Peer-reviewed version available at *Genes* **2019**, *10*, 324; doi:10.3390/genes10050324

Computational characterization of the mtORF of pocilloporid corals: insights into differences in protein structure and function among Stylophora lineages from contrasting environments

*Eulalia Banguera-Hinestroza,[1,2], Yvonne Sawall[3], Jean-Francois Flot,[1,2]

[1]. Université libre de Bruxelles, Evolutionary Biology and Ecology, Belgium.

[2]. Interuniversity Institute of Bioinformatics in Brussels – (IB)[2], Brussels, Belgium

[3]. Bermuda Institute of Ocean Sciences (BIOS), Coral Reef Ecology, Bermuda.

*Corresponding authors: ebanguer@ulb.ac.be; eulalia.banguera@gmail.com

Abstract

More than a decade ago, a new mitochondrial Open Reading Frame (mtORF) was discovered in corals of the family Pocilloporidae, which turn out to be an effective barcode gene for these corals. However, its function remains unknown. Recently, this gene revealed the existence of a hybrid *Stylophora* lineage (*RS_LinA*) inhabiting in sympatry along the environmental gradient of the Red Sea (18.5°C to 33.9°C) with its parental species (*RS_LinB*). Furthermore, in *RS_LinB*, the mtORF uncovered phylogeographic patterns that were strongly correlated with environmental variations. This was similar to the patterns unraveled by *hsp70,* suggesting that mtORF too might be involved in thermal adaptation. Here we used computational approaches to characterize the mtORF and to identify its potential role. Results showed that this gene encodes a transmembrane protein ($0.97<P< 1.00$) involved in transport ($0.80<P< 0.87$), regulation of metabolic processes ($0.70<P<0.85$), and likely in the cell-surface receptor signaling pathway ($0.56<P<0.80$). Predicted protein functions differed among *Stylophora* lineages and interestingly, in *RS_LinB* only, the protein was intrinsically disordered and displayed domains involved in cellular complexes and stress response ($0.0001< P <0.001$). These characteristics, exclusive of an endemic lineage adapted to extreme environmental fluctuations, support a role of the mtORF in stress response, speciation and adaptation.

Key words*: protein function, mtORF, Stylophora, thermal adaptation, disordered amino acids residues, pocilloporid corals, a transmembrane protein, pocilloporid corals.*

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

1. Introduction

New mitochondrial open reading frames (mtORFs) are occasionally reported in the literature; in some cases, they have been proven to encode functional proteins, with a genetic variability sufficient to unravel patterns of differentiation in populations and species from different geographic origin (e.g., unicellular algae of the class Raphidophyceae [1]). In corals, a mtORF apparently unique to three genera of the family Pocilloporidae (*Pocillopora*, *Seriatopora* and *Stylophora*) was reported by Flot and Tillier 2007 [2]. This mtORF is also found in the genus *Madracis* (this study) –of the same coral family [3]–, but it was annotated as a control region in the only mitogenome of this genus available to date [4], leading to some confusion in the literature.

The mtORF gene has been useful for the delimitation of species within *Pocillopora* [2,5,6] and has allowed the identification of paraphyletic morphospecies in *Seriatopora,* as well as fine-scale genetic structure in species belonging to this genus [7,8]. Furthermore, it has revealed strong phylogenetic and phylogeographic patterns in *Stylophora* [9,10]. This indicates that in the absence of highly variable mtDNA genes in corals [11,12], the mtORF may therefore be a useful mitochondrial barcode gene for pocilloporids [2,6].

mtDNA genes, in particular barcode genes (e.g., cytochrome oxidase subunit I gene *cox1* [13]), often unravel patterns of genetic variation that are highly correlated with changes in environmental conditions, mostly related to temperature and altitude ([13–17]. This is because they are involved in key molecular mechanisms related to mitochondrial bioenergetics. For example, they are key players in the mito-nuclear protein complexes of the oxidative phosphorylation [OXPHOS] system, which transform ADP (adenosine 5′-diphosphate) and phosphate to the cellular energy carrier adenosine 5′-triphosphate (ATP) ([18,19]. Hence, mtDNA genes play an important role in the thermal tolerance of an organism by regulating the availability of ATP whose demand is often increased in organisms thriving under environmental extremes [20–23].

Mutations at the mtDNA OXPHOS genes, enhanced by natural selection, has been found to have an important impact in the ecology and evolutionary history of multiple taxa [14–16,21,24–26]. These mutations produce mito-nuclear incompatibilities between genotypes adapted to markedly different environments, creating reproductive barriers and increasing genetic differentiation among populations (particularly in species distributed across broad geographical ranges), which are the main processes leading to speciation [13,21,23,27–29].

The involvement of mtDNA genes in adaptation to the environment and in speciation seems to be supported by the evolutionary patterns revealed by the mtORF barcode gene in the genus *Stylophora*. Recently, Banguera-Hinestroza et al. (2018) [9] analysed 827 samples of *Stylophora* covering a broad geographical range of this genus, including the full latitudinal (12 latitudes) and environmental gradient of the Red Sea (i.e. minimum and maximum Sea Surface Temperatures – SST– have been recorded between 18.5°C-29.9°C in northern areas and 23.5°C-33.9°C in southern regions [30]). They found a high number of mutational changes in the mtORF gene as well as in adjacent genes (*nd6 and atp6*). Interestingly, the three genes are part of a recombinant region that strongly differentiates a hybrid lineage (*RS_LinA*) from its parental species (*RS_LinB*), both inhabiting the entire environmental gradient of the Red Sea. Furthermore, in *RS_LinB*, the mtORF uncovered the existence of two well-differentiated populations restricted respectively to the colder northern regions or to the warmer central-southern Red Sea. The same phylogeographic pattern was observed for the *hsp70* gene encoding a heat-shock protein well known for its role in stress response and adaptation to different climatic conditions [31]. This suggests that both genes may have played a similar function in the adaptation of the ancestral and endemic *Stylophora* lineage to the different temperature regimes of the Red Sea including extremely warm conditions in the Southern region.

So far, the pocilloporids' mtORF gene has not been fully characterized and its function in the mitogenome is currently unknown, which limits our understanding of its possible role in adaptation and speciation. Therefore, the aim of the current study, is to characterize the mtORF and unravel its function, as well as to identify signals of natural selection in the mtORF proteins of species inhabiting different environments, including the extreme environments of the Red Sea.

Previous studies suggested that the mtORF gene might be a pseudogene [32,33], in which case multiple stop codons or frameshift mutations should be observed, leading to a non-functional protein [34,35]. Hence, as a first step, we analyzed hundreds of mtORF sequences from *Pocillopora*, *Seriatopora* and *Stylophora*, freely available in the NCBI data base and searched for stop codons in the translated sequences. In this analysis, we also included the only mtORF sequence available from the mitogenome of the genus *Madracis*. This approach was additionally used for identifying whether this gene encodes a functional protein, because long ORF (more than 600 bp in the case of this gene) are rarely maintained in genome, if they are not expressed [36].

As a second step, we used computational prediction methods available at the TMHMM Server [37,38] and at the PSIPRED Workbench [39] to analyze the structure and function of mt*ORF* in *Madracis*, *Pocillopora, Seriatopora*, and *Stylophora*, including the two *Stylophora* lineages previously identified from the Red Sea [9] –note that mt*ORF* in *italic* refers to the protein, while non-italic notation of ORF refers to the gene–. Both workspaces have been shown to be reliable for predicting protein structure and function for non-model species [40]. After characterizing the mt*ORF* protein, we tested whether it may be involved in local adaptation by searching for signals of selection in the sequences of *Stylophora* lineages [9], particularly those from the Red Sea. As mentioned above, these lineages occurred throughout the entire Red Sea in regions with strong environmental differences. The most northern region of the Red Sea (Gulf of Aqaba, 32°N) is characterized by the coldest temperature regime (18.5°C-28.8°C), highest salinity (41 PSU) and generally low nutrient input, while the most southern region (Farasan islands, 20°N) is characterized by an exceptionally warm temperature regime (23.5°C-33.9°C), lower salinity (37 PSU) and high nutrient input [30,41,42].

2. Material and methods

2.1. Characterization of the mtORF gene

A total of 1830 mtORF sequences of *Pocillopora*, *Seriatopora* and *Stylophora* publicly available in the NCBI data base [4,6–8,10,33,43,44] were translated using the coelenterate mitochondrial code implemented in the program MEGA [45] and scanned for stop codons. In addition, to identify whether this mtORF was also present in the genus *Madracis*, its mitogenome was downloaded from the NCBI data base (accession number: EU400212) and aligned with mitogenomes of the other three genera. After probing homology, the mt*ORF* was extracted, translated and searched for stop codons. Samples of *Stylophora* and *Pocillopora* from the Red Sea were taken exclusively from our collection [9,46] to ensure the reliability of the data. Structural differences among mtORF sequences and dissimilarities in length and composition of tandem repeats were identified using the Program Tandem repeats finder of Benson (1999) [47]. Sequences from *Stylophora* lineages included in this analysis correspond to the *Stylophora* hybrid (*RS_linA*), and sequences resembling those of the two parental species that gave origin to the hybrid lineage, this is: sequences of *Stylophora* from the Red Sea (*RS_LinB*) and sequences from *Stylophora* from the Indo-Pacific (referred here as *Stylophora pistillata*).

Second, the mtORF extracted from the full mitochondrial genomes of *Madracis mirabilis* (EU400212), *Stylophora pistillata* (accession: EU400214), *Seriatopora hystrix* (accession: EF633600), *Pocillopora damicornis* (accession: EF526303), and from *RS_LinA* and *RS_LinB* [9] were analysed using a set of protein prediction approaches available at the PSIPRED Protein Sequence Analysis Workbench (www.bioinf.cs.ucl.ac.uk/psipred [48]) and at the TMHMM Server v. 2.0. [37,38]. These methods included: (*i*) the PSIPRED approach (Protein Structure Prediction) by Jones (1999) [49], a highly accurate method that uses two feed-forward neural networks to predict secondary structure in outputs generated using PSI-BLAST (Position Specific Iterated BLAST [39,50]); (*ii*) the MEMSAT3 method (MEMbrane protein Structure And Topology 3) that predicts secondary structure of membrane protein by using multiple alignments produced by PSI-BLAST and by scoring log-likelihoods ratios through different topological models to gather the consensus [40,48,51,52]; and (*iii*) the MEMSAT-SVM approach using Support Vectors Machines, which are binary classifiers used to categorize residue preferences before combining them into a probabilistic framework [40].The outputs of these methods were

compared with the TMHMM approach that predicts transmembrane helices based on a hidden Markov model [37,38].

