frontiers in
MICROBIOLOGY

REVIEW ARTICLE
published: 07 July 2014
doi: 10.3389/fmicb.2014.00336

A genomic approach to coral-dinoflagellate symbiosis: studies of *Acropora digitifera* and *Symbiodinium minutum*

Chuya Shinzato, Sutada Mungpakdee, Nori Satoh* and Eiichi Shoguchi

Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

Edited by:
Monica Medina, Pennsylvania State University, USA

Reviewed by:
Malcolm Hill, University of Richmond, USA
Daniel J. Thornhill, Defenders of Wildlife, USA

****Correspondence:***
Nori Satoh, Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Okinawa 904-0495, Japan
e-mail: norisky@oist.jp

Far more intimate knowledge of scleractinian coral biology is essential in order to understand how diverse coral-symbiont endosymbioses have been established. In particular, molecular and cellular mechanisms enabling the establishment and maintenance of obligate endosymbiosis with photosynthetic dinoflagellates require further clarification. By extension, such understanding may also shed light upon environmental conditions that promote the collapse of this mutualism. Genomic data undergird studies of all symbiotic processes. Here we review recent genomic data derived from the scleractinian coral, *Acropora digitifera,* and the endosymbiotic dinoflagellate, *Symbiodinium minutum.* We discuss *Acropora* genes involved in calcification, embryonic development, innate immunity, apoptosis, autophagy, UV resistance, fluorescence, photoreceptors, circadian clocks, etc. We also detail gene loss in amino acid metabolism that may explain at least part of the *Acropora* stress-response. Characteristic features of the *Symbiodinium* genome are also reviewed, focusing on the expansion of certain gene families, the molecular basis for permanently condensed chromatin, unique spliceosomal splicing, and unusual gene arrangement. Salient features of the *Symbiodinium* plastid and mitochondrial genomes are also illuminated. Although many questions regarding these interdependent genomes remain, we summarize information necessary for future studies of coral-dinoflagellate endosymbiosis.

Keywords: corals, symbiosis, *Symbiodinium*, genome, transcriptome

INTRODUCTION

Coral reefs and tropical forests are places that foster the greatest diversities of organisms on the earth. Even though coral reefs occupy only ∼1% of the seas, they are estimated to harbor around one-third of all described marine species (Wilkinson, 2004), and their productivity supports around one quarter of marine fisheries. However, due to human activities and climate changes, reefs decline in abundance, and wholesale loss of reef habitats is one of the most pressing environmental issues of our time.

The major architects of coral reefs, the scleractinian corals, are anthozoan cnidarians that form obligate endosymbioses with photosynthetic dinoflagellates of the genus *Symbiodinium*. The symbionts confer upon the coral holobiont the ability to fix CO_2 and to deposit the massive aragonite (a form of calcium carbonate) skeletons that distinguish reef-building corals from other anthozoans, such as sea anemones. The association is fragile however, collapsing under stress and from disease. Molecular and cellular mechanisms underlying much of coral biology, including the establishment, maintenance, and breakdown of coral-*Symbiodinium* symbioses remain to be elucidated.

In order to investigate mechanisms that support this mutualism, genomic information from both corals and *Symbiodinium* is essential. Proteomics approaches have also been applied to coral and *Symbiodinium* studies (Drake et al., 2013; Ramos-Ailva et al., 2013). Following cloning and characterization of single genes (e.g., Berghammer et al., 1996; Miller et al., 2000), the first large molecular dataset available for a coral was a collection of ∼3000 expression sequence tags (ESTs) from the Indo-Pacific complex coral, *Acropora millepora* (Kortschak et al., 2003). Since then, several EST data sets and transcriptomics studies in corals, as well as *Symbiodinium* spp. have appeared (**Tables 1, 2**). In 2011, a draft genome of *Acropora digitifera* was decoded (**Table 1**) (Shinzato et al., 2011). Then, in 2013, a draft genome of *Symbiodinium minutum* was decoded (**Table 2**) (Shoguchi et al., 2013a). The present review describes characteristic features of these two genomes, with the hope that this information may support future studies of coral biology.

THE *ACROPORA DIGITIFERA* GENOME

The genome of *A. digitifera,* decoded using next-generation sequencing technology, is ∼420-Mbp in size, 39% G+C, and contains 23,668 predicted protein-coding loci (Shinzato et al., 2011). The coral gene set is comparable in size and composition to those of *Nematostella vectensis* (Putnam et al., 2007) and *Hydra magnipapillata* (Chapman et al., 2010). The *A. digitifera* genome browser is accessible at http://marinegenomics.oist.jp/acropora_digitifera (Koyanagi et al., 2013). Approximately 93% of *A. digitifera* genes have homologs in other metazoans (**Figure 1A**), and of these, 11% have significant homology only amongst EST data from corals (**Figure 1B**) (Hemmrich and Bosch, 2008), suggesting the presence of a considerable number of

coral-specific genes. As discussed later, the *Acropora* nuclear DNA sequences do not contain any *Symbiodinium*-related genome sequences.

EVOLUTIONARY ORIGINS OF REEF-BUILDING CORALS

Corals are morphologically very similar to sea anemones, but their evolutionary origins are obscure. Reef building scleractinians first appeared in the fossil record in the mid Triassic (∼240 MYR) (Stanley and Fautin, 2001), but were already highly diversified, suggesting much earlier origins. The availability of fully sequenced genomes for three cnidarians (*Acropora, Nematostella*, and *Hydra*) allows us to estimate the time of divergence between corals and other metazoans. Molecular phylogenetic analyses, based on an alignment of 94,200 amino acids, suggest a divergence time of 520 ∼ 490 MYR for *Acropora* and *Nematostella* (late Cambrian or early Ordovician). This implies early origin of Scleractinia indicates that corals have persisted through previous periods of dramatic environmental change, including the mass extinction event at the Permian/Triassic boundary, when global CO_2 and temperature were much higher than at present. However, molecular phylogeny of symbiotic dinoflagellates suggests that *Symbiodinium* originated in early Eocene, and that the majority of extant lineages diversified since Mid-Eocene, ∼18 MYR ago (Pochon et al., 2006). Therefore, it is far from certain

Table 1 | Published genomics and transcriptomics datasets of scleractinian corals.

Dataset	Species	Sequencing technologies	References
Genome	*Acropora digitifera*	454, Illumina	Shinzato et al., 2011
Transcriptome	*Acropora millepora*	Sanger, 454, Illumina	Moya et al., 2012
	Acropora hyacinthus	Illumina	Barshis et al., 2013
	Acropora palmata	Sanger, 454	Polato et al., 2011
	Acropora cervicornis	Illumina	Libro et al., 2013
	Porites australiensis	Illumina	Shinzato et al., 2014
	Porites astreoides	454	Kenkel et al., 2013
	Favia sp.	Illumina	Mehr et al., 2013
	Montastraea faveolata	Sanger	Schwarz et al., 2008
	Stylophora pistillata	454	Karako-Lampert et al., 2014
	Pocillopora damicornis	454	Traylor-Knowles et al., 2011

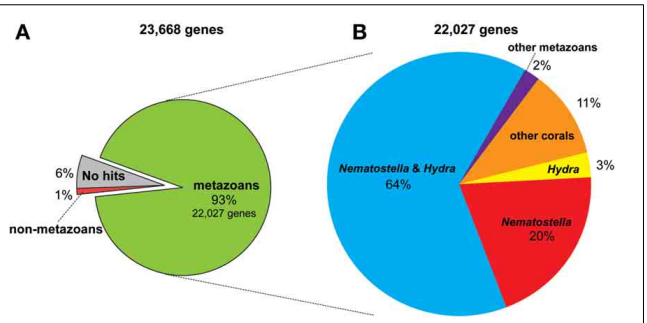


FIGURE 1 | The proportion of similarity of *Acropora digitifera* gene models to those of other metazoans (see text for the method). (A) The 93% (22,027) of the 23,668 gene models have similarity to other metazoans, while 1% of them for non-metazoans and 6% show no similarity to proteins deposited in NCBI database (NR). **(B)** Of the 22,027 metazoan-similarity genes, 64% had counterparts in both *Nematostella* and *Hydra* genomes, 20% of them for *Nematostella* but not in *Hydra*, 3% for *Hydra* but not in *Nematostella,* and 2% for other metazoans but not in *Nematostella* and *Hydra*. Eleven percent of them have corresponding ESTs of corals reported in NCBI.

Table 2 | Published genomics and transcriptomics datasets of *Symbiodinium*.

Dataset	Species (strain ID)	Clade	Host	Sequencing technologies	References
Genome	*Symbiodinium minutum* (Mf 1.05b.01)	B1	*Montastraea faveolata*	454, Illumina	Shoguchi et al., 2013a
Transcriptomes	*Symbiodinium minutum* (Mf 1.05b.01)	B1	*Montastraea faveolata*	Illumina	Shoguchi et al., 2013a
	Symbiodinium microadriaticum (CCMP2467)	A1	*Stylophora pistillata*	Illumina	Baumgarten et al., 2013
	Symbiodinium spp.	C	*Acropora hyacinthus*	Illumina	Ladner et al., 2012
	Symbiodinium spp.	D	*Acropora hyacinthus*	Illumina	Ladner et al., 2012
	Symbiodinium sp. (Mf1.05b)	B1	*Montastraea faveolata*	454	Bayer et al., 2012
	Symbiodinium sp. (CassKB8)	A	*Cassiopea* sp.	454	Bayer et al., 2012
	Symbiodinium sp.	C3K	*Acropora hyacinthus*	Illumina	Barshis et al., 2014*
	Symbiodinium sp.	D2	*Acropora hyacinthus*	Illumina	Barshis et al., 2014*
	Symbiodinium sp.	C15	*Porites australiensis*	Illumina	Shinzato et al., 2014*
	Symbiodinium kawagutii (CCMP2468)	F1	*Montipora verrucosa*	Sanger	Zhang et al., 2013
	Symbiodinium sp.	A	*Aiptasia pallida*	Sanger	Sunagawa et al., 2009
	Symbiodinium sp. (CassKB8)	A	*Cassiopea* sp.	Sanger	Voolstra et al., 2009
	Symbiodinium sp.	C3	*Acropora aspera*	Sanger	Leggat et al., 2007

**From a mixed host/symbiont cDNA library.*

that modern coral reefs can adapt to the rapid environmental changes now occurring.

TRACES OF SYMBIOSIS IN THE CORAL GENOME

Obligate endosymbiosis of corals dates from at least the mid Triassic (Stanley and Fautin, 2001), and the longevity of this association might be expected to have resulted in changes in the coral genome. However, a comprehensive search of *Acropora* nuclear DNA sequences failed to find any *Symbiodinium* DNA sequences (Shinzato et al., 2011); hence there is, as yet, no evidence for horizontal gene transfer from symbiont to host. Neither is *Symbiodinium* vertically transferred via host gametes. As a result, the symbiosis must be re-established with each generation. Nonetheless, comparative analyses imply that *Acropora* is probably metabolically dependent upon its endosymbiont.

When the metabolic repertoire of *A. digitifera* was compared using the KEGG pathway database to that of its non-symbiotic relative, *Nematostella*, it became apparent that *Acropora* lost a gene for cysteine biosynthesis. Biosynthesis of cysteine from homocysteine and/or serine requires two enzymes, cystathionine beta-synthase (Cbs) and cystathionase (cystathionine gamma-lyase) (**Table 3**). Although both the *A. digitifera* and *Nematostella* genomes encode cystathionase, the gene for Cbs could not be identified in *Acropora* despite the existence of an ortholog in *Nematostella* (**Table 3**). An extensive search of transcriptomic data available for various *Acropora* spp. (Hemmrich and Bosch, 2008) failed to identify a *Cbs* transcript in any congener. Moreover, whereas a PCR strategy confirmed the presence of *Cbs* in some other corals (*Galaxea fascicularis, Favites chinenis, Favia lizardensis*, and *Ctenactis echinata*), no amplification products could be obtained for two different *Acropora* species (**Table 3**). Although further studies of biosynthetic pathways are required, this finding raises the intriguing possibility of a metabolic basis for the obligate nature of symbiosis in *Acropora*. Differences in dependency could potentially explain not only the phenomenon of symbiont selectivity, but also the high sensitivity of *Acropora* to environmental challenges.

Table 3 | The presence or absence of a gene encoding cystathionine β-synthase (Cbs) for L-cysteine biosynthesis in corals.

	L-Homocysteine + L-Serine	*Cbs* →	L-Cystathionine	*Cth* →	L-Cysteine
Hydra magnipapillata		Yes[a]		Yes	
Nematostella vectensis		Yes[a]		Yes	
COMPLEXA					
Acropora digitifera		–[b]		Yes	
Acropora tenuis		–[c]		ND	
Acropora millepora		–[d]		Yes	
Galaxea fascicularis		Yes[c]		ND	
ROBUSTA					
Montastraea faveolata		Yes[d]		Yes	
Favia lizardensis		Yes[c]		ND	
Favites chinensis		Yes[c]		ND	
Ctenactis echinata		Yes[c]		ND	

ND, not determined.

[a] *Supported by sequenced genome and EST analyses.*

[b] *Supported by sequenced genome, EST, and PCR amplification of genome DNA.*

[c] *Supported by PCR-amplification of genome DNA.*

[d] *Supported by EST analyses.*

GENES INVOLVED IN CALCIFICATION

The coral gene repertoire, with predicted roles in skeleton deposition, is of particular interest, given the likely impact of ocean acidification resulting from rising atmospheric CO_2 on coral calcification. Surveys of the *Acropora* genome reveal the presence of genes for specific groups of proteins associated with calcification, including the eukaryotic carbonic anhydrases (Jackson et al., 2007). In general, the soluble fraction of the organic matrix (OM) in invertebrates is very rich in acidic amino acids, and has a particularly high aspartic acid composition (Sarashina and Endo, 2006). A number of candidate OM protein genes are present in the *Acropora* genome. Galaxins, first purified from the coral, *G. fascicularis*, are unique to corals and are the only coral skeletal matrix protein for which the complete primary structure has been determined (Fukuda et al., 2003). However, galaxin possesses neither acidic regions (the fraction of Asp+Asn in the galaxin is only 9.7%) nor obvious Ca^{2+} binding domains. Four genes encoding galaxin-related proteins have been identified in the *A. digitifera* genome, including two likely *A. digitifera* homologs of galaxin.

