KEGG orthology-based annotation of the predicted proteome of Acropora digitifera: ZoophyteBase - an open access and searchable database of a coral genome

Walter C Dunlap1,2, Antonio Starcevic4, Damir Baranasic4, Janko Diminic4, Jurica Zucko4, Ranko Gacesa4, Madeleine JH van Oppen1, Daslav Hranueli4, John Cullum5 and Paul F Long2,3*

Abstract

Background: Contemporary coral reef research has firmly established that a genomic approach is urgently needed to better understand the effects of anthropogenic environmental stress and global climate change on coral holobiont interactions. Here we present KEGG orthology-based annotation of the complete genome sequence of the scleractinian coral Acropora digitifera and provide the first comprehensive view of the genome of a reef-building coral by applying advanced bioinformatics.

Description: Sequences from the KEGG database of protein function were used to construct hidden Markov models. These models were used to search the predicted proteome of A. digitifera to establish complete genomic annotation. The annotated dataset is published in ZoophyteBase, an open access format with different options for searching the data. A particularly useful feature is the ability to use a Google-like search engine that links query words to protein attributes. We present features of the annotation that underpin the molecular structure of key processes of coral physiology that include (1) regulatory proteins of symbiosis, (2) planula and early developmental proteins, (3) neural messengers, receptors and sensory proteins, (4) calcification and Ca²⁺-signalling proteins, (5) plant-derived proteins, (6) proteins of nitrogen metabolism, (7) DNA repair proteins, (8) stress response proteins, (9) antioxidant and redox-protective proteins, (10) proteins of cellular apoptosis, (11) microbial symbioses and pathogenicity proteins, (12) proteins of viral pathogenicity, (13) toxins and venom, (14) proteins of the chemical defensome and (15) coral epigenetics.

Conclusions: We advocate that providing annotation in an open-access searchable database available to the public domain will give an unprecedented foundation to interrogate the fundamental molecular structure and interactions of coral symbiosis and allow critical questions to be addressed at the genomic level based on combined aspects of evolutionary, developmental, metabolic, and environmental perspectives.

Keywords: Acropora digitifera, KEGG orthology, Database, Annotation, Proteome, Genome, Coral, Symbiosis, Cnidaria

* Correspondence: paul.long@kcl.ac.uk
1Institute of Pharmaceutical Science, King’s College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, United Kingdom
2Department of Chemistry King’s College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, United Kingdom
Full list of author information is available at the end of the article

© 2013 Dunlap et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Background

All of the reef-building corals (Scleractinia; phylum Cnidaria) that create the vast calcium carbonate deposits of coral reefs have evolved an endosymbiotic partnership with photosynthetic dinoflagellates of the genus *Symbiodinium* (Dinophyceae), commonly known as zooxanthellae, which reside within the gastrodermal cells of their scleractinian host [1-3]. Coral-algal symbiosis is a cooperative metabolic adaptation necessary for survival in the shallow oligotrophic (nutrient-poor) waters of tropical and subtropical marine environments [4,5] that drives the productivity of coral reefs [6]. Coral reefs provide habitat and trophic support for many thousands of marine species, the richness of which rival the biological biodiversity of tropical rainforests [7]. Underlying the basic requirements of corals for growth, reproduction and survival are special needs to accommodate symbiont-specific host recognition, to control innate and responsive immune systems, and what is likely to emerge from future research is the extent to which the host is involved in direct regulation of its endosymbiotic populations. Much is understood about the cellular biology of cnidarian-dinoflagellate symbiosis (reviewed in [8]), but less is known at the molecular level of coral symbiology. There is little opposition to the contention that environmental and anthropogenic disturbances are causing alarming losses to coral reefs ([9] and reference therein). Threats to productivity are being imposed by the disruption of coral symbiosis (apparent as "coral bleaching") caused in response to increasing thermal stress attributed to global warming [10,11], from an increase in stress-related coral disease [12-14], from the discharge of domestic and industrial wastes, pollutants from agricultural development and the transport of sediments in terrestrial runoff [15,16], and potentially from imminent declines in coral calcification owing to rising ocean acidification [17-19]. Accordingly, we require a better understanding of the molecular stress responses and adaptive potential of corals. Such information is necessary to predict bleaching events and so better inform effective management policies for the conservation of coral reef ecosystems [20-24].