Further analyses using the PSIPRED webserver included: (*i*) The prediction of the likelihood of a transmembrane helix (TMH) being involved in the formation of pore lining regions. Those run parallel to transmembrane helices and are vital for biological processes, such as the transport of ions and molecules across the membrane [40]. (*ii*) The identification of disordered regions using the DISOPRED approach [53–55]. (*iii*) Classifying functional assignments using FFPred 2.0 [56], which uses the GOA (Gen Ontology Annotation data base) to predict Gene Ontology (GO), including macromolecular interactions, biological processes, molecular function and cellular components, searching throughout a large data set of already known proteins and annotations from Eukaryotes [56]. For each prediction, the posterior probabilities are given in terms of reliability that is composed of three characteristics: sensitivity, specificity and precision [57]. (*iv*) The prediction of structural domains involved in protein-protein interaction using pDomTHREADER [58]. (*v*) The distinction of TM regions from signal peptides (SP), using the signaIP software 4.1 [59,60] available at www.expasy.org. For a full description of the approaches see the PSIPRED web server [57].

2.2. Signatures of adaptive evolution in the mt*ORF* of *Stylophora* inhabiting different environments.

The potential role of the mt*ORF* protein in local adaptation was tested by searching for signals of selection in the translated mtORF sequences. For this, mtORF haplotypes belonging to the Red Sea *Stylophora* lineages, as well as to *Stylophora* species from other oceanic regions were used. Samples included in these analyses (N= 827) belong to the same set of samples analysed in Banguera-Hinestroza et al. (2018) [9], from which also the definition of *Stylophora* lineages, clades, subclades and populations, was taking. Briefly, two *Stylophora* lineages were distinguished within two highly divergent clades: Clade 1 included the linage *RS_LinA* (~50% of *Stylophora* corals collected within the Red Sea and Gulf of Aden) and *Stylophora* specimens from the Indo-Pacific, Madagascar, Arabian Gulf and Gulf of Aden regions. Clade 2 included

exclusively *Stylophora* specimens of the lineage *RS_LinB* (the remaining individuals from the Red Sea and Gulf of Aden areas). This Clade was further divided into two well defined groups (subclades), each grouping specimens distribute either in the northern areas of the Red Sea or in the central-southern regions.

For analyses of selection, first, all *Stylophora* sequences were analysed using regions that were unambiguously aligned among all haplotypes. In addition, sequences from Clade 1 (*RS_LinA* plus *Stylophora* from other oceanic basins) and Clade 2 (*RS_LinB*) were analysed separately. Last, selection was tested in sequences from the hybrid lineage (*RS_LinA*) and also in sequences from the northern and southern populations of the parental Red Sea lineage (*RS_LinB*).

Signatures of selection at individual sites on the mt*ORF* protein were inferred using several approaches implemented in the Datamonkey webserver (www.datamonkey.org) [61,62]. (*i)* The Fixed Effects Likelihood method (FEL), which detects selection by first estimating branch lengths and substitution rates parameters in a given phylogeny, when its corresponding coding alignment is provided. After the calculation of these two parameters the method work with fixed values to infer nonsynonymous (dN) and synonymous (dS) substitution rates per sites allowing the identification of selection along the branches of the phylogenetic tree [63]. (*ii*) The Mixed Effects Model of Evolution (MEME), which identifies episodic positive selection at individual sites [64] by combining both, fixed models and random effect models. (*iii*) The Branch-Site REL model (BS-REL -aBSREL-), a method that uses maximum likelihood computations to infer the variation of dN and dS over branches and from site to site [65,66]. Statistical significance levels were set at a *p*-value of 0.1 for MEME and FEL and at 0.5 for aBSREL as recommended by the authors.

Furthermore, the M8 model of Yang et al. (2000) [67] and the Mechanistic–Empirical Combination (MEC) model from Doron-Faigenboim and Pupko, 2006 [68] were applied to the mt*ORF* protein via the Selecton webserver [69,70]. Under these approaches, the best model that fits the data is found by performing a likelihood ratio test (LRT) over the whole alignment among several sequences. Here, a null model in which positive selection is not allowed (i.e. M8a; non-

codons with dN/dS > 1) is compared against a general model that assumes positive selection (i.e. dN/dS > 1) such as the M8 and the MEC model. If positive selection is detected, a Bayes empirical Bayes (BEB) approach is used to calculate the posterior probabilities of sites undergoing positive selection [71]. Statistical significance was tested comparing the AIC (Akaike Information Content) scores between the MEC model or M8 model against the M8a model. Note that the likelihood ratio test is used to compare both models and the significance test is passed whenever the AIC score of M8 or MEC is lower than the M8a score, using a threshold of 0.05. For details about the advantages of the MEC model, refer to Stern et al. (2007) [70].

The last test used to detect selection was the codon-based Z-test of positive selection of Nei and Gojobori, (1986) [72] implemented in MEGA 7 [45]. This test allows discrimination between positive and purifying selection by performing pairwise comparisons among all haplotypes within a specific clade/subclade against the others. This is done by comparing the relative abundance of synonymous and non-synonymous substitutions and their variances, by calculating dS and dN [45]. The null hypothesis of dN = dS and the alternative hypotheses of dN > dS (positive/diversifying selection) and dN < dS (purifying selection) are tested using one-tailed Z-test. Although this method is comparatively simple for testing selection hypotheses, it is considered to perform as well as more complex methods [73].

3. Results

3.1. General features of the mtORF-encoded protein

Stop codons were not found in the translated mtORF sequences of *Pocillopora*, *Seriatopora* and *Stylophora* (N=1830), except in 3 out of 431 sequences from *Seriatopora* (Accessions KR150027, KR150037 and KR150052), which showed an insertion of one nucleotide in position 165, 206 and 397 respectively, when aligned with all other sequences. We interpreted this as errors during base calls in the published sequences. The mtORF extracted from the mitogenome of *Madracis* was not interrupted by stop codon, as well, and showed to be homologous to the mtORF of other pocilloporids (the alignment of the mtORF of this genus with the other genera is

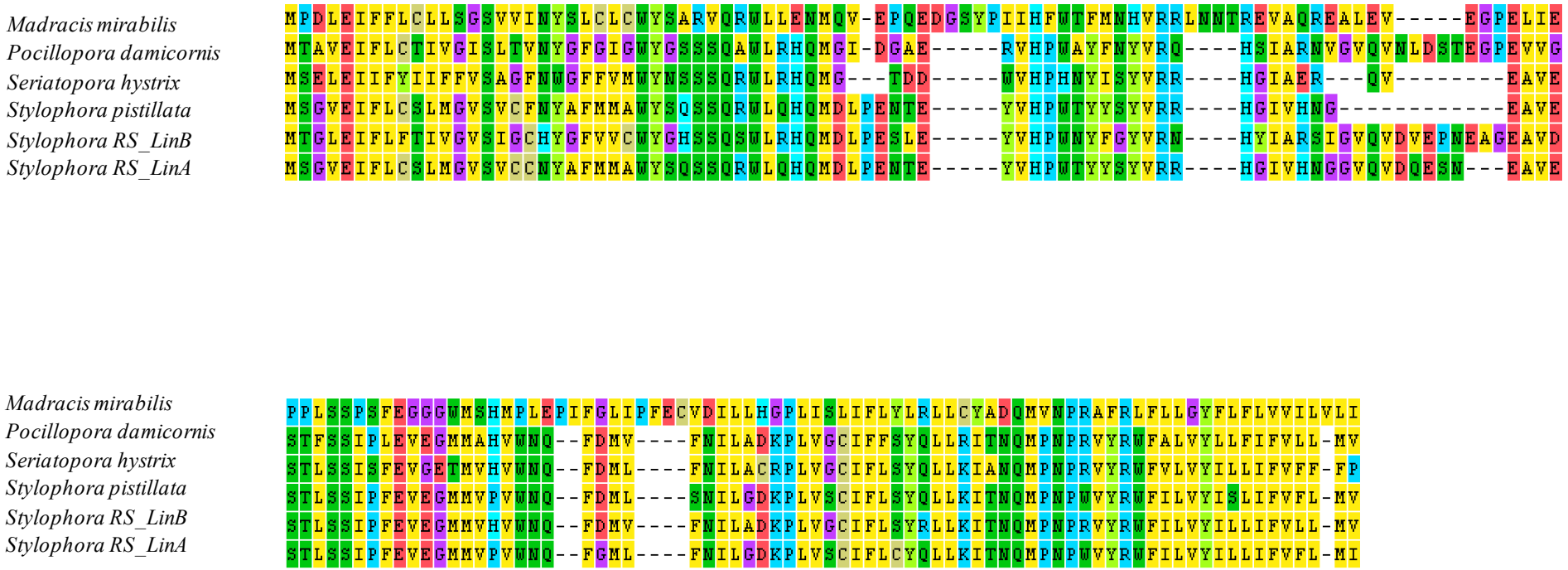
shown in Supplementary Figure *S1*). This suggests that the mtORF gene is fully functional and not a pseudogene, which was also corroborated by subsequent analyses of this study (see below).

Translating the mtORF sequences of each genus resulted in proteins of different lengths. The longest was found in *RS_LinB* with 362 amino acids, followed by *Stylophora pistillata* with 309 amino acids, *Pocillopora* with 302, *Stylophora RS_LinA* with 301, *Seriatopora* with 269, and *Madracis* with 221 amino acids. Within *Stylophora,* differences in length, particularly within the Red Sea group, were found to be associated with the presence of duplicated tandem repeats (TR), which are recorded in Table 1.

3.2. Prediction of transmembrane helices (TMH) and protein structures

Two TMH were predicted in the mtORF encoded protein (abbreviated as mt*ORF*-pro thereafter) of *Stylophora*, *Seriatopora* and *Pocillopora* and three TMH in the mt*ORF-pro* of *Madracis* by the TMHMM approach, with high posterior probabilities ($P > 0.97$) and with low probabilities of the N-term being in the cytoplasmic site of the membrane ($0,09 < P < 0.48$) (Figure 1). The presence of these helices was confirmed by the MEMSAT3 and MENSAT-SVM approaches, except for *Seriatopora*, in which three TMHs were predicted by MEMSAT-SVM. Furthermore, transmembrane topology predicted by MENSAT-SVM suggested a lack of pore-lining residues in *Madracis*, but the existence of two pore-lining helixes in *Seriatopora* and one in *Pocillopora* and *Stylophora*. No evidence for signal peptides, which are often mistaken with transmembrane helixes, was found using MEMSAT or signaIP.