TRANSCRIPTION FACTOR GENES AND SIGNALING MOLECULE GENES

Cnidarians have genes for transcription factors and signaling molecules comparable to those found in bilaterians (Technau et al., 2005; Putnam et al., 2007) and this is also true of corals (Shinzato et al., 2011). Of those, genes for Hox cluster and basic helix-loop-helix (bHLH) families have been examined in detail in the *A. digitifera* genome.

Hox genes

Hox genes are homeobox transcription factors that play a critical role in developmental patterning (McGinnis et al., 1984). They have been identified in every extant phylum except the Porifera, Ctenophora, and Placozoa. Cnidarians are the only non-bilaterian phylum with *Hox* genes; therefore they are critical to our understanding of early *Hox* cluster evolution. However, the *H. magnipapillata* genome shows no *Hox* cluster (Chapman et al., 2010) and clustering in *N. vectensis* is limited to anterior Hox genes (Chourrout et al., 2006; Putnam et al., 2007; Ryan et al., 2007), raising the question of the degree of *Hox* gene clustering in cnidarians. The *A. digitifera* genome has the most extensive *Hox* cluster reported in any cnidarian (DuBuc et al., 2012). Phylogenetic analysis revealed a total of six *Hox*, one *ParaHox*, three *Mox*, one *Eve*, and one *HlxB9* gene in the *Acropora* genome. Of the six *Hox* genes, two anterior (PG1 and PG2) linked to an *Eve* homeobox gene and an *Anthox1A* gene (**Figure 2**). Therefore, the *Hox* cluster of the cnidarian–bilaterian ancestor was more extensive than previously thought. These facts are congruent with the existence of an ancient set of constraints on the *Hox* cluster and reinforce the importance of incorporating a wide range of animal species to reconstruct critical ancestral nodes.

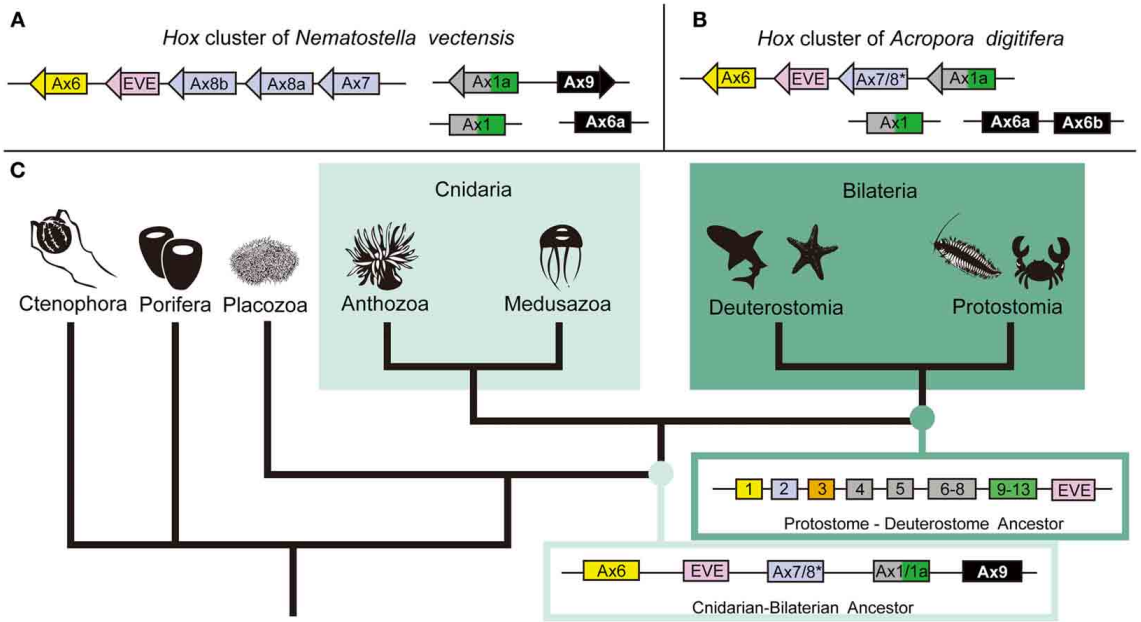


FIGURE 2 | The anthozoan complement of Hox genes and the implications of the evolution of the Hox cluster. Comparing the genomic linkage of Hox genes in the sea anemone *N. vectensis* and the staghorn coral *A. digitifera* confirms that cnidarians once had a Hox cluster that contained both anterior and posterior/central class Hox genes. **(A)** The Hox cluster of *N. vectensis* includes the anterior Hox genes Anthox6 (PG1), Anthox8b (PG2), Anthox8a (PG2), and Anthox7 (PG2) as well as the Eve homeobox gene. **(B)** The Hox cluster of *A. digitifera* includes the anterior Hox genes Anthox6 (PG1) and Anthox7/8 (PG2), and the posterior/central class Hox gene Anthox1a (PG4–14), as well as the Eve homeobox gene. Another gene HlxB9 (also named MNX) is found upstream of Anthox6 in the Hox cluster of both genomes (data not shown). **(C)** The metazoan tree of life with inferred ancestral Hox clusters. The ancestor to protostomes and deuterostomes is thought to have had two anterior class Hox genes (Hox1 and Hox2), one paralagous group 3 gene (Hox3), three central class genes (Hox4, Hox5, and Hox6–8), one posterior class Hox gene (Hox9–14), and one Eve homeobox gene. Because of the extended cluster in A. digitifera, we can now say that the cnidarian–bilaterian ancestor had, at least, two anterior class Hox genes (Anthox6 and Anthox7/8), a central/posterior class Hox gene (Anthox1/1a), and the Eve homeobox gene. It is unclear at what point the genomic rearrangement involving the Eve homeobox gene occurred. The origin of the PG3 Hox genes also is not clear. *Anthox7/8 has been categorized as a PG2 Hox gene in previous publications, but it is possible, based on our current phylogenetic analysis, that Anthox7/8 descended from a Hox gene that was lost in bilaterians. Based on the genomic orientation of these genes, we also believe the ancestor likely had a fourth Hox gene potentially related to Anthox9. For more detail information, please see DuBuc et al. (2012). Abbreviations: PG, paralogous group; Ax, Anthox.

bHLH genes

bHLH proteins constitute a large group of transcription factors that comprise a basic region for DNA binding and two α-helices, interrupted by a variable loop region, for dimerization. bHLH proteins homo- or heterodimerize to recognize and bind specific core hexa-nucleotides, and play pivotal roles in cell differentiation and proliferation (Massari and Murre, 2000; Jones, 2004). A putative full set of bHLH genes has been described in the genomes of a number of metazoans, and molecular phylogenetic analyses have identified 45 orthologous families of bHLH factors, which were categorized into six high order groups (Atchley and Fitch, 1997).

The *A. digitifera* genome contains a nearly full set of 70 bHLH transcription factors, comparable to the 68 bHLH genes in *N. vectensis* (Gyoja et al., 2012). The *Acropora* genes have been assigned to 29 previously reported orthologous families. In addition, three novel HLH orthologous families have been identified, designated pearl, amber, and peridot (Gyoja et al., 2012). Pearl and amber orthologs are present in genomes and ESTs of the Mollusca and Annelida, in addition to the Cnidaria. Peridot orthologs are present in genomes and ESTs of the Cephalochordata and the Hemichordata, in addition to the Cnidaria. These three genes have apparently been lost in the clades of *Drosophila*, *Caenorhabditis*, and *Homo sapiens*. Therefore, cnidarians provide information about alteration of transcription factor genes during animal evolution.

INNATE IMMUNITY

Innate immunity in corals is of special interest not only in the context of self-defense, but also in relation to the establishment and collapse of the obligate symbiosis with *Symbiodinium*. The coral innate immune repertoire is highly complex and more sophisticated than that of *Hydra* and *Nematostella* (**Figure 3**) (Shinzato et al., 2011; Hamada et al., 2013). For example, whereas a single canonical Toll/TLR protein is present in *N. vectensis* (Miller et al., 2007), the *Acropora* genome encodes at least four such molecules, as well as five IL-1R-related proteins, and a number of TIR-only proteins (**Figure 3A**). Likewise, the *Acropora* repertoire of NACHT/NB-ARC domains, which are characteristic of primary intracellular pattern receptors, is again highly complex—an order of magnitude more NACHT/NB-ARC domains are present in coral than in other animals, and some of these cnidarian proteins have novel domain structures.

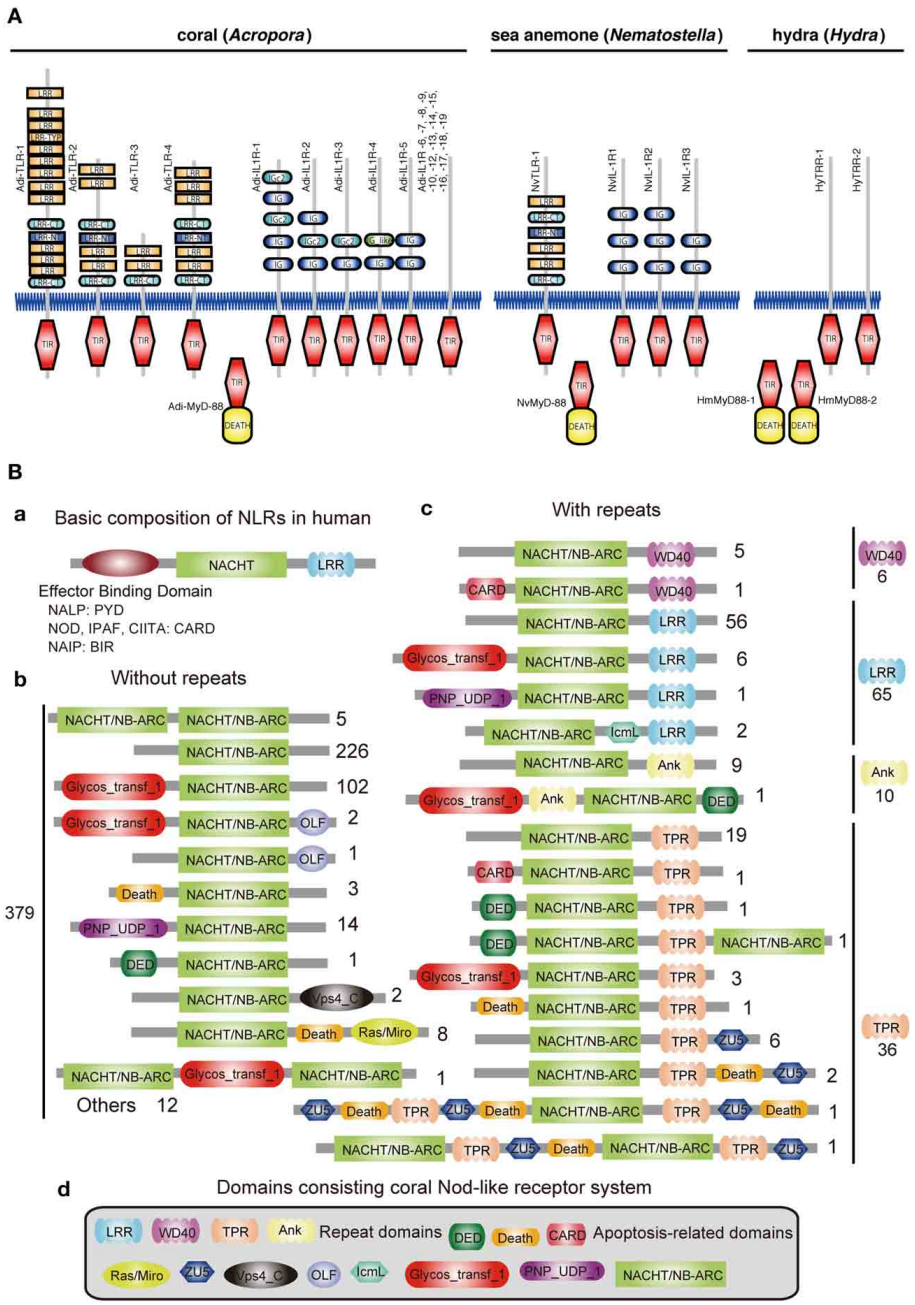


FIGURE 3 | Innate immunity of corals. (A) Repertoires of TIR domain-containing proteins of three cnidarians. The schematic representation of the domain structures of all of the TIR domain-containing proteins identified in *Acropora digitifera*, alongside the corresponding complements from *Nematostella vectensis* and *Hydra magnipapillata*. The repertoire of Toll/TLR, IL-1R-like, and TIR-only proteins is significantly more complex in the case of *A. digitifera* than in *N. vectensis* or *H. magnipapillata*. TIR, TIR domain; DEATH, DEATH domain; IG and IGc2, Ig domain; LRR, LRY-TRY, LRR-CT and LRR-NT, leucine-rich repeats. **(B)** The complexity of the NBD repertoire of *Acropora digitifera*. The figure summarizes the numbers of loci with each kind of domain architecture. (a) The basic composition of NRLs in human is, from N- to C-terminus, effector binding domain, NACHT domain and repeats (LRR). The effector binding domain components are PYD in NALPs, CARD in NODs, IPAF, CIITA and BIR in NAIP. (b) A total of 379 coral NBD loci do not encode repeat domains. Numbers to the right of schematics represent the number of loci with each specific architecture. (c) In addition, 117 loci in the coral encode NBDs and repeat domains of the WD40, LRR, Ank, or TPR types. (d) The various domains identified in the Nod-like proteins of *Acropora*.

