segment of the mitochondrial genome of the scleractinian coral *Acropora tenuis*: major differences in gene order within the anthozoan subclass Zoantharia. *Mol Biol Evol* 16: 1812–1215.
81. Knight RD, Freeland SJ, Landweber LF (2001) A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol* 2: 4.
82. Barry D, Hartigan JA (1987) Statistical analysis of Hominoid molecular evolution. *Stat Sci* 2: 191–207.
83. Kosiol C (2006) Markov models for protein sequence evolution. [Ph.D. thesis]. EMBL-European Bioinformatics Institute. Cambridge: University of Cambridge.
84. Huttley G (2009) Do genomic datasets resolve the correct relationship among the placental, marsupial and monotreme lineages? *Austral J Zool* 57: 167–174.
85. Sheffield NC, Song H, Cameron SL, Whiting MF (2009) Nonstationary evolution and compositional heterogeneity in beetle mitochondrial phylogenomics. *Syst Biol* 58: 381–394.
86. 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.

87. 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.
88. Flot JF, 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.
89. Sinniger F, Chevaldonné P, Pawlowski J (2007) Mitochondrial genome of *Savalia savaglia* (Cnidaria, Hexacorallia) and early metazoan phylogeny. *J Mol Evol* 64: 196–203.
90. Emblem A, Karlsen BO, Evertsen J, Johansen SD (2011) Mitogenome rearrangement in the cold-water scleractinian coral *Lophelia pertusa* (Cnidaria, Anthozoa) involves a long-term evolving group I intron. *Mol Phylogenet Evol* 61: 495–503.
91. Adachi J, Waddell PJ, Martin W, Hasegawa M (2000) Plastid genome phylogeny and a model of amino acid substitution for protein encoded by chloroplast DNA. *J Mol Evol* 50: 348–358.
92. Daly M, Fautin DG, Cappola VA (2003) Systematics of the Hexacorallia (Cnidaria: Anthozoa). *Zool J Linn Soc-Lond* 139: 419–437.
93. Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures Math Life Sci* 17: 57–86.