To understand how coral holobionts respond to environmental change at the molecular level, the identification of genes that may respond by transcription to stress is of primary importance [25]. Thus, the use of transcriptomic methodologies to identify stress-responsive genes has been highly successful [26-32]. Transcriptome high-throughput profiling has allowed changes in gene expression across thousands of genes to be measured simultaneously. Fuelled by data-generating power, the number of coral based studies utilising transcriptomics to investigate molecular responses to environmental stressors has expanded greatly by the acquisition of expressed sequence tag (EST) gene libraries, the fabrication of microarray biochips used to estimate levels of mRNA expression, and by direct analysis using next-generation, high-throughput sequencing. However, much of this work has been conducted using the apomictic state of pre-settlement coral larvae, so transcribed genes relevant to metamorphosis and the cytobiology of the adult polyp are limited to a few recent studies [33-36]. The transcriptome additionally does not provide the structural framework and essential regulatory elements of the functional genome for comprehensive evaluation. Recently, deep metatranscriptomic sequencing of two adult coral holobionts has been made available on searchable databases: PocilloporaBase for *Pocillopora damicornis* [36] and PcnrBase for *Platygyra carnosus* [37]. In contrast, high-throughput metaproteomic analyses to quantify the product yield of stress-response genes of the coral holobiome are yet to be widely adopted by the coral reef scientific community, despite the proteome being the ultimate measure of the coral phenotype [38,39].

The early accumulation of transcriptomic data revealed that a small proportion of coral ESTs matched genes known previously only from other kingdoms of life, implying that the ancestral animal genome contained many genes traditionally regarded as ‘non-animal’ that have been lost from most animal genomes [40]. Furthermore, an unexpected revelation from EST data is the greater extent to which coral sequences resemble human genes than those of the *Drosophila* and *Caenorhabditis* model invertebrate genomes [41,42]. Comparative genomic analysis has revealed higher genetic divergence and massive gene loss within the ecdysozoan lineages. Hence, many genes assumed to have much later evolutionary origins are likely to have been present in an ancestral or early-diverged metazoan [43]. While much of the animal kingdom remains yet to be explored, examples of the metazoan phylum Cnidaria provide a unique insight into the deep evolutionary origins of at least some vertebrate gene families [42]. Thus, the complete genomic sequence of a coral is likely to reveal many genes previously assumed to be strictly vertebrate innovations. To date, cnidarian genomes have been published for the sea anemone *N. vectensis* [42] and the hydroid *Hydra magnipapillata* [44]. Only the coral genome of *Acropora digitifera* is available without restriction on use of its published sequence [45], but the compiled sequence has not been fully annotated. At the time of this writing, the genome assembly of *Acropora millepora* has been released to the public domain [46], also without full annotation, but an embargo is imposed on use of this data that is highly restrictive to the progress of further studies. Understanding how genomic variation affects molecular and organismal biology is the ultimate justification of genome sequencing, and annotation is an essential step in this process. We envisage that unrestricted access to annotation of the *A. digitifera* genome will provide an unprecedented foundation to freely
interrogate the generic molecular structure, possible endobiotic interactions and the response of coral to environmental stress. Accordingly, we offer annotation of the predicted proteome of *A. digitifera* on the open access and searchable database, ZoophyteBase [47]. Use of the ZoophyteBase search engines will allow genes of encoded proteins to be identified that can be examined in context of the cellular physiology, processes of ecological significance, the evolutionary and developmental biology of corals and the functional metabolism of the holobiont that collectively underpin the health of coral reefs.