Differences in the mt*ORF-pro* among genera were mainly found in the predictions of secondary structure (Figure 2) and intrinsically disordered regions –IDRs– (Figure 3). Only one disordered residue was predicted in *Madracis*, *Pocillopora* and *Seriatopora*, but multiple unstructured amino acid residues were found in the mt*ORF-pro* of *Stylophora* lineages: in the ancestral *RS_LinB,* the IDR expands 234 amino acids (from residue 63 to 296), in *Stylophora pistillata* 79 amino acids (from residue 123 to 201), and in the mt*ORF-pro* of the hybrid (*RS_LinA*) it covers two small regions with a total of 27 amino acids (from residue 134 to 153 and from 160 to 166). Noticeably,



Supplementary *Figure S1.* Conserved blocks in the mt*ORF-pro* of pocilloporid corals

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 January 2019 doi:10.20944/preprints201901.0232.v1

Peer-reviewed version available at *Genes* **2019**, *10*, 324; doi:10.3390/genes10050324

Table 1. Polymorphic Tandem Repeats (TR) in the mtORF of *Stylophora*

Cluster 1	Consensus sequence	Period Size	Copy Number	Consensus Size	Percent Indels	A	C	G	T
RS_LinA	AAGTGAGTCAGAAAGTAGAGGTAGTTTGAGGGATAAATGAAGGATT	51	2.1	51	0	38	2	33	25
RS_LinA	AAGTGAGTCAGAAAGTAGAGGTAGTTTGAGGGATAAATGAGGGATT	51	4.1	51	0	39	2	32	25
Madagascar	AAGTGAGTCAGGAGTAGAGGTAGTTTGAGGGGATAATAATTAAGAGTT	51	2.1	51	0	38	2-3	33	24-25
Indo-Pacific	AAGTGAGTCAGAAAGTAGAGGTAGTTTGAGGGAAAGTGTAATTAAGGGATT	51	2.1	51	0	37	1	35	25
RS_LinA	ATGACGACAGTTTGTGTGATAATGATCGCTGTGTGTTC	39	2.6	42	4	22	7	33	37
Madagascar	TGTGCTATGACGAGAGTTTGTGTGAATGATAGATTTG	39	2.4	39	7	21	6	34	37
Madagascar	ATGACGACAGTTTGTGTGAATTGATAATCGTTTGAGTC	39	2.7	39	4	23	7	33	36
Indo-Pacific	ATGACGACAGTTTGTGTGAATTGATAATCGTTTGTGTTG	39	3.2	42	6	22	6	33	37
Indo-Pacific	ATGACGAGAGTTTGTGTGATAATGATAATCGTTTGTGT	39	2.6	39	0	24	4	32	38
RS_LinA	TGTGCTATGACGAGAGTTTG	21	3.7	20	12	20	6	36	36
RS_LinA	GTTTGTGTGATAATGATGATA	18-21	2.8-3.8	21	10-16	21-23	1-3	30-32	42-43
RS_LinA	AATGAGAGAGTTTGTGTGAT	18	2.4	21	13	29	4	33	33
RS_LinA	AGTTTGTGTGAATGAT	18	1.9	18	0	25	0	31	42
Indo-Pacific	GTTTGTGGATAATGGATC	18	4.8	19	12	21	4	32	41

Cluster 2 (RS_LinB)	Consensus Sequence	Period Size	Copy Number	Consensus Size	Percent Indels	A	C	G	T
Region within the Red Sea									
Northern and Southern	AATAGGGGATCAGTGATTC	21	6-11	21	10	32-34	9	29-31	26
Northern	TAGAATTCGGGGATTAAGTGATCAAG	27	8-20	27	5-12	34-35	9	24-27	28-30
Northern and Southern	CAATAGGGGATCAAGTGATTCAAGTA	27	4-7	27	7-15	34-35	10-11	27-29	25-26
Northern	AAGTAGGGATTAAGTGATATGTACAAGTGGGGATTAAGTGATTC	48	3.1	47	19	34	7	26	30
Northern	CAATAGGGGATCAAGTGATTCAAGAATGCGGCATCAAGTGATT	48	6-7	48	16-17	33-34	9-10	27-28	27-28
Northern	AATGGGGGATTAAGTGATCAACTACACTAGGGGATTAAGTGATCAAGTG	48	5-9	54	13-17	34-35	9-10	25-27	28-29
Northern	CAATTAAGGATCGAGTGATAAGTACAAAGTGAGTGAAGTGATTCAAGTA	51	1.9	52	6	36	10	27	24
Northern	AAGTGATCAAGTGAATTAGGGATGAAGTGATTCAAAGGGGAGCAAG TGATTCAATTAGAGGATC	69	2.6	69	0	33-34	8	29-30	27
Northern	AAGTGATCAAGTGAATTAGGGATGAAGTGATTCAATT AGGGGATCAAGTGGTTCAATTAGGGGATC	69	3.5	69	9	34	9	28	27
Northern and Southern	GGATCAAGTGATCAAGTGAATTAGGGGATCAAGTGATTCA ATTAGGGGAGCAAGTGATTCAATTAGA	69	4-6	69	3-11	33-34	9-10	26-27	28
Northern	CAATTGGAGATCAAGTGATTCAAGAATCGAGCATTAAG TGATTCAAGTAGAATGAGGGCATCAAGTGATT	75	3.3	75	5	33	9	27	28
Northern	CAATAGGGGATCAAGTGATCAAGTGAATAGGGGATCA AGTGATTCAATTAGGGGATCAAGTGATT	75	5.3	69	6	33	9	27	28
Northern	ATCAAGTGATTCATAGAATTGGGGATTAAGTGACTAG TAGAATGAGGGCATCAAGTGATTCAATTAGGAT	75	5.9	75	8	34	10	26	28
Southern	AATAGGGGATCAAGTGATCATC	21	8.9	24	16	33	10	29	26
Southern	AATTGGGGATTAAGTGATTCAACTAC	27	8-17	27	7-10	34-35	10	24-27	28
Southern	CAATTAGGGGTCAAGTGATTCAATTAGGATGAAGTGATTCAAGTA	48	2.5	48	0	34	10	28	26
Southern	AAGTGATTCAAGTATAAAGTGGGATCAAGTGATTCTAAGAGACC	48	4.8	48	12	33	9	29	27
Southern	AAGTGATTCATACAAGGGGCATCAAGTGATTCAATTAGAGC	48	5.8	48	17	34	10	26	27
Southern	TAAGTGAAATTAGGGATCAAGTGATTACTAGGGGATCAAGTGA	48	5.6	46	7	34	9	29	26
Southern	AAGTGATCAAGTGAATTAGGGATCAAGTGATTCAAGGGGAC	48	5.8	48	5	34	9	28	27
Southern	AAGTGATTCAGATACAATGGGGCATCAAGTGTTCAATAGGAATC	48	6.8	47	16	34	10	27	27
Southern	AATTCGGGGATCAAGTGATTCAAGTCAACTAGGGGATTAAGTGATCAAGTAC	48	7.4-7.9	54	13-14	34	10	27	28
Southern	AATTCGGGGATCAAGTGATTCAAGTCAACTAGGGGATTAAGTGATCAACTA	48	7.9	54	12	33	10	27	28
Southern	GGATCAAGTGATCAAGTGAATTAGGGATCAAGTGATTCAAGGGGATCA AGTGATTCAATTAGA	69	3.9	69	2	33	10	27	28
Southern	CAATAGGAGATCAAGTGTTCAAGACAACGGGGATTAAGTGATCAAGTGA ATTAGGGGATCAAGTGATT	69	5	74	7	34	10	27	27
Southern	CAATAGGAGATCAAGTGATTCAAACAATAGGGGATTAAGTGATCAAGTGA AATGAGGGCATCAAGTGATT	75	4.7	75	7	34	10	27	27
Southern	AATTAGGGGAAAGTGATTCAAAGGGGATCAAGTGATAAGTACAATTAGGGG ATCAAGTGATTCAATTAGAGGAGCAAGTGATTATAC	96	2.7	91	10	34	9	29	26
Southern	GGATCAAGTGATCAACTACAATAGGGATCAAGTGATCAAGTGAATAGG GGATCAAGTGATTCAATTGGGGGATCAAGTGGTTCAATTAGG	96	3.7	96	8	34	10	27	28

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 January 2019 doi:10.20944/preprints201901.0232.v1

Peer-reviewed version available at *Genes* **2019**, *10*, 324; doi:10.3390/genes10050324

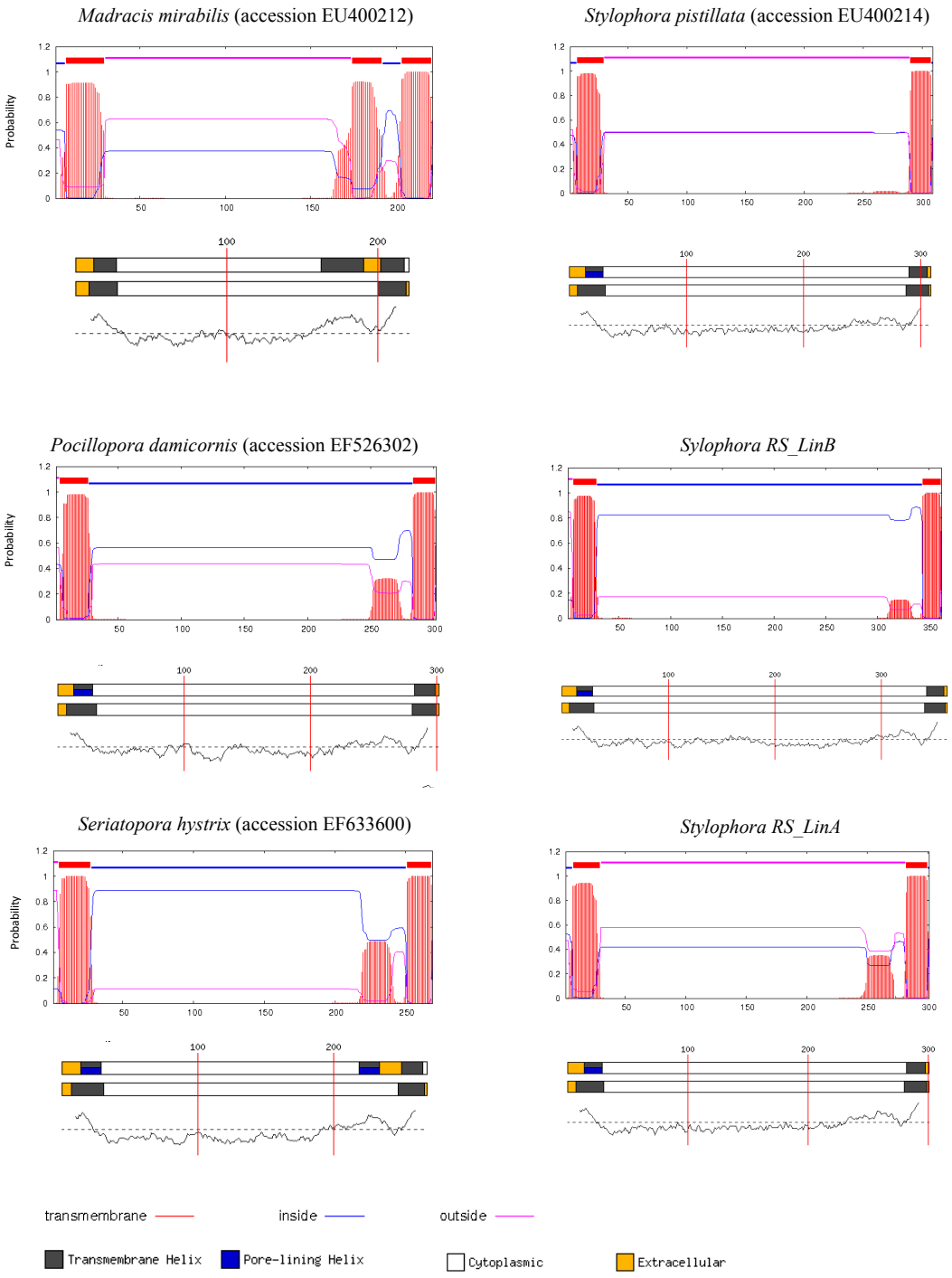


Figure 1. Top: predicted TM segments by the TMHHM approach. This plot shows the overall probability that a residue sits in a helix, inside of it, or outside (Sonnhammer et al., 1998; Krogh et al., 2001). Bottom: MEMSAT3 and MEMSAT-SVM predictions for the query sequence. The curve represent a raw output for the support vector machines. Dashed line shows the prediction threshold (PSIPRED server; Buchan et al., 2013; Nugent, T. & Jones, D.T. (2009).