In the vertebrate innate immune system, ∼20 tripartite nucleotide oligomerization domain (NOD)-like receptor proteins that are defined by the presence of NAIP, CIIA, HET-E, and TP1 (NACHT) domains, a C-terminal leucine-rich repeat (LRR) domain, and one of three types of N-terminal effector domain, are known to function as primary intracellular pattern recognition molecules (**Figure 3B**) (Hamada et al., 2013). Surveying the coral genome demonstrates a larger number of NACHT- and

related domain nucleotide-binding adaptors shared by APAF-1, R proteins, and CED-4 (NB-ARC)-encoding loci (∼500) than in other metazoans, and also a surprising diversity of domain combinations among coral NACHT/NB-ARC-containing proteins (**Figure 3B**). N-terminal effector domains include apoptosis-related domains, caspase recruitment domains (CARD), death effector domains (DED), and Death, and C-terminal repeat domains, such as LRRs, tetratricopeptide repeats, ankyrin repeats, and WD40 repeats. Many of the predicted coral proteins that contain a NACHT/NB-ARC domain also contain a glycosyl transferase group 1 domain, a novel domain combination first found in metazoans. Phylogenetic analyses suggest that the NACHT/NB-ARC domain inventories of various metazoan lineages, including corals, are largely products of lineage-specific expansions. Many of the NACHT/NB-ARC loci are organized in pairs or triplets in the *Acropora* genome, suggesting that the large coral NACHT/NB-ARC repertoire has been generated at least in part by tandem duplication (Hamada et al., 2013). In addition, shuffling of N-terminal effector domains may have occurred after diversification of specific NACHT/NB-ARC-repeat domain types. These attributes illustrate the extraordinary complexity of the innate immune repertoire of corals, which may reflect adaptation to a symbiotic lifestyle in a uniquely complex and challenging environment.

APOPTOSIS

The apoptotic network of *A. digitifera* is comparable in complexity to those of "higher" animal taxa, including vertebrates (**Figure 4A**) (Shinzato et al., 2011). Seven Bcl-2 family members containing multiple domains, four IAP family members, 25 caspases, a single APAF-1, four Death receptors, three Death ligands, and 32 members of the TRAF adaptor family are present in the *Acropora* genome (**Figure 4B**). These numbers are generally comparable to those in the *Nematostella* genome. The TRAF family in *Acropora* and *Nematostella* and the caspases in *Acropora* are overrepresented relative to humans. While no BH3-only members of the Bcl-2 family have been identified (**Figure 4B**), this may be a consequence of the small size of the BH3 domain and the extent of sequence divergence in these proteins. Failure to detect adaptors with Death domains may reflect the low level of domain conservation characteristic of this family.

AUTOPHAGY

The *A. digitifera* genome contains orthologs of ATG1, ATG2, ATG3, ATG4, ATG5, ATG6, ATG7, ATG8, ATG9, ATG 10, ATG12, ATG13, ATG14, ATG16, ATG18, ATG24, TOR, Vsp34, and Vsp15, but no counterparts of the yeast-specific proteins ATG11, ATG15, ATG17, ATG19, ATG20, ATG21, ATG22, ATG23, ATG26, ATG27, and ATG29 (Shinzato et al., 2011) (**Figure 5**). The *Acropora* genome also encodes orthologs of human UVRAG, SH3GLB1, DRAM, AMBRA1, RB1CC1, and ATG101 (**Figure 5**), which are also absent in yeast.

GENES INVOLVED IN UV-DAMAGE PROTECTION

Reef-building corals typically inhabit shallow and relatively clear tropical waters and are therefore constantly exposed to high levels of UV irradiation. Since high solar radiation sometimes causes coral bleaching (Gleason and Wellington, 1993), one intriguing question is how corals protect themselves against UV-damage. UV-absorbing substances potentially act as photoprotective compounds. These include mycosporine-like amino acids (MAAs), scytonemin, carotenoids, and others of unknown chemical structure (Shick et al., 1999; Reef et al., 2009). Although some photoprotective compounds have been isolated from corals (Rastogi et al., 2010), it is often unclear whether symbiotic dinoflagellates and/or bacteria produce the photoprotective compounds, or whether the corals themselves can independently synthesize them.

MAAs

A recent study of the cyanobacterium, *Anabaena variabilis*, identified a four-gene cluster (encoding DHQS-like, O-MT, ATP-grasp, and NRPS-like enzymes) that converts pentose-phosphate metabolites into shinorine, one of MAAs (**Figure 6**) (Balskus and Walsh, 2010). A search of cnidarian gene models for components of the shinorine gene cluster revealed that this four-gene pathway is present in both *Acropora* and *Nematostella*, but not in *Hydra* (Shinzato et al., 2011). This strongly suggests that both *Acropora* and *Nematostella* can synthesize shinorine by themselves, which may be a precursor for photoprotective compounds.

In addition, molecular phylogenetic analyses show that homologous proteins in *Acropora* have more sequence similarities to those of bacteria and dinoflagellates (Shinzato et al., 2011). These genes might have been acquired via horizontal gene transfer (Starcevic et al., 2008). For example, during the evolution of cnidarian stinging cells, a subunit of bacterial poly-γ-glutamate (PGA) synthase was transferred to an animal ancestor via horizontal gene transfer (Denker et al., 2008). It has been proposed that in marine environments, horizontal gene transfer is important in adapting to ecological vagaries (Keeling, 2009).

Scytonemin

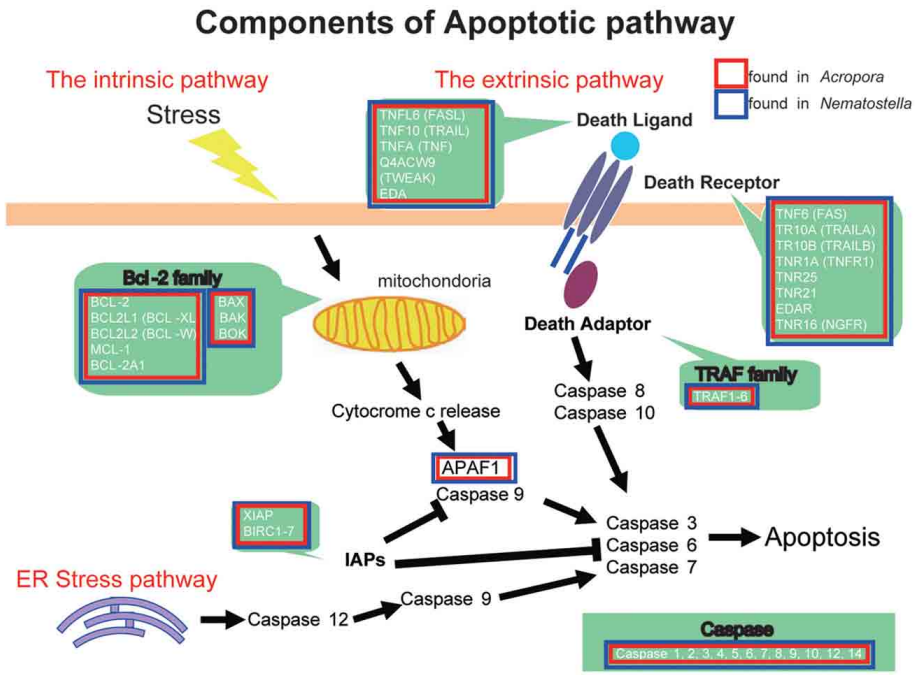
The UV-blocker, scytonemin, is found exclusively in cyanobacteria. In *Nostoc punctiforme,* its biosynthesis is controlled by a cluster of 18 genes (**Figure 7**) (Soule et al., 2007; Balskus and Walsh, 2008). The cluster comprises one subcluster of genes involved in aromatic amino acid biosynthesis, and a novel subcluster of genes of unknown function (Soule et al., 2009). The former includes *tyrA*, *dsbA*, *aroB*, *trpE*, *trpC*, *trpA*, *tyrP*, *trpB*, *trpD*, and *aroG* (**Figure 7B**). The latter includes *scyA*, *scyB*, *scyC*, *scyD*, *scyE*, and *scyF* (**Figure 7B**).

The *A. digitifera* genome contains only six of the 18 genes: namely, *scyA*, *scyB*, *scyF*, *dsbA*, *aroB*, and *tyrP* (**Figure 7**) (Shoguchi et al., 2013c). This result suggests that coral cannot synthesize scytonemin independently. Molecular phylogenetic analyses indicate that coral *scyA* and *scyB* are associated with bacterial genes for acetolactate synthase and glutamate dehydrogenase, respectively. This suggests that these enzymes are coupled with PGA/amino acid biosynthesis in corals. In addition, *scyA*, *scyB*, and *aroB* (*DHQS-like*) are likely to have originated by horizontal transfer from bacteria.

Glyoxylate cycle enzymes: malate synthase and isocitrate lyase

Glyoxylate cycle enzymes play a role in lipid metabolism in plant seeds (Kornberg and Beevers, 1957). Although this pathway has

A



B

Number of Apotosis- related genes

Protein Family	Human Protein	Domain (Pfam ID)	Ad	Nv	Hs
Bcl-2	BCL-2, BCL2L1 (BCL-XL), BCL2L2 (BCL-W), MCL-1, B2LA1, B2L10, B2L13, BAX, BAK, BOK	Bcl-2 (PF00452)	10	18	10
IAP	XIAP, BIRC1-7	BIR (PF00653)	6	7	8
caspase	caspase1,2,3,4,5,6,7,8,9,10,12,14	Peptidase_C14 (PF00656)	15	16	12
APAF-1	APAF	NB-ARC (PF00931)	1	1	1
Death receptor	TNF6 (FAS), TR10A (TRAILA), TR10E (TRAILB), TNR1A (TNFR1). TNR25, TNR21, EDAR, TNR16 (NGFR)	TNFR_c6 (PF00020)	6	7	8
Death ligand	TNFL6 (FASL), TNF10 (TRAIL), TNFA (TNF), Q4ACW9 (TWEAK), EDA	TNF (PF00229)	7	6	5
Adaptor TRAF family	TRAF1-6	MATH (PF00917)	31	27	6

FIGURE 4 | (A) Schematic presentation of cellular components involved in the pathways of apoptosis, based on human genes. The extrinsic pathway, intrinsic pathway, and ER stress pathway are three major pathways of apoptosis. Major families are shown by green background. Families found in the *Acropora digitifera* genome are boxed by red and those of *Nematostella vectensis* by blue. **(B)** The number of apoptosis-related family members in the genome of *A. digitifera* (Ad), *N. vectensis* (Nv), and *Homo sapiens* (Hs). The *Acropora* and *Nematostella* genomes contain apoptosis-related genes of which numbers are comparable to those of the human genome, except for a larger number of adaptor TRAF family in the cnidarians.

not been found in animal lineages, nematode genomes contain genes encoding enzymes involved in the pathway (Liu et al., 1995). Interestingly, the *A. digitifera* genome contains one *isocitrate lyase* (*ICL*) gene and two *malate synthase* (*MS*) genes. Orthology between *Acropora* and *Nematostella* is supported by molecular phylogenetic analysis (Shoguchi et al., 2013c). The genes, *ICL* and *MS1*, are aligned head-to-head in tandem. In addition, by comparisons between neighboring genes, synteny in the region is also conserved. The anthozoan genes form a clade with bacterial *ICL*. Therefore, the origin of anthozoan

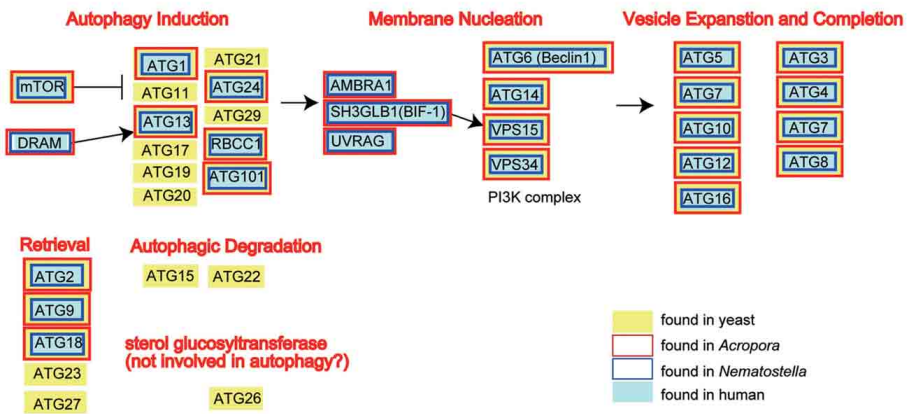


FIGURE 5 | Schematic presentation of the autophagy pathway, and human (gray backgroud) and yeast genes (*Saccharomyces cerevisiae*; yellow background) involved in the pathway. The pathway is composed of autophagy induction, membrane nucleation, vesicle expansion and completion, retrieval and autophagic degradation. Genes found in the *Acropora digitifera* genome are boxed by red, those of *Nematostella* by blue. It is obvious that all the human autophagy-related genes have counterparts in *Acropora* and *Nematostella*. In contrast, autophagy-related genes that are found only in the yeast cannot be found in the cnidarian geneomes.