**Construction and content**

ZoophyteBase is an open access and searchable database of complete annotation of the predicted proteome of the coral *A. digitifera* [48]. It was constructed using the MEGGASENSE system, which is a general system for constructing annotation databases with different sorts of input data (DNA reads, assembled genomes, predicted proteomes) and the possibility of using different combinations of analysis tools to create the annotation (Gacesa et al, in preparation). In the case of ZoophyteBase, hidden Markov model (HMM) profiles [49] were chosen as the annotation tool rather than the more common BLAST searches [50]. HMM profiles are constructed from multiple alignments of protein families and contain information about conserved differences in amino acid residues as well as deletions and insertions [49]. This is particularly important for a coral database, as corals are evolutionarily distant to most other organisms. This means that known homologous sequences present in the databases will usually have relatively low similarity, making BLAST searches inaccurate. The statistical information in an HMM profile gives more sensitive and accurate detection of sequence homology. An additional advantage of HMM profiles is that the statistical significance of hits (the expected value) is much more accurate than that calculated by BLAST programs.

The quality of sequence annotation is limited by the accuracy of information provided in any database used. It is well known that there are many problems with annotation in the large uncurated databases such as the NCBI GenBank nr sequences. Widely accepted, the most accurate database for functional annotation is the KEGG database [51]. The KEGG database organises sequences as groups of KEGG orthologues. These are sets of homologous sequences from as wide a range of organisms as possible having an assigned molecular function. These functions are arranged in a hierarchical fashion and grouped in biological pathways. The sequences belonging to KEGG orthologues were used to construct HMM profiles for annotating the coral sequences. Accordingly, the 23,524 predicted proteins encoded in the coral genome were analysed using HMM profiles. If a protein showed a highly significant correlation (“hit”) to a single HMM profile, this was used to create a “trusted” annotation of the sequence. Choosing a cut-off for this criterion is not trivial, because longer sequences tend to have more significant e-values. For construction of ZoophyteBase the criterion 1e-5 was used. This resulted in 19,044 predicted proteins giving “trusted” sequence annotation. For many of these proteins there were two or more highly significant hits to established HMM profiles. In these cases, the most significant correlation was used to construct our “best-fit” annotation file, but other hits can be viewed by the database user so that expert knowledge can be employed to override the automatic annotation function. In 8,004 out of 19,044 predicted proteins which were annotated, more than one annotation was assigned based on non-overlapping regions within the protein which were used to construct the “best-fit” annotation file. We interpreted these as “fusion” events generated by the *in silico* protein prediction method used, and these proteins were treated as multiple instead of single encoded proteins. Hence, this analysis resulted in the annotation of 33,195 proteins in total, generated from the original 23,524 predicted coral proteins. This is a very conservative annotation scheme, so it can be assumed that most of the annotations are biologically meaningful. Almost 81% (19,044 out of 23,524) of the predicted proteome was assigned using this method.

**Utility**

The MEGGASENSE system was used to generate a web interface for ZoophyteBase. The home page (Figure 1A) allows the use of several functions. A text version of the entire annotation can be downloaded for manual inspection. There is a proteome overview that gives statistics about the database and a breakdown of the annotated functions into different categories of genes. A particularly useful feature of ZoophyteBase is the ability to use text queries employing a search engine that provides a relevant inquiry in the absence of an exact match between key words of a search and those described for a functional protein. The search engine uses text from the KEGG-database, PubMed and other sources to establish links between query words to access protein data using an intelligent Google-like search engine implemented by the search platform Lucene/Solr [52]. This helps to overcome the common problem that different terminology is used by different groups of researchers. The use of this search function is illustrated by using the query “phagocytosis” (Figure 1B). This inquiry finds 42 hits to KEGG orthologue profiles. One of the hits corresponds to amphiphysin (a synaptic vesicle protein) with annotation of two protein homologues encoded in the coral genome. On the data page there is a brief description of the function of amphiphysin together with a PUBMED literature reference. The sequences of the predicted coral
proteins (Figure 1C) can be retrieved, and it is also possible to analyse such data with computer aided drug design methods [53] to look for conserved domains. There are also two tools for the user to examine matches to protein sequences. The user can carry out a BLAST search against the coral protein sequence or analyse the predicted sequence against HMM profiles used to annotate the coral proteome. These tools require only the user to paste their query into the sequence window.