Madracis mirabilis
(accession EU400212)

Pocillopora damicornis
(accession EF526302)

Seriatopora hystrix
(accession EF633600)

Stylophora pistillata
(accession EU400214)

Stylophora RS_LinB

Stylophora RS_LinA

Figure 2. Secondary structure predictions by PSIPRED(Jones 1999)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 January 2019 doi:10.20944/preprints201901.0232.v1

Peer-reviewed version available at *Genes* **2019**, *10*, 324; doi:10.3390/genes10050324

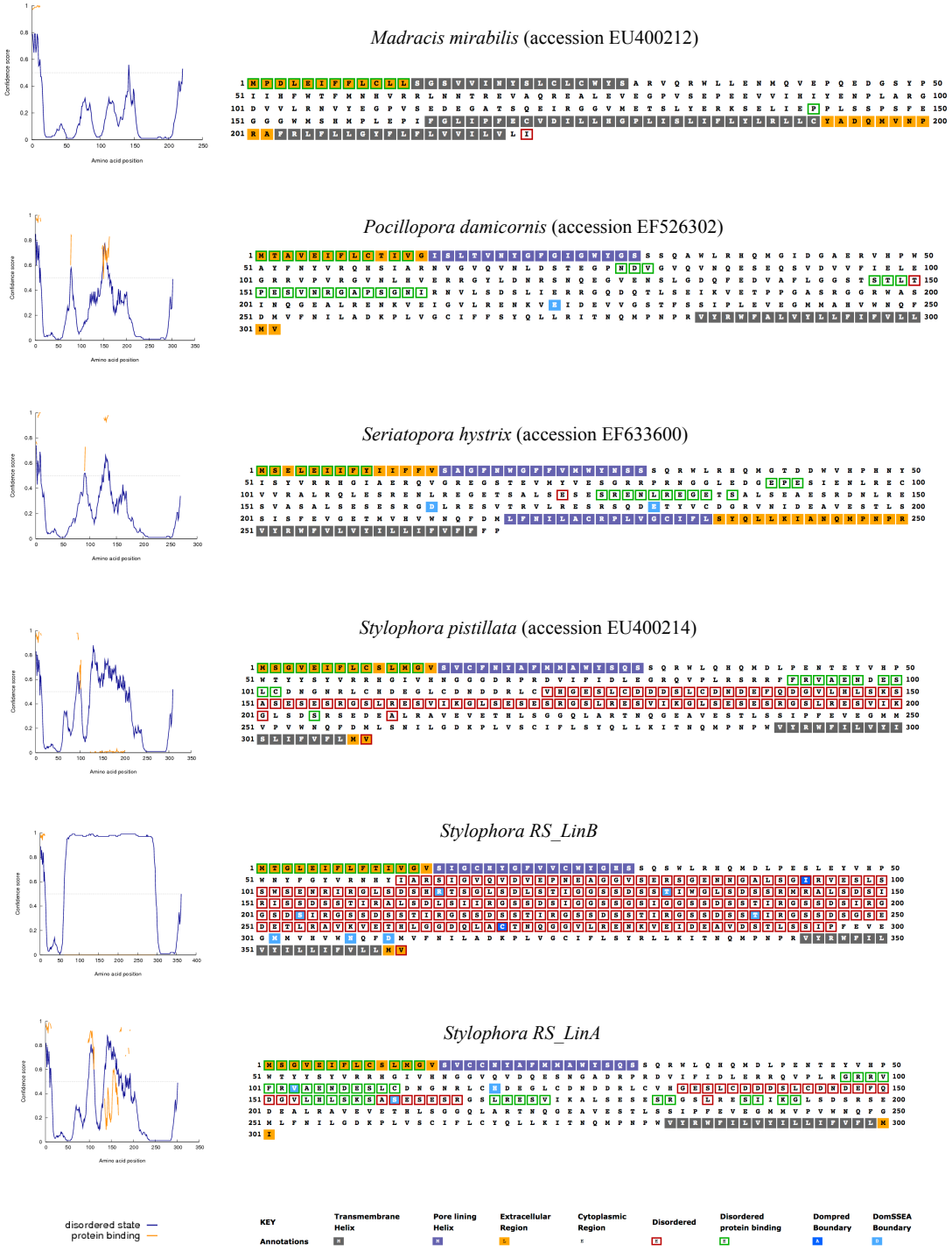


Figure 3. Left: Intrinsic disorder profile predicted by DISOPRED3. Disordered amino acids are indicated by a blue line above the grey dashed line (confidence score higher than 0.5). The confidence of disordered protein binding residues are denoted by an orange line (Buchan et al., 2013). Right: TM helix Map predicted by MEMSAT.

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

in *RS_LinB* the IDR coincides with a core region revealing a high number of duplicated polymorphic Tandem Repeats (TR) that were different to those found in *RS_LinA*, in length and nucleotide composition (Table 1).

Three protein domains were identified in *RS_LinB* with 4218 PSIBLAST hits that were not predicted in other members of this coral family, including other *Stylophora* lineages. Searches against the 95 million protein domains available in the CATH data base, using the pDomTHREADER approach, showed that the mt*ORF-pro* of *RS_LinB* had the best matches (levels of confidence: $0.0001 < P < 0.001$) with annotated domains involved in: (*i*) the structural integrity of a complex or its assembly within or outside a cell as per molecular function determined by Gene Ontology –GO = 0005198– (CATH domain = 1s58A00; p-value = 8×10^{-5}; from amino acid 1 to 299), (*ii*) host-bacteria/virus interactions (CATH domain = 1ocyA02, p-value = 5×10^{-5}; from amino acid 147-286), and (*iii*) domains that play a role in cell to cell, cell to matrix interactions and response to stress (CATH domain = 1ux6A01; p-value = 1×10^{-4}; from amino acid 136 to 292)

3.3. Predictions of molecular function, biological process, and cellular localization

Three molecular functions: catalytic activity ($0.53 < P < 0.77$), hydrolase activity acting on acid anhydrides ($0.56 < P < 0.92$), and cytoskeletal protein binding ($0.59 < P < 0.68$) were predicted for the mt*ORF-pro* in all pocilloporids. However, broad differences were found in their posterior probabilities (Table 2). For example, in the mt*ORF-pro* of *Stylophora RS_LinB*, catalytic function was predicted with low probabilities ($P = 0.53$) and the same was true for hydrolase activity acting on acid anhydrides in the mt*ORF-pro* of *Seriatopora* ($P = 0.56$).

Noticeably, in most genera the highest probabilities ($P > 0.7$) were given to different molecular function. In *Seriatopora*, ATP binding presented the highest probability ($P = 0.832$). In *Madracis*, the highest scores were found for ion channel activity ($P = 0.913$), hydrolase activity acting on acid anhydrides ($P = 0.902$) and ion transmembrane transporter activity ($P = 0.88$).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 January 2019 doi:10.20944/preprints201901.0232.v1

Peer-reviewed version available at *Genes* **2019**, *10*, 324; doi:10.3390/genes10050324

Table 2. Prediction of molecular function and biological processes by Psipred

Molecular Function		**Posterior Probabilities**					
GO term		***Madracis***	***Pocillopora***	***Seriatopora***	***Stylophora RS_LinB***	***Stylophora RS_LinA***	***Stylophora pistillata***
GO:0005216	ion channel activity	0.913	-	0.628	-	-	0.581
GO:0016817	**hydrolase activity, acting on acid anhydrides**	**0.902**	**0.862**	**0.561**	**0.824**	**0.695**	**0.626**
GO:0015075	ion transmembrane transporter activity	0.88	-	0.654	-	-	-
GO:0022890	inorganic cation transmembrane transporter activity	0.864	-	-	-	-	-
GO:0008324	cation transmembrane transporter activity	0.858	-	-	-	-	-
GO:0005524	ATP binding	0.833	-	0.832	-	0.739	0.502
GO:0046873	metal ion transmembrane transporter activity	0.787	-	-	-	-	-
GO:0015077	monovalent inorganic cation transmembrane transporter activity	0.749	-	-	-	-	-
GO:0003824	**catalytic activity**	**0.748**	**0.762**	**0.771**	**0.528**	**0.739**	**0.681**
GO:0016818	**hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides**	**0.707**	**0.602**	**0.602**	**0.666**	**0.695**	**0.619**
GO:0022857	transmembrane transporter activity	0.706	0.56	0.696	-	-	0.502
GO:0005261	cation channel activity	0.688	-	-	-	-	-
GO:0005215	transporter activity	0.67	0.548	0.684	-	0.561	0.515
GO:0035639	purine ribonucleoside triphosphate binding	0.639	0.538	0.706	-	0.733	0.538
GO:0008092	**cytoskeletal protein binding**	**0.588**	**0.592**	**0.598**	**0.682**	**0.62**	**0.676**
GO:0000166	nucleotide binding	-	0.52	0.71	-	0.72	-
GO:0001882	nucleoside binding	-	-	0.69	-	0.752	-
GO:0032549	ribonucleoside binding	-	-	0.682	-	0.779	-
GO:0017076	purine nucleotide binding	-	-	0.604	-	0.779	0.557
GO:0022891	substrate-specific transmembrane transporter activity	-	-	0.68	-	-	-
GO:0030554	adenyl nucleotide binding	-	-	0.648	-	0.547	-
GO:0016301	kinase activity	-	0.53	0.549	0.617	0.857	0.655
Biological Process							
GO:0006810	**transport**	**0.862**	**0.847**	**0.874**	**0.803**	**0.824**	**0.82**
GO:0034220	ion transmembrane transport	0.845	-	0.689	-	-	-
GO:0019222	**regulation of metabolic process**	**0.813**	**0.853**	**0.795**	**0.711**	**0.849**	**0.791**
GO:0007166	**cell surface receptor signaling pathway**	**0.804**	**0.63**	**0.564**	**0.79**	**0.662**	**0.717**
GO:0009117	nucleotide metabolic process	0.728	-	0.697	-	-	-
GO:0051649	**establishment of localization in cell**	**0.694**	**0.692**	**0.721**	**0.538**	**0.698**	**0.659**
GO:0051641	**cellular localization**	**0.647**	**0.631**	**0.64**	**0.632**	**0.643**	**0.636**

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

In *Pocillopora* and *RS_LinB,* the highest score was given to hydrolase activity acting on acid anhydrides (P = 0.862 and P = 0.814 respectively). In *RS_LinA*, kinase activity (P = 0.857), purine nucleotide binding (P = 0.779), and ribonucleoside binding (P = 0.779) had the highest scores. In *Stylophora pistillata* catalytic activity and cytoskeletal protein binding were predicted as the most probable functions, with moderated probabilities (both P = 0.68). Ion transmembrane transport activities were predicted for *Madracis* and *Seriatopora*, but not for other genera. Moreover, kinase activity was suggested for the mt*ORF-pro* of all genera, except *Madracis*, while ribonucleoside binding was predicted only for *RS_LinA* and *Seriatopora* (Table 2).