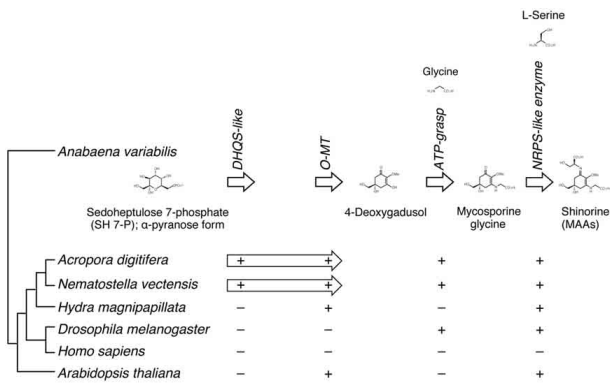


FIGURE 6 | The genes required for the biosynthesis of shinorine are present in anthozoan cnidarians. (Upper) The organization of gene cluster involved in the biosynthetic pathway of the photo-protective molecule shinorine, a mycosporine-like amino acid, in the cyanobacterium Anabaena variabilis. (Lower) The presence of corresponding genes in various organisms is indicated by "+." The Acropora and Nematostella genomes contain homologs of each of the four genes, in which DHQS-like and O-MT are fused each other.

genes may be different from those of nematode glyoxylate cycle enzymes.

FLUORESCENT PROTEINS

Corals exhibit diverse colors, which depend largely on fluorescent proteins (Matz et al., 1999, 2006). Four basic colors of fluorescent proteins present in corals include cyan (CFP), green (GFP), and red (RFP), and a non-fluorescent blue/purple chromoprotein (Kelmanson and Matz, 2003; Field et al., 2006). Fluorescent proteins are usually composed of ∼230 amino acids. Corals are able to synthesize several different fluorescent or colored moieties from amino acids within fluorescent proteins, via two or three consecutive autocatalytic reactions. While CFP and GFP possess the same chromophore, individual chromophores can differ dramatically in spectroscopic characteristics (Henderson and Remington, 2005; Lukyanov et al., 2006).

The *A. digitifera* genome contains one, five, one, and three candidate genes for CFP, GFP, RFP, and chromoprotein, respectively, (Shinzato et al., 2012). The CFP and GFP genes are clustered in an ∼80-kb genomic region, suggesting that they originated from an ancestral gene by tandem duplication. Since CFP and GFP possess the same chromophore, this gene clustering may provide the first genomic evidence for a common origin of the two proteins. Comparisons of the fluorescent protein genes of closely related coral species suggest an expansion of chromoprotein genes in the *A. digitifera* genome, and of RFP genes in the *A. millepora* genome. RNA-seq analysis shows that *A. digitifera* fluorescent protein genes are expressed during embryonic and larval stages and in adults, suggesting that these genes play a variety of roles in coral physiology.

A wide variety of roles have been attributed to coral fluorescent proteins, including modulating the efficiency of photosynthesis and photoprotection for the symbionts (e.g., Salih et al., 2000) as well as antioxidant functions (Bou-Abdallah et al., 2006; Palmer et al., 2009). Along with cataloging the coral fluorescent protein repertoire, functions of these proteins should be investigated by future studies, especially in the context of molecular mechanisms involved in environmental stress responses of corals, which are associated with collapse of coral-*Symbiodinium* symbiosis.

PHOTORECEPTORS AND CIRCADIAN CLOCK GENES

Corals exhibit circadian behaviors, which play a pivotal role in timing of spawning. However, little is known about the

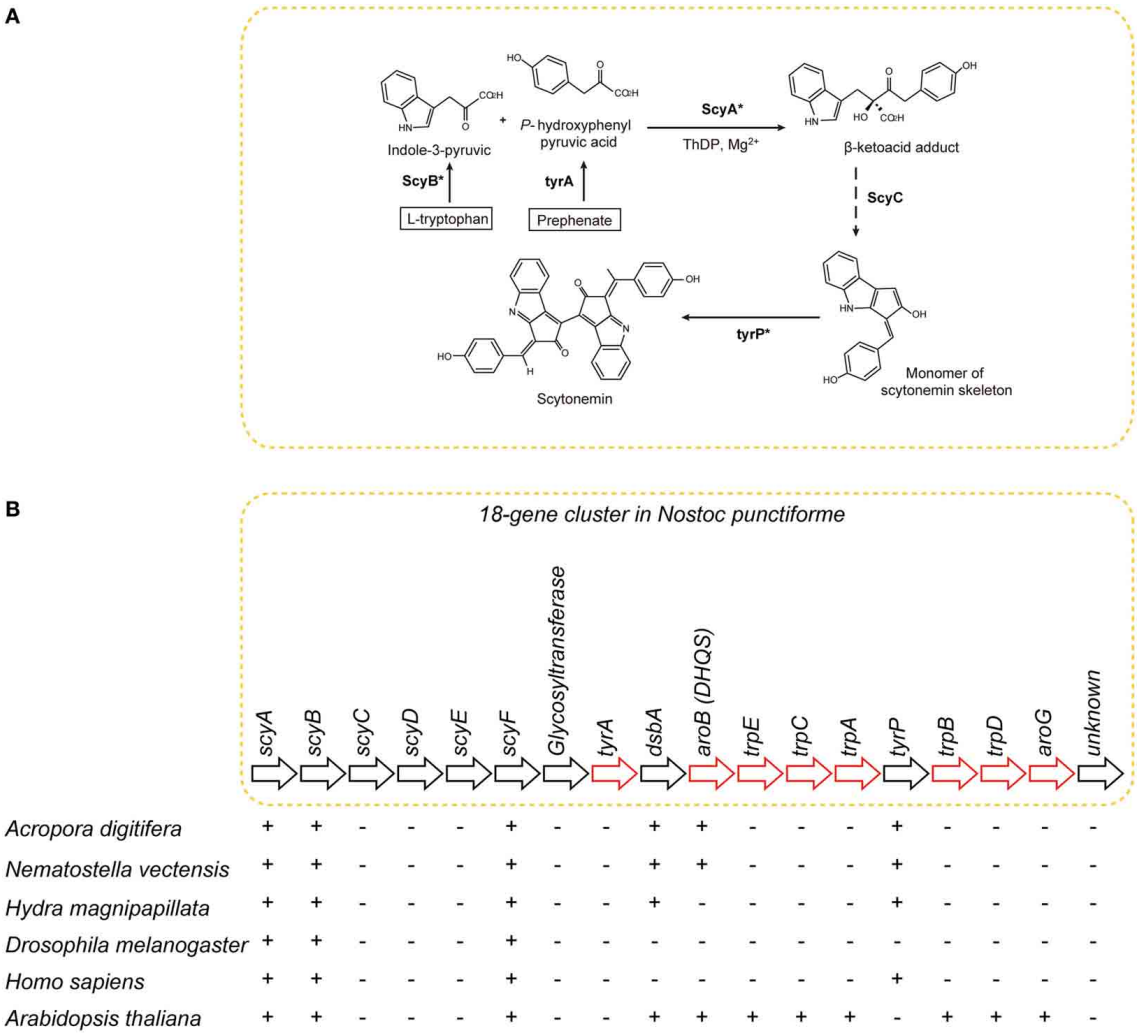


FIGURE 7 | The distribution of genes associated with biosynthesis of scytonemin in cyanobacteria, cnidarians, and other metazoans. (A) Pathways of biosynthesis of the photo-protective molecule scytonemin in the cyanobacterium *Nostoc punctiforme* (Soule et al., 2007; Rastogi et al., 2010). Genes encoding the enzymes indicated with asterisks were identified in the *A. digitifera* genome. **(B)** Schematics showing the organization of the scytonemin gene cluster. Genes indicated by red arrows encode enzymes involved in the biosynthesis of aromatic amino acids. The presence of corresponding genes in various organisms is indicated by "+," indicating that a TBLASTN search against *N. punctiforme* as query showed significant hits. Anthozoan genomes encode a gene, *aroB* homolog, involved in aromatic metabolism that is not found in higher metazoans.

molecular mechanisms underlying the regulation of these behaviors. Microarray analysis of *Acropora-Symbiodinium* suggested complex diel cycles of gene expression (Levy et al., 2011). The *A. digitifera* genome contains seven opsin and three cryptochrome (photoreceptor) genes (**Figure 8**) (Shoguchi et al., 2013b). Two genes from each family likely underwent tandem duplication in the coral lineage. In addition, *A. digitifera* has orthologs to *Drosophila* and mammalian circadian clock genes: four *clock*, one *bmal/cycle*, three *pdp1-like*, one *creb/atf*, one *sgg/zw3*, two *ck2alpha*, one *dco* (*csnk1d/cnsk1e*), one *slim/BTRC*, and one *grinl* (**Figure 8**). However, *Acropora* is unlikely to have *vrille*, *rev-erv*α*/nr1d1*, *bhlh2*, *vpac2*, *adcyap1*, *or adcyaplr1* orthologs (**Figure 8**). Intriguingly, an extensive survey failed to find homologs of *period* and *timeless*, although it found one *timeout* gene. When the coral genes were compared to orthologous genes in *N. vectensis*, a similar repertoire of circadian clock genes was apparent, although *A. digitifera* contains more clock genes and fewer photoreceptor genes than *N. vectensis* (**Figure 8**). This suggests that the circadian clock system was established in a common ancestor of corals and sea anemones, and diversified by tandem gene duplications and the loss of paralogous genes in each lineage. Future studies should examine how the coral circadian clock functions without *period*.

SYMBIODINIUM GENOME

Coral symbionts are all *Symbiodinium* spp. belonging to the phylum Dinoflagellata. Dinoflagellates are unicellular eukaryotes, 10–100 μm in diameter, and characterized by two flagella and a unique cell covering referred to as the theca.

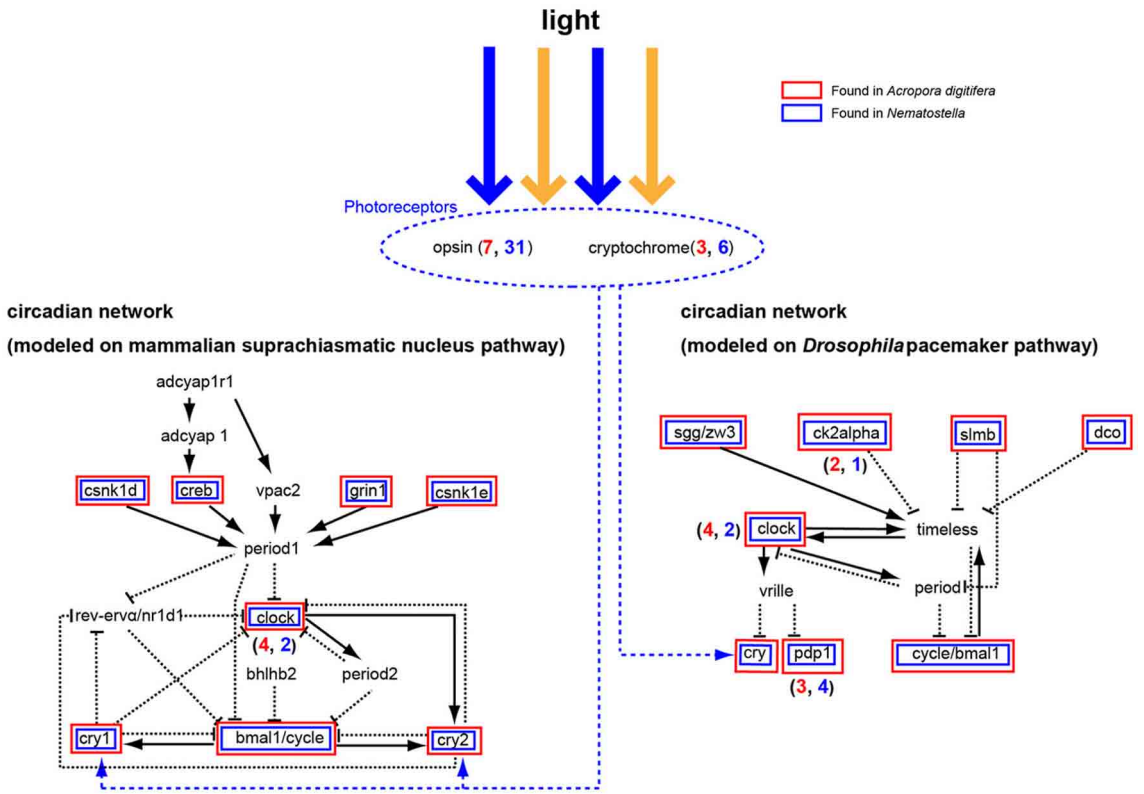


FIGURE 8 | A circadian network modeled on the mammalian suprachiasmatic nucleus pathway (left) and the *Drosophila melanogaster* pacemaker pathway (right). Some cryptochromes detect blue light and function in the network for ∼24 h. The *Acropora digitifera* genome contains seven opsin genes and three cryptochrome genes (blue number shows *A. digitifera* gene number; red number, *N. vectensis*). *Acropora* genes in the pathway are shown in red boxes and *Nematostella* genes in blue. The A. digitifera genome is likely to lack period, bhlhb2, vrille, timeless, rev-erva/nr1d1, adcyap1r1, adcyap1, and vpac2.