In this manuscript we demonstrate the utility of ZoophyteBase by presenting predicted gene-encoded proteins revealed by annotation of the A. digitifera genome that have physiological, biological and environmental significance. We discuss features of importance in coral physiology: (1) regulatory proteins of symbiosis, (2) planula and early developmental proteins, (3) neural messengers, receptors and sensory proteins, (4) calcification and Ca2+-signalling proteins, (5) plant-derived proteins, (6) proteins of nitrogen metabolism, (7) DNA repair proteins, (8) stress response proteins, (9) antioxidant and redox-protective proteins, (10) proteins of cellular apoptosis, (11) microbial symbioses and pathogenicity proteins, (12) proteins of viral pathogenicity, (13) toxins and venom, (14) proteins of the chemical defencesome and (15) coral epigenetics.

Discussion

Regulatory proteins of symbiosis

Metabolic cooperation is a key feature of coral-algal symbiosis that allows reef-building corals to inhabit the often nutrient-poor waters of tropical oceans [54]. In this phototrophic symbiosis, fixed carbon produced by resident algae is released to the host for nutrition, and the algal symbionts benefit by acquiring the inorganic nutrient wastes of host metabolism [2,55]. The symbiotic
dinoflagellates reside and proliferate within a specialised phagosome (the symbiosome) maintained within host gastrodermal cells. This arrangement requires complex biochemical coordination by the coral at various metabolic stages that includes endocytosis (phagocytosis) by post-settlement polyps to acquire algal symbionts, accord symbiosome recognition to arrest phagosomal maturation for sustained organelle homeostasis, activate symbiophagy or exocytosis to eliminate damaged symbionts [56,57], and regulate apoptotic or exocytotic pathways to remove excess or impaired populations, all of which have long been recognised as essential to preserve the stability of coral symbiosis [58]. Although these processes are poorly understood in corals, it has been realised from studies of the sea anemone Aiptasia pulchella, a related anthozoan also containing Symbiodinium sp. endosymbionts, that the persistence of algal-containing symbiosomes in Cnidaria relies on the exclusion or retention of small Rab GTPase family proteins that are key regulatory components of vesicular trafficking and membrane fusion in eukaryotic cells [59]. Significantly, ApRab3 and ApRab4 accumulate in the biogenesis of maturing symbiosomes of A. pulchella [60,61], and mature symbiosomes enveloping healthy dinoflagellates have tethered ApRab5 [62], a checkpoint antagonist of downstream ApRab7 and ApRab11 proteins that would otherwise direct autophagy of the symbiont cargo [63,64].