The mt*ORF-pro* for these genera differed also in their aliphatic indices, a measure of the thermostability of the protein, where a higher value indicates a higher thermostability [74]: 112.4 for Madracis, 94.4 for *Pocillopora*, 88.3 for *RS_LinA*, 87.2 for *RS_LinB*, 85. 4 for *Stylophora pistillata* and 84.0 for *Seriatopora.*

Two main biological processes were suggested for the role of mt*ORF-pro* in all genera and *Stylophora* lineages (Table 2): transport ($0.80 < P < 0.87$) and regulation of metabolic processes ($0.71 < P < 0.85$), except in *RS_LinB* where cell surface receptor signaling pathway had the highest posterior probability after transport (P = 0.79). Other processes were predicted with moderated probabilities and two were only suggested for the mt*ORF-pro* of *Madracis* and *Seriatopora* (i.e. ion transmembrane transport and nucleotide metabolic process; Table 2). Furthermore, predictions for cellular localization yielded a high probability and reliability for this protein being an intrinsic and integral component of the membrane ($0.97 < P < 1.00$) in all genera, with moderated to high probabilities of being part of an organelle membrane (i.e. mitochondria; $0.70 < P < 0.80$).

3.4. Signatures of selection in the *mtORF-pro* of *Stylophora* corals

Amino acid sites under positive/diversifying selection were identified with the MEME approach for Clade 1 sequences (i.e. sequences from *RS_LinA* plus specimens from Madagascar, Indo-Pacific and Indian Ocean; at 2 sites) and when sequences from *RS_LinA* and *RS_LinB* were

evaluated together (at 5 sites). However, positive selection was not detected when the alignment included all *Stylophora* sequences (Clade 1 and Clade 2; evaluating only unambiguously aligned regions) or when sequences from each *Stylophora* lineage were analysed separately. In contrast, the M8 and MEC model recovered the highest number of single amino acid sites under positive selection in most cases. However, only the MEC model showed significant values for most comparisons, this is: (*i*) in the mt*ORF-pro* of *Stylophora* specimens within Clade 1 –at 1 site– (AIC score for MEC: 2846.897 was lower than the M8a score: 2852.35), (*ii*) when the alignment included sequences from both Red Sea lineages (*RS_LinA* and *RS_LinB*; AIC score for MEC: 1468.55; AIC score for M8a: 1472.37), and (*iii)* when sequences from all *Stylophora* specimens were analysed (i.e. the MEC score = 1737.074; M8a score = 1745.86).

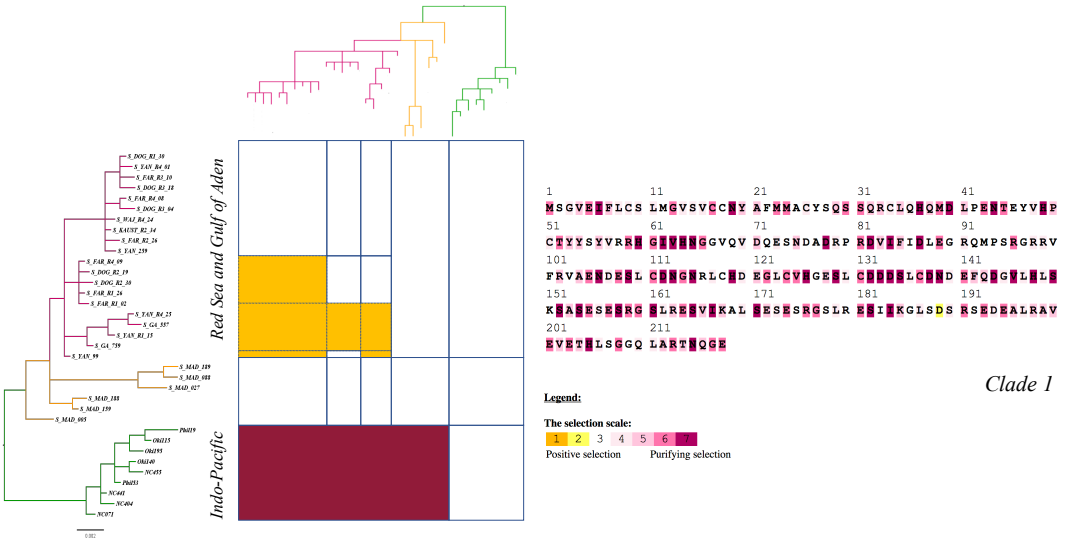
In addition, the analyses of sequences from the southern and northern group of *RS_LinB*, separately, indicated multiple sites under positive selection in the mt*ORF-pro* of the southern group (AIC score for MEC: 2313.488, AIC score for M8a: 2317.705), which was corroborated by the analyses of the Nei-Gojobori method with high-probabilities ($0.003 < P < 0.042$). Positive selected sites were not found in the northern group of this lineage nor in sequences from *RS_LinA* alone (except in the mt*ORF-pro* of specimens which haplotypes have the highest frequencies in the southern Red Sea). A graphical view of positively and negatively selected sites using the MEC approach, as well as a graphical summary for the outputs of the Nei-Gojobori method are included in Figure 4 (Full matrices are available from the authors).

Amino acid sites under pervasive negative/purifying selection were found in all alignments of the mt*ORF-pro* using the FEL method: at 11 sites for Clade 1 plus Clade 2 sequences; at 7 sites for Clade 1 sequences, at 6 sites for Clade 2 sequences (*RS_LinB*) and also at 6 sites when *RS_LinA* and *RS_LinB* sequences were placed together. One site was found under purifying selection in the mt*ORF-pro* of the northern group of *RS_LinB*, as well as in the southern group. Furthermore, purifying selection was also detected in the mt*ORF-pro* of Indo-Pacific specimens by the Nei-Gojobori method ($0.003 < P < 0.042$). aBSREL did not detect sites under selection along the branches of the phylogenetic tree of *Stylophora*.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 January 2019 doi:10.20944/preprints201901.0232.v1

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

Stylophora CLADE 1 (*RS_LinA* plus *Stylophora pistillata)*



Stylophora CLADE 2 (*RS_LinB*)

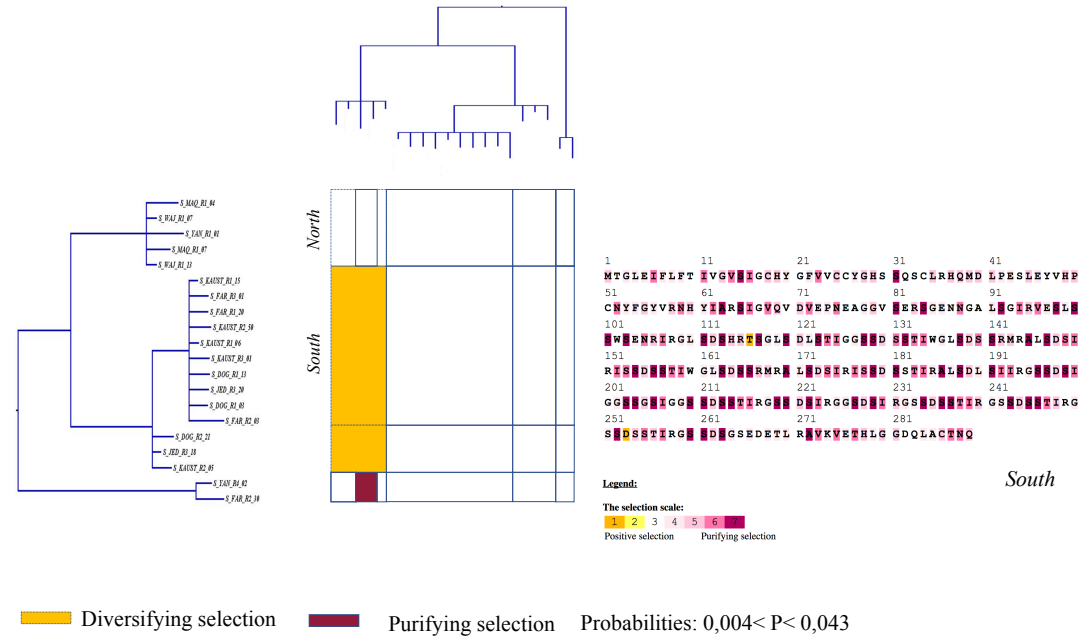


Figure 4. Left: a graphical view of the Codon-based test of purifying and positive/diversifying selection (Nei and Gojobori 1986) between the sequences belonging to haplotypes in each *Stylophora* Clade. Right: Sites under positive and purifying selection predicted by the MEC approach (Doron-Faigenboim and Pupko, 2006); for clade 1 (above) and for the southern group of Clade 2 (below).

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

4. Discussion

4.1 mt*ORF* protein structure in the coral family Pocilloporidae

This study confirms that the mtORF gene encodes a transmembrane protein in pocilloporid corals, which is in agreement with its preliminary characterization by Flot and Tillier (2007) [2]. No stop codons were found in the translated sequences and all protein prediction approaches applied (TMHMM, MEMSAT and MEMSAT-SVM) predicted two or three different TMHs in all four investigated pocilloporid genera, with high posterior probabilities ($P > 0.98$ in the TMHMM approach). This finding was reinforced by Gene Ontology (GO terms) analysis predicting that this protein is an integral ($P > 0.97$) and intrinsic component of the membrane ($P > 0.99$), likely of the mitochondria ($P > 0.7$).