Approximately half of them are photosynthetic (Graham and Wilcox, 2000). Dinoflagellates belong to the well-supported Superphylum Alveolata, which also includes ciliates and apicomplexans, such as the malarial parasite, *Plasmodium falciparum* (Burki et al., 2007). Each alveolate lineage has had a distinct evolutionary trajectory with regard to nuclear genome organization, resulting in three divergent outcomes (Gardner et al., 2002; Eisen et al., 2006). Ciliates contain two nuclei, a somatic macronucleus and a micronucleus for reproduction, and they lack plastids. Apicomplexans, due to their parasitic life style in most species, have substantially reduced genomes, with highly degenerate plastids known as apicoplasts (Wilson et al., 1996). Dinoflagellate nuclei have permanently condensed liquid-crystalline chromosomes that lack nucleosomes (**Figures 9A,B**) (Bouligand and Norris, 2001). In addition, recent studies of partial dinoflagellate genome data show repeated gene copies arranged in tandem arrays (Bachvaroff and Place, 2008), trans-splicing of messenger RNAs (Lidie and van Dolah, 2007; Zhang et al., 2007), and a reduced role for transcriptional regulation, compared to other eukaryotes (Erdner and Anderson, 2006; Moustafa et al., 2010). Given these remarkable characteristics, elucidating the structure and composition of dinoflagellate genomes is essential to understanding their packaging of chromosomal DNA and expression of encoded genes. However, dinoflagellates possess some of the largest eukaryotic nuclear genomes (1500–245,000 megabases [Mbp] in size), which have previously thwarted whole-genome sequencing (Lin, 2011; Wisecaver and Hackett, 2011). In 2013, the genome of a culturable dinoflagellate, *S. minutum,* was decoded (Shoguchi et al., 2013a).

THE NUCLEAR GENOME

The genome of *S. minutum* is estimated at ∼1500 Mbp. Approximately 40-fold coverage of the genome yielded a ∼616 Mbp assembly (Shoguchi et al., 2013a). A large quantity of RNA-seq sequences were assembled into 63,104 unique transcripts, 26,691 of which encode complete open reading frames. Gene prediction yielded 41,925 protein models, 77.2% of which (32,366 gene models) are supported by RNA-seq data. In addition, the vast majority of the transcriptome is encoded in the 616-Mbp draft assembly, suggesting that these contigs represent the euchromatin-like region of the *Symbiodinium* genome (http://marinegenomics.oist.jp/genomes/gallery). DNA transposons, retrotransposons, and tandem repeats comprise 0.5, 1.1, and 4.6% of the assembled genome, respectively. The GC-content of the *Symbiodinium* nuclear genome was 44%. This is comparable to GC-content of metazoans and green plants, but

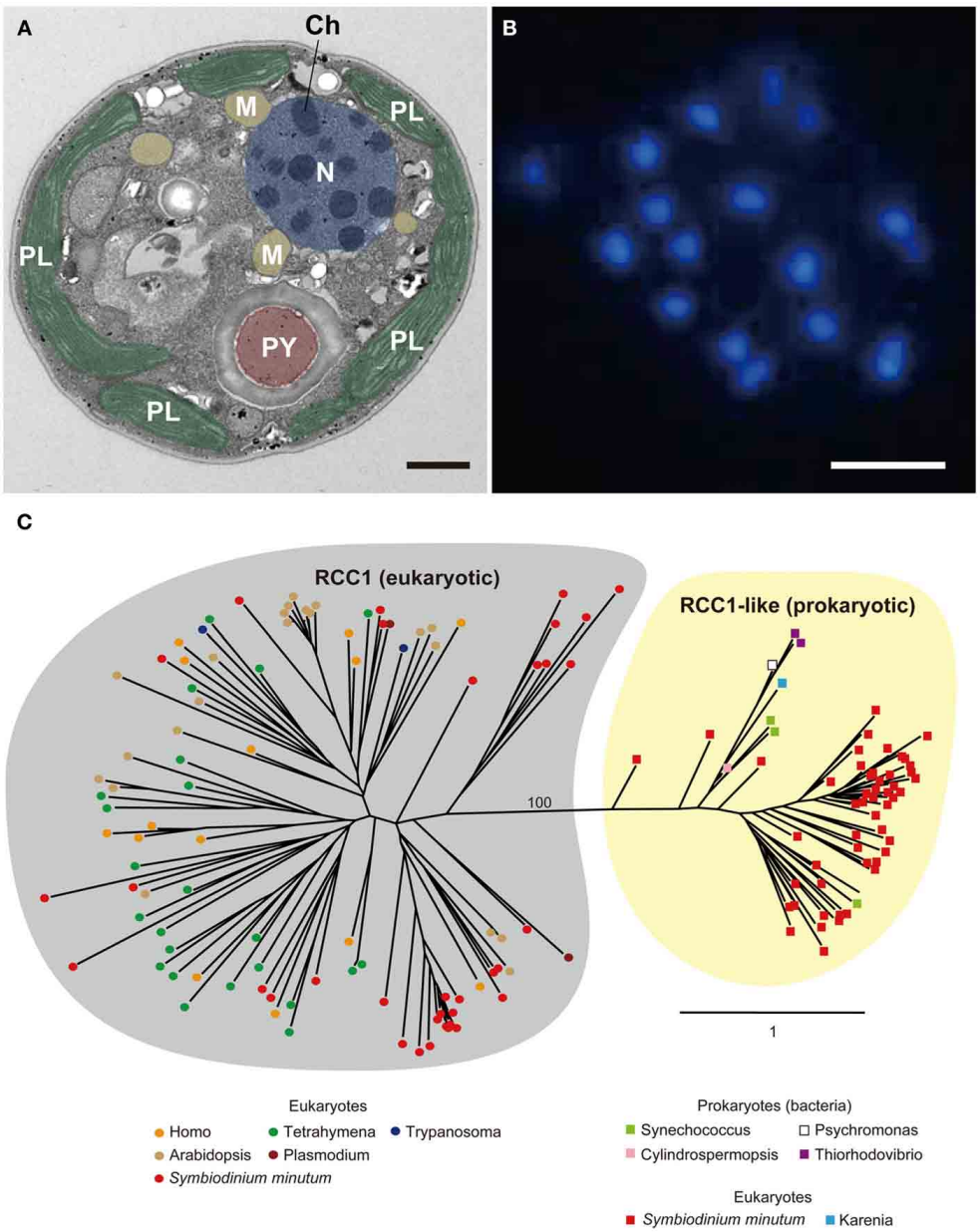


FIGURE 9 | (A) Electron micrograph showing permanently condensed chromosomes (Ch) of *Symbiodinium minutum*. The nucleus (N) is shown in purple, plastids (PL) in green, mitochondria (M) in orange, and pyrenoid (PY) in brown. Scale bar, 1 μm. **(B)** DAPI staining of the nucleus showing permanently condensed chromosomes of *S. minutum*. Scale bar, 1 μm. **(C)** RCC1 proteins are eukaryotic proteins that bind to chromatin and play an important role in the regulation of gene expression. A maximum-likelihood phylogeny of 86 RCC1 family proteins encoded in the *S. minutum* genome is shown. The two distinct groupings of eukaryotic RCC1 proteins and prokaryotic RCC1-like proteins are supported by 100% bootstrap duplication. Bar indicates an amino acid substitution per site.

contrasts strongly with the AT-rich genomes of other alveolates, such as apicomplexans [*P. falciparum*, 19% GC (Gardner et al., 2002)] and ciliates [*Tetrahymena thermophile*, 22% GC (Eisen et al., 2006)], respectively.

Gene content of the dinoflagellate genome

Of 41,925 gene models, 20,983 (50%) encode proteins with known domains. One of the largest dinoflagellate protein families is the EF-hand family, a large family of calcium-binding proteins characterized by a helix-loop-helix structural domain. The second largest dinoflagellate family contains ankyrin repeats, one of the most common protein-protein interaction motifs in nature. When the *Symbiodinium* gene families are compared with those of other eukaryotes, *Symbiodinium* shares a considerable number of homologous genes with *Homo* and *Arabidopsis*, although ~46% of predicted proteins are novel or *Symbiodinium*-specific.

Specific gene expansion in the Symbiodinium genome

Dinoflagellates have been predicted to possess 38,000–87,000 protein-coding genes (Hou and Lin, 2009). The presence of a larger number of genes in the *S. minutum* genome (41,925) is likely caused by lineage-specific expansion of genes by duplication (Hou and Lin, 2009). Orthologous gene clustering analyses indicate that 1064 groups (10,912 genes) in the *Symbiodinium* genome have likely resulted from such events. One striking finding is that the regulator of chromosome condensation family protein (RCC1) is highly expanded (discussed below). Calcium channel and calmodulin families are also expanded. Because the largest domain was the EF-hand subgroup of calcium-binding proteins, Ca^{2+} metabolism is clearly of great importance in *Symbiodinium*.

Molecular basis of permanently condensed chromatin

As mentioned above, dinoflagellate nuclei are characterized by permanently condensed, liquid-crystalline chromosomes (**Figures 9A,B**), and dinoflagellate chromosomal organization is a fundamental issues that is still not fully understood (Lin, 2011). In eukaryotes, histone proteins are involved in chromatin modulation, whereas in prokaryotes, histone-like proteins serve this function. The *S. minutum* genome contains both eukaryotic histone genes and prokaryotic histone-like genes, although orthologs of histone H1 are not found in the genome (Shoguchi et al., 2013a). All four core-histone genes (H2A, H2B, H3, and H4) are duplicated. In addition, there are 15 histone-like proteins similar to those found in bacteria.

In addition to enlargement of the genome, a dinoflagellate, *Hermatodinium* sp., gains a novel family of nucleoproteins from an algal virus, termed dinoflagellate/viral nucleoprotein (DVNP) (Gornik et al., 2012). The *Symbiodinium* genome contains 19 genes that appear homologous to DVNPs, suggesting a role for this type of protein in *Symbiodinium* chromosome structure.

The RCC1 proteins (RCC1 superfamily in eukaryotes and RCC1-like repeat proteins in both prokaryotes and eukaryotes) bind to chromatin and play an important role in the regulation of gene expression (Dasso, 1993). As mentioned above, genes for RCC1 have the third highest degree of expansion in the *Symbiodinium* genome, and a total of 189 genes are present in the *Symbiodinium* genome (Shoguchi et al., 2013a). When 86 of these proteins are used for molecular phylogenic analyses, two distinct clusters become evident. One, with 34 *Symbiodinium* proteins consists of those orthologous to eukaryotes, including alveolates, plants, and animals (**Figure 9C**, left), whereas the other includes 52 proteins with similarities to prokaryotes, including cyanobacteria and proteobacteria (**Figure 9C**, right). This result potentially explains the characteristic architecture of dinoflagellate chromosomes, although the manner in which they interact with each other to establish and maintain the permanently condensed chromosomes remains to be studied.

Unique spliceosomal splicing

Although previous reports have suggested that introns are relatively uncommon in dinoflagellate genes (Okamoto et al., 2001; Hoppenrath and Leander, 2010), genes of *S. minutum* are highly intron-rich. 39,970 of the 41,925 genes (95%) are composed of multiple exons. The average number of exons per gene reaches 19.6, and some genes contain more than 200 introns (Shoguchi et al., 2013a). In addition, spliceosomal introns of *Symbiodinium* are unique among eukaryotic genomes. In other eukaryotes, introns are excised under the GT-AG rule, wherein GT and AG are used as recognition nucleotides at 5′ and 3′ splice sites, respectively, (**Figure 10**). In contrast, *Symbiodinium* uses GC and GA at the 5′ donor splice site, in addition to GT (**Figure 10**). GC usage frequency is nearly equal to that of GT. The presence of these 5′ splice sites provides the first evidence in eukaryotes that the majority of mRNA splicing does not always follow the GT-AG rule. Another feature of *Symbiodinium* splicing is that the 3′ acceptor splice site, AG, is frequently followed by the nucleotide G (**Figure 10**), although a similar phenomenon is known in human minor alternative splice sites (Thanaraj and Clark, 2001).

Key steps in RNA splicing are performed by spliceosomes, acting in concert with five small nuclear RNA molecules (snRNAs; *U1*, *U2*, *U4*, *U5*, and *U6*). The five major snRNAs recognize nucleotide sequences that specify where splicing is to occur, and they participate in spliceosome chemistry (Rogozin et al., 2012). In the *Plasmodium* and *Tetrahymena* genomes, snRNAs are scattered throughout the genome, whereas in metazoans and green plants, two different types of the five major snRNAs are sometimes tandemly aligned (Wang and Brendel, 2004; Marz et al., 2008). In contrast, in the *Symbiodinium* genome, all five snRNAs, *U1, U2, U4, U5*, and *U6* occur in a cluster, in addition to other snRNAs scattered across about 70 locations. This is the first discovery of an snRNA gene cluster in a eukaryote genome. It has been reported that *trans*-splicing of messenger RNAs is common in dinoflagellates (Lin, 2011; Wisecaver and Hackett, 2011). The *Symbiodinium* genome contains spliced-leader (SL) genes with a conserved SL sequence.

Unique arrangement of genes in the genome

The *Symbiodinium* genome is also unique in the context of gene arrangement (Shoguchi et al., 2013a). In contrast to the random arrangement of protein-coding genes in the genomes of *Tetrahymena*, *Plasmodium*, *Arabidopsis,* and *Homo,* those of the *Symbiodinium* and *Trypanosoma* genomes show a clear tendency for tandem and unidirectional gene alignment. The grade of change in gene direction was searched using a 10-gene sliding window (**Figure 11**). Graphs of these data for *Plasmodium*, *Tetrahymena*, *Arabidopsis*, and *Homo* show a peak between 4 and 5 changes in orientation, indicating the frequency of strand switch regions (SSRs) between genes in head-to-head or tail-to-tail orientations (**Figure 11**). In contrast, *Symbiodinium* and *Trypanosoma* show a cluster (**Figure 11**). This indicates a strong tendency for tandem alignment of genes or clustering of unidirectionally aligned genes in the *Symbiodinium* and *Trypanosoma* genomes.