The secondary structure of the mt*ORF-pro* differed among genera and particularly among *Stylophora* lineages (Figure 2) and the prediction of IDRs varied largely among pocilloporids. For example, IDRs were absent in the proteins of *Madracis*, *Pocillopora* and *Seriatopora* (only one amino acid residue was predicted as disordered), but a long stretch of IDRs were predicted in the protein of *Stylophora RS_LinB* (234 disordered amino acid residues out of 362). This contrasted with the number of IDRs predicted for *Stylophora pistillata* (79 residues out of 309) and *RS_LinA* (27 residues out of 301). Furthermore, in *RS_LinB*, IDRs coincided with a region rich in duplicated TR (Table 1) that generated long insertion-deletions (indels) among sequences.

Long IDRs, as the predicted here for *RS_LinB*, are recognized to facilitate macromolecular interactions, participating actively in the assembly of signaling complexes by interacting with structured domains in other proteins [54,75–77]. They are also essential for increasing the involvement of a range of proteins in the diversity of cellular processes required during adaptation, particularly to variable environmental conditions [78]. In line with this, one of the most outstanding findings in this study was the presence of protein domains ($0.0001 < P < 0.001$) that were exclusively found in the mt*ORF-pro* of *RS_LinB* (likely the oldest *Stylophora* within the Red Sea [9]) and were predicted to be involved in the assembly of cellular complexes (CATH

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

389 domain code: 1s58A00), cell-to-cell and cell-to-matrix interactions, and response to stress
390 (CATH domain code: 1ux6A0). This is highly congruent with the occurrence of IDRs in the
391 protein of this lineage, suggesting a key role of the *mtORF-pro* in mitochondrial complexes,
392 which are key in the regulation of multiple processes related with thermal adaptation [22].
393
394 Furthermore, in most metazoans IDRs are associated with regulation of cellular processes,
395 including intracellular signaling, membrane fusion and transport, and signal transduction [54,77],
396 which were the main processes predicted for biological function of the *mtORF-pro* of *RS_LinB*
397 (i.e. transport and cell surface receptor signaling pathway [$P = 0.8$]).
398
399 Our findings, particularly those associated to the mt*ORF-pro* of *Stylophora RS_LinB*, therefore,
400 support the hypothesis that differences in protein structure among genera and *Stylophora* lineages
401 are likely the results of dissimilarities in the environmental pressures that each group faced
402 during its evolution. Although experimental studies would be necessary to understand the role of
403 IDRs and TR in the genome of *RS_LinB*, previous studies in other eukaryotes have suggested that
404 IDRs and indels not only play a role in genome evolution but are also key in the regulation of
405 gene expression during the exposition of organisms to different stressors [79,80]. Moreover, TRs
406 are a common characteristic of the genome of organisms evolving under strong selective
407 pressures, environmental stress, drastic environmental changes, global warming and/or that are
408 exposed to new environments [81–83]. In humans, for example, mutations that affect TR number
409 have been found to increase the fitness of the cell when exposed to stressful conditions (e.g.,
410 cold, heat, hypoxia, oxidative stress) by adjusting the regulatory network which enhances protein
411 activity and gene expression [84]. All these previous findings support our assumption of a key
412 role of the mt*ORF-pro* and of the mitogenome in adaptation to changing environments, as
413 suggested by Banguera-Hinestroza et al. (2018) [9].
414
415 **4.2. Differences in mt*ORF-pro* function among pocilloporids**

416 The molecular function of the mt*ORF-pro* was predicted to differ among genera and among
417 *Stylophora* lineages. Most functions related with ion and cation transmembrane transporter

Peer-reviewed version available at Genes 2019, 10, 324; doi:10.3390/genes10050324

418 activities (0.75 < P < 0.91; Table 2), were only predicted for the mt*ORF-pro* of *Madracis*.
419 Moreover, for this genus, as well as for *Pocillopora* and *RS_LinB*, hydrolase activity acting on
420 acid anhydrides, catalyzing the transmembrane movement of substances, was predicted with the
421 highest probability (for all P > 0.82; only second after ion channel activity in *Madracis*). This
422 function, however, showed low probabilities for *Seriatopora*, *RS_LinA* and *Stylophora pistillata*
423 (0. 56 < P < 0.69). In fact, the analyses of the mt*ORF-pro* of these genera yielded the highest
424 probabilities for different molecular functions: ATP binding (P = 0.83) for *Seriatopora*; kinase
425 activity (P = 0.86) for *Stylophora RS_LinA*; and catalytic and cytoskeletal protein binding
426 activities (both P = 0.68) for *Stylophora pistillata*.
427
428 The dissimilarities in the predicted protein function of mt*ORF*-pro and the differences in
429 posterior probabilities when several functions were predicted, are in line with differences found
430 in protein structure among genera. This may be the result of mutational and structural changes
431 occurring in this protein during the divergence of pocilloporids, also likely linked to different
432 environmental settings (i.e. in *Stylophor*a lineages). Although this hypothesis is highly
433 speculative and experimental analyses are required to confirm these differences, it is well known
434 that proteins may perform different functions depending on their biological context, such as
435 cellular location, the substrate to which they bind and their interactions with other proteins to
436 form molecular complexes [85]. In this respect, it is broadly accepted that proteins involved in
437 mitochondrial complexes, which is likely the case of this mt*ORF-pro*, are prone to structural
438 changes, which not only impact protein function, but also enhance the protein flexibility that is
439 required for adaptation to different conditions (see review by Fields, 2001 [86]).
440
441 **4.3. Signatures of selection in the mt*ORF-pro* of *Stylophora* lineages**
442
443 Our data show that the mt*ORF-pro* of *Stylophora* lineages carries signatures of both
444 positive/diversifying and negative/purifying selection a multiple amino acid sites. Interestingly,
445 the MEC approach detected signal of positive selection only in the protein of the southern group
446 of *RS_LinB* (specimens inhabiting the warmest regions of the Red Sea), a signature that was
447 confirmed by the Nei-Gojobori method (Figure 3). In contrast, no positive selected sites were

448 detected in the mt*ORF-pro* of the northern group (i.e. restricted to the coldest areas). In this
449 group, as well as in sequences from *RS_LinA* specimens, multiple sites were found under
450 purifying selection (except in protein sequences of those haplotypes with the highest frequency in
451 the southern Red Sea, Figure 3). The findings of positive selection in the mt*ORF-pro* of
452 specimens inhabiting the hottest areas of the Red Sea are congruent with the results based in
453 protein structure and function, supporting the hypothesis of a role of this protein in thermal
454 adaptation to high temperatures, highlighting also its functional importance.

455 As suggested by previous phylogenetic studies [9], *RS_LinB* represent an ancestral lineage,
456 endemic of the Red Sea, with an ancient history within this region. Although the first occurrence
457 of this lineage in the Red Sea is unknown, it is broadly accepted that coral reef ecosystems
458 accompanied the evolution of this basin since its incipient stage in the early Miocene and have
459 been affected by the multiple and strong climatic variations occurring in the region throughout
460 time [87]. Pocilloporid corals, as well as those from other coral families, inhabiting the Red Sea
461 are likely descendants of those that entered the region during the Pliocene-Pleistocene epochs
462 [87–91], which overcame the strong climatic fluctuations in salinity and temperature of the
463 glacial e interglacial periods [92–94] by surviving in refugia in the northern and southern Red Sea
464 (Reviewed in DiBattista et al., 2016 [95]).

465 The above explain the signals of adaptive evolution found in the mt*ORF-pro* of *RS_LinB*, for
466 which the extreme fluctuation in temperatures in the southern Red Sea may have imposed strong
467 selective pressures. Moreover, the lack of positive selection in the protein of *RS_LinA* that
468 contrasted with the multiple amino acid sites found under purifying selection, may be related
469 with its recent history in the Red Sea, in agreement with its hybrid origin. In fact, more recent
470 *Stylophora* lineages likely entered the Red Sea within the last 4-7k years, when environmental
471 conditions were similar as today [96], passing the bottleneck of high temperature waters in the
472 southern Red Sea (up to 34°C) and then spreading north along with a range of other coral species
473 [97]. Purifying selection, therefore, likely acted as the main force against deleterious mutations in
474 the mt*ORF-pro* of this lineage, shaping the mitochondrial diversity needed for its survival in such
475 range of environmental conditions.

476 **Conclusions**

477

478 Our study not only offers insights into the coding nature of the mtORF gene in pocilloporid
479 corals, but add to our knowledge of the main characteristics of this protein. Striking differences
480 were found among the mt*ORF* proteins of pocilloporid corals, with the most noticeable being
481 those among *Stylophora* lineages, which include differences in predicted structure and functions.
482 The most differentiated protein was that of *RS_LinB*, a lineage that was hypothesized to have
483 conquered the extreme environments of the Red Sea during an early colonization, which date is
484 still unknown, but likely predating other *Stylophora* lineages. All the characteristics found in the
485 mt*ORF-pro* of the *RS_LinB* (i.e. high frequency of TR, long stretches of disordered residues and
486 the presence CATH domains related to stress response and protein complexes), plus the signals
487 of selection, support our hypothesis of the role of this protein, and therefore of the mitogenome,
488 in the adaptation of *Stylophora* corals to extreme environments and fluctuating conditions.
489
490 This study opens the door for further studies looking at the role of mitochondria and
491 mitochondrial genes in coral species adaptation and diversification. Corals and coral reefs are
492 severely threatened by climate change (e.g. sea surface temperature rise), hence, more detailed
493 studies on coral adaptive mechanisms are required to understand the potential responses of corals
494 in both ecological and evolutionary scales. We consider the understanding of protein-protein
495 interaction at the mitochondrial level of particular relevance, since they appear to play a key role
496 in coral adaption to strong environmental changes.

497 **Funding statement**:

498 EBH's postdoctoral position in Belgium was funded by an *Action de Recherche Concertée*
499 (ARC) grant of the Fédération Wallonie-Bruxelles to JFF. The study was further supported by the
500 King Abdullah University of Science and Technology (KAUST), Saudi Arabia and the bi-lateral
501 project "The Jeddah Transect" of the King Abdulaziz University (KAU), Saudi Arabia, and the
502 Helmholtz Center for Ocean Research (GEOMAR), Germany (YS).

Acknowledgments: We thank Christian Voolstra and the Bioscience Core Lab at KAUST for sharing their facilities, and Abdulmoshin Al-Sofyani at KAU for supporting coral sampling. Thanks also to Sandra Cervantes Arango, Dario Ojeda Alayon and Patrick Mardulyn, for useful discussions and advices.

Author Contributions: Conceptualization, Eulalia Banguera-Hinestroza, Yvonne Sawall and Jean-François Flot; Formal analysis, Eulalia Banguera-Hinestroza; Funding acquisition, Jean-François Flot; Investigation, Eulalia Banguera-Hinestroza, Yvonne Sawall and Jean-François Flot; Methodology, Eulalia Banguera-Hinestroza; Writing – original draft, Eulalia Banguera-Hinestroza; Writing – review & editing, Eulalia Banguera-Hinestroza, Yvonne Sawall and Jean-François Flot.