Genes involved in the basic transcriptional machinery

Although the *S. minutum* genome is unique in regard to permanently condensed chromosomes, spliceosomal splicing, and unidirectionally aligned genes, the genome contains highly conserved

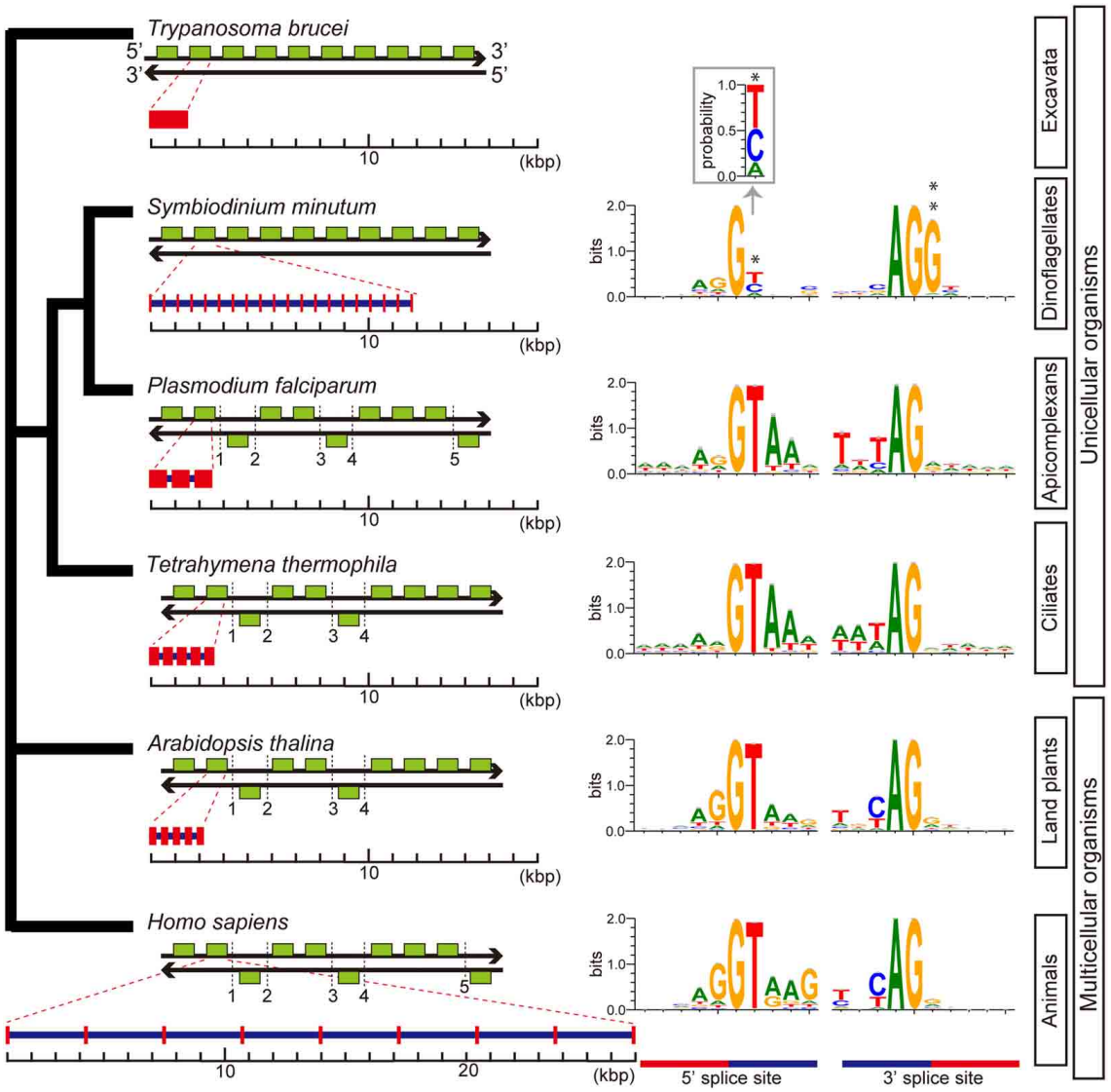


FIGURE 10 | Architecture of genes and splice site motifs in the nuclear genomes of representative eukaryotes and alveolates. Green boxes indicate typical patterns of 10-gene arrangements with the number of strand switch regions (SSRs), although the SSRs shown here are not always typical. Patterns are based on the analyses shown in **Figure 6**. Gene architecture shows average gene lengths (exons in red and introns in blue) with the average intron number per gene. The sequence motif of the splice site is illustrated using WebLogo. Only two genes with spliceosomal introns in *Trypanosoma brucei* have been reported, but the motif was not shown. The unusual gene organization on the same strand of DNA shows similarities between *Symbiodinium* and *Trypanosoma*. Additionally, analyses of intron-richness and the weakness of 5′ splice site signals (asterisk) indicate that *Symbiodinium* has the most unusual genome organization found in a eukaryote genome to date. The probability of position 2 at the 5′ splice site is shown in inset. A double asterisk shows G conserved at the 3′ splice site.

basic transcriptional machinery components, including RNA polymerase I, II, and III, basal transcription factors, such as TFIID and TATA-binding protein (TBP), and transcription elongation factors (Shoguchi et al., 2013a). In contrast, the genome contains a few sequence-specific transcription factors, including 19 gene models with AP2 domain(s), 15 models with HMG-box domain(s), eight models with zf-C2H2 domain(s), and others. These results suggest constant, steady transcription of *Symbiodinium* genes with fewer genes under sequence-specific transcriptional control.

CHLOROPLAST (PLASTID) GENOME

Chloroplasts (plastids) are common photosynthetic organelles in eukaryotic algae and land plants. Plastids first may have arisen when non-photosynthetic eukaryotic hosts acquired cyanobacterial endosymbionts by a process termed "primary endosymbiosis" (Howe et al., 2008; Keeling, 2010). Other non-photosynthetic eukaryotes may have subsequently acquired endosymbionts from photosynthetic eukaryotes to create secondary plastids (Howe et al., 2008; Keeling, 2010). In some lineages including dinoflagellates, secondary plastids may have been lost and replaced with

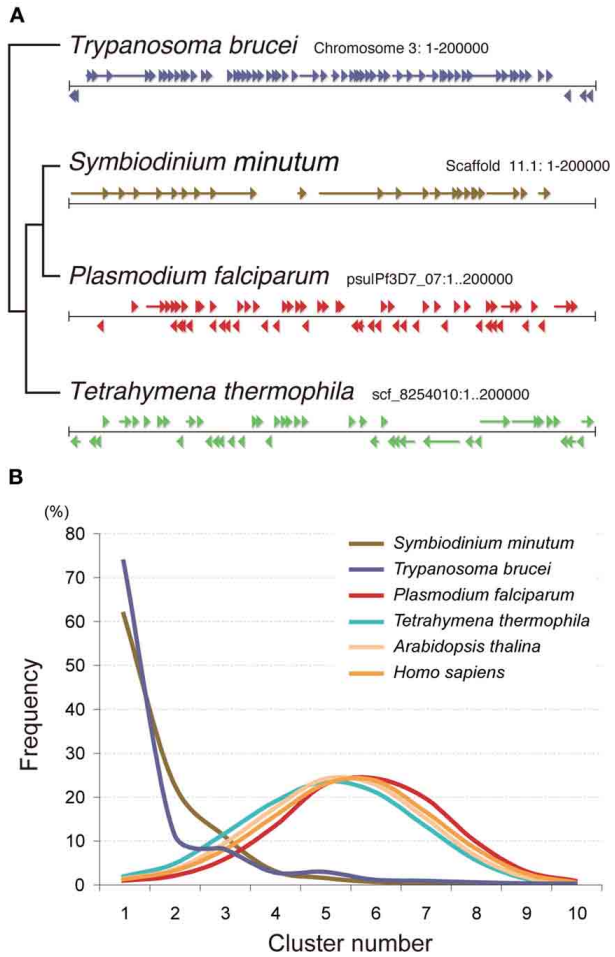


FIGURE 11 | Nuclear gene arrangement in the *Symbiodinium minutum* genome. (A) Examples of gene arrangement in the 200-kbp nuclear genome are shown and compared among *Symbiodinium minutum* (dinoflagellate), *Trypanosoma brucei* (euglenozoan), *Plasmodium falciparum* (apicomplexa), and *Tetrahymena thermophila* (ciliate). In contrast to the genomes of *Tetrahymena*, *Plasmodium*, and *Acropora,* that show a random arrangement of protein-coding genes (arrowheads and arrows), the genomes of *S. minutum* and *Trypanosoma* are arranged into a large directional gene cluster in a head-to-tail orientation. **(B)** A search of the directional gene cluster using a 10-gene window shows the strong tendency toward unidirectional alignment of genes in the *S. minutum* and *Trypanosoma* genomes. Each line represents a frequency histogram for changes in the gene orientation between successive genes in the genome. The X-axis represents the number of orientation changes as one moves through windows of 10 genes. For examples as indicating random orientation, the poisson distributions with $\mu = 4.5$ (average) and 0.2 are shown.

secondary endosymbiotic plastids or other primary endosymbiotic plastids, resulting in tertiary plastids (Allen et al., 2011).

Evolutionary changes in plastid genomes in alveolates are dramatic. Ciliates lost plastids and became heterotrophic, while parasitic apicomplexans retain unpigmented plastid remnants termed apicoplasts. On the other hand, two species closely related to apicomplexans, *Chromera velia* and *Vitrella brassicaformis*, are photosynthetic. Their plastid genomes retain ancestral characteristics of both apicomplexan and dinoflagellate plastids and probably share a common red algal endosymbiont (Janouskovec et al., 2010). Interestingly, rapidly evolving dinoflagellate plastids show a great variety of reduced stages. Their gene content has been dramatically diminished by large-scale transfer of genes to the nucleus, leaving only 12–17 genes in the plastids (Howe et al., 2008). Conventional plastid genomes have all genes physically linked in one molecule, typically 120–200 kb in size (Keeling, 2010), while dinoflagellate plastid genes reside on small plasmids of 2.2–6 kb, termed "minicircles" (Zhang et al., 1999), containing a few genes and a core, non-coding region, which is conserved within species and plays a regulatory role (Zhang et al., 2002; Leung and Wong, 2009; Wisecaver and Hackett, 2011). Moreover, a number of unusual post-transcriptional RNA modifications, including the addition of 3′ terminal poly(U)tracts, occur in the ancestral chloroplasts of dinoflagellates. Extensive RNA editing occurs in some dinoflagellates (Zauner et al., 2004; Wang and Morse, 2006; Dang and Green, 2009), employing diverse editing types that have not been observed in mammals and plants. This leads to speculation about the functional connection between poly(U)tailing and RNA editing in dinoflagellate plastid transcripts (Dang and Green, 2009).

In *S. minutum*, 95 of 109 plastid-associated genes have been transferred to the nuclear genome and subsequently expanded by gene duplication (Mungpakdee et al., 2014). Only 14 genes remain in plastids, as DNA minicircles. Each *Symbiodinium* minicircle (1.8–3.3 kb) contains one gene and a conserved non-coding region containing putative promoters and RNA-binding sites. Nine types of RNA editing, including a novel G/U type, are evident in minicircle transcripts, but not in genes transferred to the nucleus. In contrast to DNA editing sites in dinoflagellate mitochondria, which tend to be highly conserved across all taxa, editing sites employed in DNA minicircles are highly variable from species to species. Editing is crucial for core photosystem protein function. It restores evolutionarily conserved amino acids and increases peptidyl hydropathy. RNA editing is also likely to increase protein plasticity necessary to initiate photosystem complex assembly.

MITOCHONDRIAL GENOME

In most metazoans, mitochondrial genomes are 13–20-kb, compact, circular molecules, containing 12–13 proteins, 24–25 tRNAs, and 2 rRNAs. As in the case of plastid genomes, mitochondrial genomes also dramatically changed during evolution. Ciliates (*Tetrahymena* and *Ichthyophthirius*) have linearly mapped mitochondrial genomes of 43 kb with a normal gene number (Burger et al., 2000), while only 3 protein-coding genes and fragmented rRNAs organized as part of linear repeats of about 6–7 kbp are found in parasitic apicomplexans (*Plasmodium, Babesia*, and *Theileria*) (Hikosaka et al., 2012). Gene content of dinoflagellate mitochondrial genomes is comparable to that of apicomplexans (Slamovits et al., 2007), but with highly fragmented and rearranged genome structure (Waller and Jackson, 2009).

A 49-kmer assembly of only high coverage (>100) Illumina paired-end reads of a dinoflagellate, *S. minutum*, revealed two candidate mitochondrial scaffolds, two linear DNAs (19,577 and 291,368 bp) (Mungpakdee et al., unpublished data). Blast and

transcriptome mapping show that one contains only *cox1* and the other *cob*, *cox3,* and 6 fragmented of large subunit (LSU) rRNA genes. Fragments of small subunit (SSU) rRNA and tRNA genes are not found in the *Symbiodinium* mitochondrial genome. The evolution of the mitochondrial genome in *Symbiodinium*, as well as in other dinoflagellates requires further investigation to reach some consensus.

CONCLUSION

Genomic information is essential for future studies of molecular and cellular mechanisms underlying the establishment, maintenance, and breakdown of obligate endosymbiosis of corals with photosynthetic dinoflagellates *Symbiodinium*. In general, the coral genome is unique in that frequent horizontal gene transfer is evident in UV-protection genes. In addition, *Symbiodinium* is one of diverse dinoflagellates in regard to nuclear, plastid, and mitochondrial genomes. At present, many questions about endosymbiosis remain to be answered, but genomic information will greatly facilitate future studies of coral-dinoflagellate endosymbiosis.

ACKNOWLEDGMENTS

Our genome project of both the coral and *Symbiodinium* was supported by a Grant-in-Aids from MEXT (No. 23128515 to Eiichi Shoguchi) and JSPS (No. 24241071 to Nori Satoh) of Japan, and OIST internal fund. We thank all members of our Unit and the DNA Sequence Section of OIST for their enormous help in the project, and Dr. Steven Aird for his help in preparing the manuscript.