REFERENCES

1. Higashi, A.; Nagai, S.; Salomon, P.S.; Ueki, S. A unique, highly variable mitochondrial gene with coding capacity of *Heterosigma akashiwo*, class Raphidophyceae. *J. Appl. Phycol.* **2017**, *29*, 2961–2969.
2. Flot, J.F.; Tillier, S. The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: The putative D-loop and a novel ORF of unknown function. *Gene* **2007**, *401*, 80–87.
3. Veron J.E.N.; Pichon, M. Scleractinia of eastern Australia. Part I : Families Thamnasteriidae, Astrocoeniidae, Pocilloporidae. *Aust. Gov. Publ. Serv.* **1976**, 208.
4. Chen, C.; Chiou, C.Y.; Dai, C.F.; Chen, C.A. Unique mitogenomic features in the scleractinian family pocilloporidae (Scleractinia: Astrocoeniina). *Mar. Biotechnol.* **2008**. *10*, 538-553
5. Flot, J.F.; Magalon, H.; Cruaud, C.; Couloux, A.; Tillier, S. Patterns of genetic structure among Hawaiian corals of the genus *Pocillopora* yield clusters of individuals that are compatible with morphology. *C. R. Biol.* **2008**, *331*, 239–247.
6. Johnston, E.C.; Forsman, Z.H.; Flot, J.-F.; Schmidt-Roach, S.; Pinzón, J.H.; Knapp, I.S.S.; Toonen, R.J. A genomic glance through the fog of plasticity and diversification in

Pocillopora. *Sci. Rep.* **2017**, *7*, 5991.

7. Flot, J.F.; Licuanan, W.Y.; Nakano, Y.; Payri, C.; Cruaud, C.; Tillier, S. Mitochondrial sequences of Seriatopora corals show little agreement with morphology and reveal the duplication of a tRNA gene near the control region. *Coral Reefs* **2008**, *27*, 789–794.
8. Warner, P.A.; Van Oppen, M.J.H.; Willis, B.L. Unexpected cryptic species diversity in the widespread coral *Seriatopora hystrix* masks spatial-genetic patterns of connectivity. *Mol. Ecol.* **2015**, *24*, 2993–3008.
9. Banguera-Hinestroza, E.; Sawall, Y.; Al-Sofyani, A.; Mardulyn, P.; Fuertes-Aguilar, J.; Cárdenas-Henao, H.; Jimenez-Infante, F.; Voolstra, R.C.; Flot, J.F. mtDNA recombination indicative of hybridization suggests a role of the mitogenome in the adaptation of reef-building corals to extreme environments. *BioRxiv 462069; doi https//doi.org/10.1101/462069* **2018**.
10. Flot, J.F.; Blanchot, J.; Charpy, L.; Cruaud, C.; Licuanan, W.Y.; Nakano, Y.; Payri, C.; Tillier, S. Incongruence between morphotypes and genetically delimited species in the coral genus *Stylophora*: phenotypic plasticity, morphological convergence, morphological stasis or interspecific hybridization? *BMC Ecol.* **2011**, *11*, 22.
11. Shearer, T.L.; Oppen, M.J.H. Van; Romano, S.L.; Wörheide, G. Slow mitochondria DNA sequence evolution in the Anthozoa. *Mol. Ecol.* **2002**, *11*, 2475–2487.
12. Shearer, T.L.; Coffroth, M.A. DNA BARCODING: Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Mol. Ecol. Resour.* **2008**, *8*, 247–255.
13. Hill, G.E. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. *Ecol. Evol.* **2016**, *6*, 5831–5842.
14. Cheviron, Z.A.; Brumfield, R.T. Genomic insights into adaptation to high-altitude environments. *Heredity.* **2012**, *108*, 354–361.
15. Morales, H.E.; Pavlova, A.; Joseph, L.; Sunnucks, P. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Mol. Ecol.* **2015**, *24*, 2820–37.
16. Scott, G.R.; Schulte, P.M.; Egginton, S.; Scott, A.L.M.; Richards, J.G.; Milsom, W.K. Molecular Evolution of Cytochrome c Oxidase Underlies High-Altitude Adaptation in the Bar-Headed Goose. *Mol. Biol. Evol.* **2011**, *28*, 351–363.

17. Silva, G.; Lima, F.P.; Martel, P.; Castilho, R. Thermal adaptation and clinal mitochondrial DNA variation of European anchovy. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20141093–20141093.
18. Saccone, C.; Lanave, C.; De Grassi, A. Metazoan OXPHOS gene families: evolutionary forces at the level of mitochondrial and nuclear genomes. *Biochim. Biophys. Acta* **2006**, *1757*, 1171–8.
19. Saraste, M. Oxidative Phosphorylation at the fin de sicle. **1999**, *283*, 1488–1493.
20. Ben Slimen, H.; Schaschl, H.; Knauer, F.; Suchentrunk, F. Selection on the mitochondrial ATP synthase 6 and the NADH dehydrogenase 2 genes in hares (*Lepus capensis* L., 1758) from a steep ecological gradient in North Africa. *BMC Evol. Biol.* **2017**, *17*, 46.
21. Gershoni, M.; Templeton, A.R.; Mishmar, D. Mitochondrial bioenergetics as a major motive force of speciation. *BioEssays* **2009**, *31*, 642–650.
22. Lajbner, Z.; Pnini, R.; Camus, M.F.; Miller, J.; Dowling, D.K. Experimental evidence that thermal selection shapes mitochondrial genome evolution. *Sci. Rep.* **2018**, *8*, 9500.
23. Sunnucks, P.; Morales, H.E.; Lamb, A.M.; Pavlova, A.; Greening, C. Integrative approaches for studying mitochondrial and nuclear genome co-evolution in oxidative phosphorylation. *Front. Genet.* **2017**, *8*, 1–12.
24. Baris, T.Z.; Wagner, D.N.; Dayan, D.I.; Du, X.; Blier, P.U.; Pichaud, N.; Oleksiak, M.F.; Crawford, D.L. Evolved genetic and phenotypic differences due to mitochondrial-nuclear interactions. *PLOS Genet.* **2017**, *13*, e1006517.
25. Finch, T.M.; Zhao, N.; Korkin, D.; Frederick, K.H.; Eggert, L.S. Evidence of Positive Selection in Mitochondrial Complexes I and V of the African Elephant. *PLoS One* **2014**, *9*, e92587.
26. Rand, D.M.; Haney, R.A.; Fry, A.J. Cytonuclear coevolution: the genomics of cooperation. *Trends Ecol. Evol.* **2004**, *19*, 645–653.
27. Hill, G.E. Mitonuclear ecology. *Mol. Biol. Evol.* **2015**, *32*, 1917–1927.
28. Pavlova, A.; Amos, J.N.; Joseph, L.; Loynes, K.; Austin, J.J.; Keogh, J.S.; Stone, G.N.; Nicholls, J.A.; Sunnucks, P. Perched at the mito-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an australian bird. *Evolution.* **2013**, *67*, 3412–3428.

29. Hill, G.E. The mitonuclear compatibility species concept. *Auk* **2017**, *134*, 393–409.
30. Osman, E.O.; Smith, D.J.; Ziegler, M.; Kürten, B.; Conrad, C.; El-Haddad, K.M.; Voolstra, C.R.; Suggett, D.J. Thermal refugia against coral bleaching throughout the northern Red Sea. *Glob. Chang. Biol.* **2018**.
31. Mayer, M.P.; Bukau, B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* **2005**, *62*, 670–684.
32. Arrigoni, R.; Benzoni, F.; Terraneo, T.I.; Caragnano, A.; Berumen, M.L. Recent origin and semi-permeable species boundaries in the scleractinian coral genus *Stylophora* from the Red Sea. *Sci. Rep.* **2016**, *6*, 34612.
33. Stefani, F.; Benzoni, F.; Yang, S.Y.; Pichon, M.; Galli, P.; Chen, C.A. Comparison of morphological and genetic analyses reveals cryptic divergence and morphological plasticity in *Stylophora* (Cnidaria, Scleractinia). *Coral Reefs.* **2011**, *30*, 1033–1049.
34. Tutar, Y. Pseudogenes. *Comp. Funct. Genomics* **2012**, *2012*, 1–4.
35. Xiao, J.; Sekhwal, M.K.; Li, P.; Ragupathy, R.; Cloutier, S.; Wang, X.; You, F.M. Pseudogenes and their genome -wide prediction in plants. Int. J. Mol. Sci. 2016, 17, 1991.
36. Pohl, M.; Theiβen, G.; Schuster, S. GC content dependency of open reading frame prediction via stop codon frequencies. *Gene* **2012**, *511*, 441–446.
37. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L.. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *J. Mol. Biol.* **2001**, *305*, 567–580.
38. Sonnhammer, E.L.; von Heijne, G.; Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings. Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 175–82.
39. McGuffin, L.J.; Bryson, K.; Jones, D.T. The PSIPRED protein structure prediction server. *Bioinformatics* **2000**, *16*, 404–405.
40. Nugent, T.; Jones, D.T. Transmembrane protein topology prediction using support vector machines. *BMC Bioinformatics* **2009**, *10*, 159.
41. Kürten, B.; Al-Aidaroos, A.M.; Struck, U.; Khomayis, H.S.; Gharbawi, W.Y.; Sommer, U. Influence of environmental gradients on C and N stable isotope ratios in coral reef biota of the Red Sea, Saudi Arabia. *J. Sea Res.* **2014**. *85*, 379–394

623 42. Sawall, Y.; Al-Sofyani, A.; Hohn, S.; Banguera-Hinestroza, E.; Voolstra, C.R.; Wahl, M.
624 Extensive phenotypic plasticity of a Red Sea coral over a strong latitudinal temperature
625 gradient suggests limited acclimatization potential to warming. *Sci. Rep.* **2015**, *5*, 8940.
626 43. Sinniger, F.; Prasetia, R.; Yorifuji, M.; Bongaerts, P.; Harii, S. *Seriatopora* diversity
627 preserved in upper mesophotic coral ecosystems in southern japan. *Front. Mar. Sci.* **2017**,
628 *4*, 1–9.
629 44. Van Oppen, M.J.H.; Bongaerts, P.; Underwood, J.N.; Peplow, L.M.; Cooper, T.F. The role
630 of deep reefs in shallow reef recovery: An assessment of vertical connectivity in a
631 brooding coral from west and east Australia. *Mol. Ecol.* **2011**, *20*, 1647–1660.
632 45. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis
633 Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874.
634 46. Robitzch, V.; Banguera-Hinestroza, E.; Sawall, Y.; Al-Sofyani, A.; Voolstra, C.R.
635 Absence of genetic differentiation in the coral *Pocillopora verrucosa* along environmental
636 gradients of the Saudi Arabian Red Sea. *Front. Mar. Sci.* **2015**, *2*.
637 47. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids
638 Res. **1999**, *27*, 573–80.
639 48. Buchan, D.W.A.; Ward, S.M.; Lobley, A.E.; Nugent, T.C.O.; Bryson, K.; Jones, D.T.
640 Protein annotation and modelling servers at University College London. *Nucleic Acids*
641 *Res.* **2010**, *38*, W563–W568.
642 49. Jones, D.T. Protein secondary structure prediction based on position-specific scoring
643 matrices 1 1Edited by G. Von Heijne. *J. Mol. Biol.* **1999**, *292*, 195–202.
644 50. Altschul, S. Gapped BLAST and PSI-BLAST: a new generation of protein database search
645 programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
646 51. Jones, D.T. Improving the accuracy of transmembrane protein topology prediction using
647 evolutionary information. *Bioinformatics* **2007**, *23*, 538–544.
648 52. Jones, D.T.; Taylor, W.R.; Thornton, J.M. A model recognition approach to the prediction
649 of all-helical membrane protein structure and topology. *Biochemistry* **1994**, *33*, 3038–
650 3049.
651 53. Jones, D.T.; Ward, J.J. Prediction of disordered regions in proteins from position specific
652 score matrices. *Proteins Struct. Funct. Genet.* **2003**, *53*, 573–578.

653 54. Van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R.J.; Daughdrill, G.W.; Dunker, A.K.;
654 Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D.T.; et al. Classification of Intrinsically
655 Disordered Regions and Proteins. *Chem. Rev.* **2014**, *114*, 6589–6631.
656 55. Ward, J.J.; McGuffin, L.J.; Bryson, K.; Buxton, B.F.; Jones, D.T. The DISOPRED server
657 for the prediction of protein disorder. *Bioinformatics* **2004**, *20*, 2138–2139.
658 56. Minneci, F.; Piovesan, D.; Cozzetto, D.; Jones, D.T. FFPred 2.0: Improved homology-
659 independent prediction of gene ontology terms for eukaryotic protein sequences. *PLoS*
660 *One* **2013**, *8*, e63754.
661 57. Buchan, D.W.A.; Minneci, F.; Nugent, T.C.O.; Bryson, K.; Jones, D.T. Scalable web
662 services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **2013**, *41*,
663 W349–W357.
664 58. Lobley, A.; Sadowski, M.I.; Jones, D.T. pGenTHREADER and pDomTHREADER: new
665 methods for improved protein fold recognition and superfamily discrimination.
666 *Bioinformatics* **2009**, *25*, 1761–1767.
667 59. Nielsen, H. Predicting Secretory Proteins with SignalP. In *Protein Function Prediction:*
668 *Methods and Protocols*; Kihara, D., Ed.; Springer New York. 2017; pp. 59–73.
669 60. Petersen, T.N.; Brunak, S.; Von Heijne, G.; Nielsen, H. SignalP 4.0: Discriminating signal
670 peptides from transmembrane regions. *Nat. Methods* 2011, *8*, 785–786.
671 61. Delport, W.; Poon, A.F.Y.Y.; Frost, S.D.W.W.; Kosakovsky Pond, S.L. Datamonkey
672 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*
673 **2010**, *26*, 2455–2457.
674 62. Kosakovsky Pond, S.L.; Frost, S.D.W. Datamonkey: Rapid detection of selective pressure
675 on individual sites of codon alignments. *Bioinformatics* **2005**, *21*, 2531–2533.
676 63. Kosakovsky Pond, S.L.; Pond, S.L.K.; Frost, S.D.W.; Kosakovsky Pond, S.L.; Frost,
677 S.D.W.; Pond, S.L.K.; Frost, S.D.W. Not so different after all: A comparison of methods
678 for detecting amino acid sites under selection. *Mol. Biol. Evol.* **2005**, *22*, 1208–1222.
679 64. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Kosakovsky Pond,
680 S.L. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.*
681 **2012**, *8*, e1002764.
682 65. Kosakovsky Pond, S.L.; Murrell, B.; Fourment, M.; Frost, S.D.W.; Delport, W.; Scheffler,

683 K. A Random Effects Branch-Site model for detecting episodic diversifying selection.
684 *Mol. Biol. Evol.* **2011**, *28*, 3033–3043.
685 66. Smith, M.D.; Wertheim, J.O.; Weaver, S.; Murrell, B.; Scheffler, K.; Kosakovsky Pond,
686 S.L. Less is more: An adaptive branch-site random effects model for efficient detection of
687 episodic diversifying selection. *Mol. Biol. Evol.* **2015**, *32*, 1342–1353.
688 67. Yang, Z.; Bielawski, J.P. Statistical methods for detecting molecular adaptation. *Trends*
689 *Ecol. Evol.* **2000**, *15*, 496–503.
690 68. Doron-Faigenboim, A.; Pupko, T. A Combined Empirical and Mechanistic Codon Model.
691 *Mol. Biol. Evol.* **2006**, *24*, 388–397.
692 69. Doron-Faigenboim, A.; Stern, A.; Mayrose, I.; Bacharach, E.; Pupko, T. Selection: A
693 server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics* **2005**,
694 *21*, 2101–2103.
695 70. Stern, A.; Doron-Faigenboim, A.; Erez, E.; Martz, E.; Bacharach, E.; Pupko, T. Selecton
696 2007: advanced models for detecting positive and purifying selection using a Bayesian
697 inference approach. *Nucleic Acids Res.* **2007**, *35*, W506–W511.
698 71. Yang, Z. Bayes Empirical Bayes Inference of Amino Acid Sites Under Positive Selection.
699 *Mol. Biol. Evol.* **2005**, *22*, 1107–1118.
700 72. Nei, M.; Gojobori, T. Simple methods for estimating the numbers of synonymous and
701 nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **1986**, *3*, 418–426.
702 73. Hughes, A.L.; Friedman, R. Codon-based tests of positive selection, branch lengths, and
703 the evolution of mammalian immune system genes. *Immunogenetics* **2008**, *60*, 495–506.
704 74. Merkler, D.J.; Farrington, G.K.; Wedler, F.C. Protein thermostability. Correlations
705 between calculated macroscopic parameters and growth temperature for closely related
706 thermophilic and mesophilic bacilli. Int. J. Pept. Protein Res. **1981**, *18*, 430–42.
707 75. Oldfield, C.J.; Dunker, A.K. Intrinsically Disordered Proteins and Intrinsically Disordered
708 Protein Regions. *Annu. Rev. Biochem.* **2014**, *83*, 553–584.
709 76. Shammas, S.L. Mechanistic roles of protein disorder within transcription. *Curr. Opin.*
710 *Struct. Biol.* **2017**, *42*, 155–161.
711 77. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and
712 regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29.

78. Sharma, M.; Pandey, G.K. Expansion and function of repeat domain proteins during stress and development in plants. *Front. Plant Sci.* **2016**, *6*, 1–15.
79. Belinky, F.; Cohen, O.; Huchon, D. Large-scale parsimony analysis of metazoan indels in protein-coding genes. *Mol. Biol. Evol.* **2010**, *27*, 441–451.
80. Light, S.; Sagit, R.; Sachenkova, O.; Ekman, D.; Elofsson, A. Protein expansion is primarily due to indels in Intrinsically Disordered Regions. *Mol. Biol. Evol.* **2013**, *30*, 2645–2653.
81. Kashi, Y.; King, D.G. Has simple sequence repeat mutability been selected to facilitate evolution? *Isr. J. Ecol. Evol.* **2007**, *52*, 331–342.
82. Kashi, Y.; King, D.G. Simple sequence repeats as advantageous mutators in evolution. *Trends Genet.* **2006**, *22*, 253–259.
83. King, D.G.; Soller, M.; Kashi, Y. Evolutionary tuning knobs. *Endeavour* **1997**, *21*, 36–40.
84. Chatterjee, N.; Lin, Y.; Santillan, B.A.; Yotnda, P.; Wilson, J.H. Environmental stress induces trinucleotide repeat mutagenesis in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 3764–9.
85. Jeffery, C.J. Moonlighting proteins. *Trends Biochem. Sci.* **1999**, *24*, 8–11.
86. Fields, P.A. Review: Protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **2001**, *129*, 417–31.
87. Taviani, M. Post-Miocene reef faunas of the Red Sea: glacio-eustatic controls. In *Sedimentation and Tectonics in Rift Basins Red Sea:- Gulf of Aden*; Springer Netherlands: Dordrecht, 1998; pp. 574–582.
88. Bruggemann, J.H.; Buffler, R.T.; Guillaume, M.M.M.; Walter, R.C.; Von Cosel, R.; Ghebretensae, B.N.; Berhe, S.M. Stratigraphy, palaeoenvironments and model for the deposition of the Abdur Reef Limestone: Context for an important archaeological site from the last interglacial on the Red Sea coast of Eritrea. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2004**, *203*, 179–206.
89. Casazza, L.R. Pleistocene reefs of the Egyptian Red Sea: environmental change and community persistence. *PeerJ* **2017**, *5*, e3504.
90. Dullo, W.C. Facies, fossil record, and age of pleistocene reefs from the Red Sea (Saudi Arabia). *Facies* **1990**, *22*, 1–45.

743 91. El-Sorogy, A. S. Contributions to the Pleistocene coral reefs of the Red Sea Coast, Egypt.
744 *Arab Gulf Journal of Scientific Research*, **2008**, *26*, 63–85.
745 92. Rohling, E.J.; Grant, K.; Hemleben, C.; Siddall, M.; Hoogakker, B.A.A.; Bolshaw, M.;
746 Kucera, M. High rates of sea-level rise during the last interglacial period. *Nat. Geosci.*
747 **2008**, *1*, 38–42.
748 93. Siddall, M.; Rohling, E.J.; Almogi-Labin, A.; Hemleben, C.; Meischner, D.; Schmelzer, I.;
749 Smeed, D. a Sea-level fluctuations during the last glacial cycle. *Nature* **2003**, *423*, 853–
750 858.
751 94. Taviani, M.; López Correa, M.; Zibrowius, H.; Montagna, P.; McCulloch, M.; Ligi, M.
752 Last glacial deep-water corals from the Red Sea. In Proceedings of the Bulletin of Marine
753 Science; **2007**, *81*, 361–370.
754 95. DiBattista, J.D.; Howard Choat, J.; Gaither, M.R.; Hobbs, J. P.A.; Lozano-Cortés, D.F.;
755 Myers, R.F.; Paulay, G.; Rocha, L.A.; Toonen, R.J.; Westneat, M.W.; et al. On the origin
756 of endemic species in the Red Sea. *J. Biogeogr.* **2016**, *43*, 13–30.
757 96. Trommer, G.; Siccha, M.; Rohling, E.J.; Grant, K.; Van der Meer, M.T.J.; Schouten, S.;
758 Hemleben, C.; Kucera, M. Millennial-scale variability in Red Sea circulation in response
759 to Holocene insolation forcing. *Paleoceanography* **2010**. *25*
760 97. Fine, M.; Gildor, H.; Genin, A. A coral reef refuge in the Red Sea. *Glob. Chang. Biol.*
761 **2013**, *19*, 3640–3647.