REFERENCES

Allen, J. F., de Paula, W. B., Puthiyaveetil, S., and Nield, J. (2011). A structural phylogenetic map for chloroplast photosynthesis. *Trends Plant Sci.* 16, 645–655. doi: 10.1016/j.tplants.2011.10.004

Atchley, W. R., and Fitch, W. M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 94, 5172–5176. doi: 10.1073/pnas.94.10.5172

Bachvaroff, T. R., and Place, A. R. (2008). From stop to start: tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. *PLoS ONE* 3:e2929. doi: 10.1371/journal.pone.0002929

Balskus, E. P., and Walsh, C. T. (2008). Investigating the initial steps in the biosynthesis of cyanobacterial sunscreen scytonemin. *J. Am. Chem. Soc.* 130, 15260–15261. doi: 10.1021/ja807192u

Balskus, E. P., and Walsh, C. T. (2010). The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science* 329, 1653–1656. doi: 10.1126/science.1193637

Barshis, D., Ladner, J., Oliver, T., and Palumbi, S. R. (2014). Lineage specific transcriptional profiles of *Symbiodinium* spp. unaltered by heat stress in a coral host. *Mol. Biol. Evol.* 31, 1343–1352. doi: 10.1093/molbev/msu107

Barshis, D. J., Ladner, J. T., Oliver, T. A., Seneca, F. O., Traylor-Knowles, N., and Palumbi, S. R. (2013). Genomic basis for coral resilience to climate change. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1387–1392. doi: 10.1073/pnas.1210224110

Baumgarten, S., Bayer, T., Aranda, M., Liew, Y. J., Carr, A., Micklem, G., et al. (2013). Integrating microRNA and mRNA expression profiling in *Symbiodinium microadriaticum*, a dinoflagellate symbiont of reef-building corals. *BMC Genomics* 14:704. doi: 10.1186/1471-2164-14-704

Bayer, T., Aranda, M., Sunagawa, S., Yum, L. K., Desalvo, M. K., Lindquist, E., et al. (2012). *Symbiodinium* transcriptomes: genome insights into the dinoflagellate symbionts of reef-building corals. *PLoS ONE* 7:e35269. doi: 10.1371/journal.pone.0035269

Berghammer, H., Hayward, D., Harrison, P., and Miller, D. J. (1996). Nucleotide sequence of ub52 from the cnidarian Acropora millepora reveals high evolutionary conservation. *Gene* 178, 195–197. doi: 10.1016/0378-1119(96)00353-8

Bou-Abdallah, F., Chasteen, N. D., and Lesser, M. P. (2006). Quenching of superoxide radicals by green fluorescent protein. *Biochim. Biophys. Acta* 1760, 1690–1695. doi: 10.1016/j.bbagen.2006.08.014

Bouligand, Y., and Norris, V. (2001). Chromosome separation and segregation in dinoflagellates and bacteria may depend on liquid crystalline states. *Biochimie* 83, 187–192. doi: 10.1016/S0300-9084(00)01211-6

Burger, G., Zhu, Y., Littlejohn, T. G., Greenwood, S. J., Schnare, M. N., Lang, B. F., et al. (2000). Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *J. Mol. Biol.* 297, 365–380. doi: 10.1006/jmbi.2000.3529

Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjaeveland, A., Nikolaev, S. I., Jakobsen, K. S., et al. (2007). Phylogenomics reshuffles the eukaryotic supergroups. *PLoS ONE* 2:e790. doi: 10.1371/journal.pone.0000790

Chapman, J. A., Kirkness, E. F., Simakov, O., Hampson, S. E., Mitros, T., Weinmaier, T., et al. (2010). The dynamic genome of Hydra. *Nature* 464, 592–596. doi: 10.1038/nature08830

Chourrout, D., Delsuc, F., Chourrout, P., Edvardsen, R. B., Rentzsch, F., Renfer, E., et al. (2006). Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* 442, 684–687. doi: 10.1038/nature04863

Dang, Y., and Green, B. R. (2009). Substitutional editing of *Heterocapsa triquetra* chloroplast transcripts and a folding model for its divergent chloroplast 16S rRNA. *Gene* 442, 73–80. doi: 10.1016/j.gene.2009.04.006

Dasso, M. (1993). RCC1 in the cell cycle: the regulator of chromosome condensation takes on new roles. *Trends Biochem. Sci.* 18, 96–101. doi: 10.1016/0968-0004(93)90161-F

Denker, E., Bapteste, E., Le Guyader, H., Manuel, M., and Rabet, N. (2008). Horizontal gene transfer and the evolution of cnidarian stinging cells. *Curr. Biol.* 18, R858–R859. doi: 10.1016/j.cub.2008.07.031

Drake, J. L., Mass, T., Haramaty, L., Zelzion, E., Bhattacharya, D., and Falkowski, P. G. (2013). Proteomic analysis of skeletal organic matrix from stony coral Stylophora pistillata. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3788–3789. doi: 10.1073/pnas.1301419110

DuBuc, T. Q., Ryan, J. F., Shinzato, C., Satoh, N., and Martindale, M. Q. (2012). Coral comparative genomics reveal expanded Hox cluster in the cnidarian-bilaterian ancestor. *Integr. Comp. Biol.* 52, 835–841. doi: 10.1093/icb/ics098

Eisen, J. A., Coyne, R. S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J. R., et al. (2006). Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol.* 4:e286. doi: 10.1371/journal.pbio.0040286

Erdner, D. L., and Anderson, D. M. (2006). Global transcriptional profiling of the toxic dinoflagellate *Alexandrium fundyense* using massively parallel signature sequencing. *BMC Genomics* 7:88. doi: 10.1186/1471-2164-7-88

Field, S. F., Bulina, M. Y., Kelmanson, I. V., Bielawski, J. P., and Matz, M. V. (2006). Adaptive evolution of multicolored fluorescent proteins in reefbuilding corals. *J. Mol. Evol.* 62, 332–339. doi: 10.1007/s00239-005-0129-9

Fukuda, I., Ooki, S., Fujita, T., Murayama, E., Nagasawa, H., Isa, Y., et al. (2003). Molecular cloning of a cDNA encoding a soluble protein in the coral exoskeleton. *Biochem. Biophys. Res. Commun.* 304, 11–17. doi: 10.1016/S0006-291X(03)00527-8

Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511. doi: 10.1038/nature01097

Gleason, D. F., and Wellington, G. M. (1993). Ultraviolet-radiation and coral bleaching. *Nature* 365, 836–838. doi: 10.1038/365836a0

Gornik, S. G., Ford, K. L., Mulhern, T. D., Bacic, A., McFadden, G. I., and Waller, R. F. (2012). Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates. *Curr. Biol.* 22, 2303–2312. doi: 10.1016/j.cub.2012.10.036

Graham, L., and Wilcox, L. (2000). *Algae.* Upper Saddle River, NJ: Prentice-Hall.

Gyoja, F., Kawashima, T., and Satoh, N. (2012). A genomewide survey of bHLH transcription factors in the coral *Acropora digitifera* identifies three novel orthologous families, pearl, amber, and peridot. *Dev. Genes Evol.* 222, 63–76. doi: 10.1007/s00427-012-0388-6

Hamada, M., Shoguchi, E., Shinzato, C., Kawashima, T., Miller, D. J., and Satoh, N. (2013). The complex NOD-like receptor repertoire of the coral *Acropora digitifera* includes novel domain combinations. *Mol. Biol. Evol.* 30, 167–176. doi: 10.1093/molbev/mss213

Hemmrich, G., and Bosch, T. C. (2008). Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. *Bioessays* 30, 1010–1018. doi: 10.1002/bies.20813

Henderson, J. N., and Remington, S. J. (2005). Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from *Anemonia majano*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12712–12717. doi: 10.1073/pnas.0502250102

Hikosaka, K., Tsuji, N., Watanabe, Y., Kishine, H., Horii, T., Igarashi, I., et al. (2012). Novel type of linear mitochondrial genomes with dual flip-flop inversion system in apicomplexan parasites, *Babesia microti* and *Babesia rodhaini*. *BMC Genomics* 13:622. doi: 10.1186/1471-2164-13-622

Hoppenrath, M., and Leander, B. S. (2010). Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences. *PLoS ONE* 5:e13220. doi: 10.1371/journal.pone.0013220

Hou, Y., and Lin, S. (2009). Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes. *PLoS ONE* 4:e6978. doi: 10.1371/journal.pone.0006978

Howe, C. J., Nisbet, R. E., and Barbrook, A. C. (2008). The remarkable chloroplast genome of dinoflagellates. *J. Exp. Bot.* 59, 1035–1045. doi: 10.1093/jxb/erm292

Jackson, D. J., Macis, L., Reitner, J., Degnan, B. M., and Worheide, G. (2007). Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis. *Science* 316, 1893–1895. doi: 10.1126/science.1141560

Janouskovec, J., Horak, A., Obornik, M., Lukes, J., and Keeling, P. J. (2010). A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10949–10954. doi: 10.1073/pnas.1003335107

Jones, S. (2004). An overview of the basic helix-loop-helix proteins. *Genome Biol.* 5, 226. doi: 10.1186/gb-2004-5-6-226

Karako-Lampert, S., Zoccola, D., Salmon-Divon, M., Katzenellenbogen, M., Tambutte, S., Bertucci, A., et al. (2014). Transcriptome analysis of the scleractinian coral *Stylophora pistillata*. *PLoS ONE* 9:e88615. doi: 10.1371/journal.pone.0088615

Keeling, P. J. (2009). Functional and ecological impacts of horizontal gene transfer in eukaryotes. *Curr. Opin. Genet. Dev.* 19, 613–619. doi: 10.1016/j.gde.2009.10.001

Keeling, P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365, 729–748. doi: 10.1098/rstb.2009.0103

Kelmanson, I. V., and Matz, M. V. (2003). Molecular basis and evolutionary origins of color diversity in great star coral *Montastraea cavernosa (Scleractinia: Faviida)*. *Mol. Biol. Evol.* 20, 1125–1133. doi: 10.1093/molbev/msg130

Kenkel, C. D., Meyer, E., and Matz, M. V. (2013). Gene expression under chronic heat stress in population of the mustard hill coral (*Porites astreoides*) from different thermal environments. *Mol. Ecol.* 16, 4322–4334. doi: 10.1111/mec.12390

Kornberg, H. L., and Beevers, H. (1957). Mechanism of conversion of fat to carbohydrate in castor beans. *Nature* 180, 35–36. doi: 10.1038/180035a0

Kortschak, R. D., Samuel, G., Saint, R., and Miller, D. J. (2003). EST analysis of the cnidarian Acropora millepora reveals extensive gene loss and rapid sequence divergence in the model invertebrates. *Curr. Biol.* 13, 2190–2195 doi: 10.1016/j.cub.2003.11.030

Koyanagi, R., Takeuchi, T., Hisata, K., Gyoja, F., Shoguchi, E., Satoh, N., et al. (2013). MarinegenomicsDB: an integrated genome viewer for community-based annotation of genomes. *Zoolog. Sci.* 30, 797–800. doi: 10.2108/zsj.30.797

Ladner, J. T., Barshis, D. J., and Palumbi, S. R. (2012). Protein evolution in two co-occurring types of *Symbiodinium*: an exploration into the genetic basis of thermal tolerance in *Symbiodinium* clade D. *BMC Evol. Biol.* 12:217. doi: 10.1186/1471-2148-12-217

Leggat, W., Hoegh-Guldberg, O., Dove, S., and Yellowlees, D. (2007). Analysis of an EST library from the dinoflagellate (*Symbiodinium* sp.) symbiont of reef-building corals. *J. Phycol.* 43, 1010–1021. doi: 10.1111/j.1529-8817.2007.00387.x

Leung, S. K., and Wong, J. T. (2009). The replication of plastid minicircles involves rolling circle intermediates. *Nucleic Acids Res.* 37, 1991–2002. doi: 10.1093/nar/gkp063

Levy, O., Kaniewska, P., Alon, S., Eisenberg, E., Karako-Lampert, S., Bay, L. K., et al. (2011). Complex diel cycles of gene expression in coral-algal symbiosis. *Science* 331, 175. doi: 10.1126/science.1196419

Libro, S., Kaluziak, S. T., and Vollmer, S. V. (2013). RNA-seq profiles of immune related genes in the staghorn coral *Acropora cervicornis* infected with white band disease. *PLoS ONE* 8:e81821. doi: 10.1371/journal.pone.0081821

Lidie, K. B., and van Dolah, F. M. (2007). Spliced leader RNA-mediated trans-splicing in a dinoflagellate, *Karenia brevis*. *J. Eukaryot. Microbiol.* 54, 427–435. doi: 10.1111/j.1550-7408.2007.00282.x

Lin, S. J. (2011). Genomic understanding of dinoflagellates. *Res. Microbiol.* 162, 551–569. doi: 10.1016/j.resmic.2011.04.006

Liu, F. Z., Thatcher, J. D., Barral, J. M., and Epstein, H. F. (1995). Bifunctional glyoxylate cycle protein of caenorhabditis-elegans—a developmentally-regulated protein of intestine and muscle. *Dev. Biol.* 169, 399–414. doi: 10.1006/dbio.1995.1156

Lukyanov, K. A., Chudakov, D. M., Fradkov, A. F., Labas, Y. A., Matz, M. V., and Lukyanov, S. (2006). Discovery and properties of GFP-like proteins from non-bioluminescent anthozoa. *Methods Biochem. Anal.* 47, 121–138. doi: 10.1002/0471739499.ch6

Marz, M., Kirsten, T., and Stadler, P. F. (2008). Evolution of spliceosomal snRNA genes in metazoan animals. *J. Mol. Evol.* 67, 594–607. doi: 10.1007/s00239-008-9149-6

Massari, M. E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* 20, 429–440. doi: 10.1128/MCB.20.2.429-440.2000

Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., et al. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat. Biotechnol.* 17, 969–973. doi: 10.1038/13657

Matz, M. V., Marshall, N. J., and Vorobyev, M. (2006). Are corals colorful? *Photochem. Photobiol.* 82, 345–350. doi: 10.1562/2005-08-18-RA-653

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A., and Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* 308, 428–433. doi: 10.1038/308428a0

Mehr, S. F. P., DeSalle, R., Kao, H. T., Narechania, A., Han, Z., Tchernov, D., et al. (2013). Transcriptome deep-sequencing and clustering of expressed isoforms from Favia corals. *BMC Genomics* 14:546. doi: 10.1186/1471-2164-14-546

Miller, D. J., Hayward, D. C., Reece-Hoyes, J. S., Scholten, I., Catmull, J., Gehring, W. J., et al. (2000). Pax gene diversity in the basal cnidarian *Acropora millepora* (Cnidaria, Anthozoa): implications for the evolution of the Pax gene family. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4475–4480. doi: 10.1073/pnas.97.9.4475

Miller, D. J., Hemmrich, G., Ball, E. E., Hayward, D. C., Khalturin, K., Funayama, N., et al. (2007). The innate immune repertoire in cnidaria–ancestral complexity and stochastic gene loss. *Genome Biol.* 8:R59. doi: 10.1186/gb-2007-8-4-r59

Moustafa, A., Evans, A. N., Kulis, D. M., Hackett, J. D., Erdner, D. L., Anderson, D. M., et al. (2010). Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. *PLoS ONE* 5:e9688. doi: 10.1371/journal.pone.0009688

Moya, A., Huisman, L., Ball, E. E., Hayward, D. C., Grasso, L. C., Chua, C. M., et al. (2012). Whole transcriptome analysis of the coral *Acropora millepora* reveals complex responses to CO(2)-driven acidification during the initiation of calcification. *Mol. Ecol.* 21, 2440–2454. doi: 10.1111/j.1365-294X.2012.05554.x

Mungpakdee, S., Shinzato, C., Takeuchi, T., Kawashima, T., Koyanagi, R., Hisata, K., et al. (2014). Massive gene transfer and extensive RNA editing of a symbiotic dinoflagellate plastid genome. *Genome Biol. Evol.* 6, 1408–1422. doi: 10.1093/gbe/evu109

Okamoto, O. K., Liu, L., Robertson, D. L., and Hastings, J. W. (2001). Members of a dinoflagellate luciferase gene family differ in synonymous substitution rates. *Biochemistry* 40, 15862–15868. doi: 10.1021/bi011651q

Palmer, C. V., Modi, C. K., and Mydlarz, L. D. (2009). Coral fluorescent proteins as antioxidants. *PLoS ONE* 4:e7298. doi: 10.1371/journal.pone.0007298

Pochon, X., Montoya-Burgos, J. I., Stadelmann, B., and Pawlowski, J. (2006). Molecular phylogeny, evolutionary rates, and divergence timing of the symbiotic dinofagellate genus Symbiodimium. *Mol. Phylogenet. Evol.* 38, 20–30. doi: 10.1016/j.ympev.2005.04.028

Polato, N. R., Vera, J. C., and Baums, I. B. (2011). Gene discovery in the threatened elkhorn coral: 454 sequencing of the *Acropora palmata* transcriptome. *PLoS ONE* 6:e28634. doi: 10.1371/journal.pone.0028634

Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., et al. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317, 86–94. doi: 10.1126/science.1139158

Ramos-Silva, P., Kaandorp, J., Huisman, L., Marie, B., Zanella-Cleon, I., Guichard, N., et al. (2013). The skeletal proteomes of the coral Acropora millepora: the evolution of calcification by co-option and domain shaffling. *Mol. Biol. Evol.* 30, 2099–2112. doi: 10.1093/molbev/mst109

Rastogi, R. P., Richa, Sinha, R. P., Singh, S. P., and Hader, D. P. (2010). Photoprotective compounds from marine organisms. *J. Ind. Microbiol. Biotechnol.* 37, 537–558. doi: 10.1007/s10295-010-0718-5

Reef, R., Dunn, S., Levy, O., Dove, S., Shemesh, E., Brickner, I., et al. (2009). Photoreactivation is the main repair pathway for UV-induced DNA damage in coral planulae. *J. Exp. Biol.* 212, 2760–2766. doi: 10.1242/jeb.031286

Rogozin, I. B., Carmel, L., Csuros, M., and Koonin, E. V. (2012). Origin and evolution of spliceosomal introns. *Biol. Direct* 7, 11. doi: 10.1186/1745-6150-7-11

Ryan, J. F., Mazza, M. E., Pang, K., Matus, D. Q., Baxevanis, A. D., Martindale, M. Q., et al. (2007). Pre-bilaterian origins of the Hox cluster and the Hox code: evidence from the sea anemone, *Nematostella vectensis*. *PLoS ONE* 2:e153. doi: 10.1371/journal.pone.0000153

Salih, A., Larkum, A., Cox, G., Kuhl, M., and Hoegh-Guldberg, O. (2000). Fluorescent pigments in corals are photoprotective. *Nature* 408, 850–853. doi: 10.1038/35048564

Sarashina, I., and Endo, K. (2006). Skeletal matrix proteins of invertebrate animals: comparative analysis of their amino acid sequences. *Paleontol. Res.* 10, 311–336. doi: 10.2517/prpsj.10.311

Schwarz, J. A., Brokstein, P. B., Voolstra, C., Terry, A. Y., Miller, D. J., Szmant, A. M., et al. (2008). Coral life history and symbiosis: functional genomic resources for two reef building Caribbean corals, *Acropora palmata* and *Montastraea faveolata*. *BMC Genomics* 9:97. doi: 10.1186/1471-2164-9-97

Shick, J. M., Romaine-Lioud, S., Ferrier-Pages, C., and Gattuso, J. P. (1999). Ultraviolet-B radiation stimulates shikimate pathway-dependent accumulation of mycosporine-like amino acids in the coral *Stylophora pistillata* despite decreases in its population of symbiotic dinoflagellates. *Limnol. Oceanogr.* 44, 1667–1682.

Shinzato, C., Inoue, M., and Kusakabe, M. (2014). A snapshot of a coral "holobiont": a transcriptome assembly of the scleractinian coral, *Porites*, captures a wide variety of genes from both the host and symbiotic zooxanthellae. *PLoS ONE* 9:e85182. doi: 10.1371/journal.pone.0085182

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., et al. (2011). Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476, 320–323. doi: 10.1038/nature10249

Shinzato, C., Shoguchi, E., Tanaka, M., and Satoh, N. (2012). Fluorescent protein candidate genes in the coral *Acropora digitifera* genome. *Zoolog. Sci.* 29, 260–264. doi: 10.2108/zsj.29.260

Shoguchi, E., Shinzato, C., Kawashima, T., Gyoja, F., Mungpakdee, S., Koyanagi, R., et al. (2013a). Draft assembly of the Symbiodinium minutum nuclear genome reveals dinoflagellate gene structure. *Curr. Biol.* 23, 1399–1408. doi: 10.1016/j.cub.2013.05.062

Shoguchi, E., Tanaka, M., Shinzato, C., Kawashima, T., and Satoh, N. (2013b). A genome-wide survey of photoreceptor and circadian genes in the coral, Acropora digitifera. *Gene* 515, 426–431. doi: 10.1016/j.gene.2012.12.038

Shoguchi, E., Tanaka, M., Takeuchi, T., Shinzato, C., and Satoh, N. (2013c). Probing a coral genome for components of the photoprotective scytonemin biosynthetic pathway and the 2-aminoethylphosphonate pathway. *Mar. Drugs* 11, 559–570. doi: 10.3390/md11020559

Slamovits, C. H., Saldarriaga, J. F., Larocque, A., and Keeling, P. J. (2007). The highly reduced and fragmented mitochondrial genome of the early-branching dinoflagellate *Oxyrrhis marina* shares characteristics with both apicomplexan and dinoflagellate mitochondrial genomes. *J. Mol. Biol.* 372, 356–368. doi: 10.1016/j.jmb.2007.06.085

Soule, T., Palmer, K., Gao, Q., Potrafka, R. M., Stout, V., and Garcia-Pichel, F. (2009). A comparative genomics approach to understanding the biosynthesis of the sunscreen scytonemin in cyanobacteria. *BMC Genomics* 10:336. doi: 10.1186/1471-2164-10-336

Soule, T., Stout, V., Swingley, W. D., Meeks, J. C., and Garcia-Pichel, F. (2007). Molecular genetics and genomic analysis of scytonemin biosynthesis in Nostoc punctiforme ATCC 29133. *J. Bacteriol.* 189, 4465–4472. doi: 10.1128/JB.01816-06

Stanley, G. D. Jr., and Fautin, D. G. (2001). Paleontology and evolution. The origins of modern corals. *Science* 291, 1913–1914. doi: 10.1126/science.1056632

Starcevic, A., Akthar, S., Dunlap, W. C., Shick, J. M., Hranueli, D., Cullum, J., et al. (2008). Enzymes of the shikimic acid pathway encoded in the genome of a basal metazoan, *Nematostella vectensis*, have microbial origins. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2533–2537. doi: 10.1073/pnas.0707388105

Sunagawa, S., Wilson, E. C., Thaler, M., Smith, M. L., Caruso, C., Pringle, J. R., et al. (2009). Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. *BMC Genomics* 10:258. doi: 10.1186/1471-2164-10-258

Technau, U., Rudd, S., Maxwell, P., Gordon, P. M., Saina, M., Grasso, L. C., et al. (2005). Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. *Trends Genet.* 21, 633–639. doi: 10.1016/j.tig.2005.09.007

Thanaraj, T. A., and Clark, F. (2001). Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. *Nucleic Acids Res.* 29, 2581–2593. doi: 10.1093/nar/29.12.2581

Traylor-Knowles, N., Granger, B. R., Lubinski, T. J., Parikh, J. R., Garamszegi, S., Xia, Y., et al. (2011). Production of a reference transcriptome and transcriptomic database (PocilloporaBase) for the cauliflower coral, *Pocillopora damicornis*. *BMC Genomics* 12:585. doi: 10.1186/1471-2164-12-585

Voolstra, C. R., Sunagawa, S., Schwarz, J. A., Coffroth, M. A., Yellowlees, D., Leggat, W., et al. (2009). Evolutionary analysis of orthologous cDNA sequences from cultured and symbiotic dinoflagellate symbionts of reef-building corals (Dinophyceae: *Symbiodinium*). *Comp. Biochem. Physiol. Part D Genomics Proteomics* 4, 67–74. doi: 10.1016/j.cbd.2008.11.001

Waller, R. F., and Jackson, C. J. (2009). Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *Bioessays* 31, 237–245. doi: 10.1002/bies.200800164

Wang, B. B., and Brendel, V. (2004). The ASRG database: identification and survey of *Arabidopsis thaliana* genes involved in pre-mRNA splicing. *Genome Biol.* 5:R102. doi: 10.1186/gb-2004-5-12-r102

Wang, Y., and Morse, D. (2006). Rampant polyuridylylation of plastid gene transcripts in the dinoflagellate Lingulodinium. *Nucleic Acids Res.* 34, 613–619. doi: 10.1093/nar/gkj438

Wilkinson, C. (2004). *Status of Coral Reefs of the World*. Townsville, QLD: Australian Institute of Marine Studies.

Wilson, R. J., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A., et al. (1996). Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* 261, 155–172. doi: 10.1006/jmbi.1996.0449

Wisecaver, J. H., and Hackett, J. D. (2011). Dinoflagellate genome evolution. *Annu. Rev. Microbiol.* 65, 369–387. doi: 10.1146/annurev-micro-090110-102841

Zauner, S., Greilinger, D., Laatsch, T., Kowallik, K. V., and Maier, U. G. (2004). Substitutional editing of transcripts from genes of cyanobacterial origin in the dinoflagellate *Ceratium horridum*. *FEBS Lett.* 577, 535–538. doi: 10.1016/j.febslet.2004.10.060

Zhang, H., Hou, Y., Miranda, L., Campbell, D. A., Sturm, N. R., Gaasterland, T., et al. (2007). Spliced leader RNA trans-splicing in dinoflagellates. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4618–4623. doi: 10.1073/pnas.0700258104

Zhang, H., Zhuang, Y., Gill, J., and Lin, S. (2013). Proof that dinoflagellate spliced leader (DinoSL) is a useful hook for fishing dinoflagellate transcripts from mixed microbial samples: *Symbiodinium kawagutii* as a case study. *Protist* 164, 510–527. doi: 10.1016/j.protis.2013.04.002

Zhang, Z., Cavalier-Smith, T., and Green, B. R. (2002). Evolution of dinoflagellate unigenic minicircles and the partially concerted divergence of their putative replicon origins. *Mol. Biol. Evol.* 19, 489–500. doi: 10.1093/oxfordjournals.molbev.a004104

Zhang, Z., Green, B. R., and Cavalier-Smith, T. (1999). Single gene circles in dinoflagellate chloroplast genomes. *Nature* 400, 155–159. doi: 10.1038/22099

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 April 2014; accepted: 17 June 2014; published online: 07 July 2014.
*Citation: Shinzato C, Mungpakdee S, Satoh N and Shoguchi E (2014) A genomic approach to coral-dinoflagellate symbiosis: studies of Acropora digitifera and Symbiodinium minutum. Front. Microbiol. **5**:336. doi: 10.3389/fmicb.2014.00336*
This article was submitted to Microbial Symbioses, a section of the journal Frontiers in Microbiology.