Our annotation of the A. digitifera genome reveals sequences encoding putative Rab homologues of the Ras superfamily of proteins (Table 1). In a comparison of cnidarian Rab proteins, eight proteins of A. digitifera matched homologues of Aiptasia pulchella, twenty-nine matched proteins encoded by the aposymbiotic freshwater H. magnipapillata and the aposymbiotic anemone N. vectensis genomes, while seven Rab and Rab-interacting proteins of A. digitifera did not match other cnidarian proteins (Table 2). Significantly, the eight homologues of A. digitifera that matched exclusively Rab proteins of A. pulchella included homologues of the aforementioned ApRab3, ApRab4 and ApRab5 proteins attributed to the maintenance of healthy symbiosomes in Aiptasia, while homologues of the autophagic ApRab7 and ApRab11 proteins are found also in N. vectensis. While Rab GTPase proteins and their effector proteins coordinate consecutive stages of endocytic vesicular transport [65,66], soluble N-ethylmaleimide-sensitive factor attachment receptor (SNARE) proteins are essential for Rab assembly to complete endosomal fusion of vesicle membranes [67], a process by which Rab proteins impart specificity by binding distinct Rab and SNARE partner proteins prior to membrane fusion [68]. Genes encoding syntaxin-like SNARE proteins have been unambiguously identified [69] from coral EST database libraries constructed from expressed mRNA isolated from various early life stages of A. millepora, A. palmata and Orbicella faveolata (= Monastrea faveolata), as well as from the genome of the sea anemone N. vectensis [70]. In meta- zoans, vacuolar r-SNARE receptor proteins comprise the syntaxin, synaptobrevin and VAMP family proteins, of which there are eight syntaxin and syntaxin-binding proteins (plus two plant-like syntaxins). Additionally, there are one t-SNARE target protein to direct vacuolar morphogenesis, two synaptoosomal proteins, one synaptoosomal complex ZIP1 protein (yeast homologue), one synap tobrevin membrane protein of secretory vesicles, ten vesicle-associated membrane proteins (VAMPs), a vacuolar protein-8 regulator of autophagy, four vacuolar-sorting proteins and two SEC22 vesicle trafficking protein encoded in the genome of A. digitifera (Table 1), many of which may interact to provide metabolic transport between the endoplasmic reticulum and Golgi apparatus [71]. Included in this vast but yet unexplored repertoire of vacuolar-acting proteins are the syntaxin-binding amisyn and tomosyn regulators of SNARE complex assembly and disassembly [72,73], which may control membrane fusion in the phagocytic establishment and dissociation of coral symbiosis.

In the final step of exocytosis there is a cytosolic influx of calcium which binds to synaptotagmin to activate completion of membrane SNARE protein assembly with exocytic docking to form the conducting channel for trans-membrane vesicular transport on activation by vesicle-fusing ATPase [74]. As synaptotagmin proteins are not included in the KEGG database, Zoophytebase was used for BLAST searches with all known synaptotagmin sequences [27]. Synaptotagmin proteins from A. digitifera were found having similarity to homologues from diverse invertebrate and vertebrate organisms, including one from the human genome (Table 3). Other Ca²⁺-sensing proteins of A. digitifera, such as calcmodulin and the calcium binding protein CML, are given with calcification and Ca²⁺-signalling proteins.

Intriguingly, annotation of the A. digitifera genome reveals a host cell factor (K14966), but this is not related to the elusive “host factor” of symbiosis demonstrated to be present in tissue homogenates of corals and other marine invertebrates that harbor Symbiodinium spp. endosymbionts [75-77]. Instead, this mammalian transcriptional coactivator host cell factor (HFC-1) is known to mediate the enhancer-promoter assemblies of herpes simplex (HSV) and varicella zoster (VZV) viruses for activation of the latent state for replication [78], such that the coral HCF homologue may have similar relevance as a viral checkpoint transcriptional coactivator of virulence in A. digitifera. HCF-1 expression is coupled also to chromatin modification [79,80] suggesting that the coral protein homologue may have an additional role in
| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.06849      | K06110         | Exocyst complex component 3 |
| v1.00063; v1.01826 | K06111     | Exocyst complex component 4 |
| v1.06336; v1.06337; v1.15354 | K07195 | Exocyst complex component 7 |
| v1.04340 [+ 4 other sequence copies] | K14966 | Host cell factor |
| v1.01629; v1.19166 | K12481 | Rabenosyn-5 |
| v1.18447 [+ 26 other sequence copies] | K07976 | Rab family, other (similar to Rab-6B) |
| v1.02380 | K12480 | Rab GTPase-binding effector protein-1 |
| v1.01032 | K13883 | Rab-interacting lysosomal protein |
| v1.14682; v1.03256; v1.07709 | K12484 | Rab11 family-interacting protein-1/2/5 |
| v1.13055; v1.13176; v1.16348 | K12485 | Rab11 family-interacting protein-3/4 |
| v1.01275 | K07932 | Rab-like protein-2B |
| v1.17629 [+ 13 other sequence copies] | K07933 | Rab-like protein-3 |
| v1.03299; v1.09653 | K07934 | Rab-like protein-4 |
| v1.08498 | K07935 | Rab-like protein-5 |
| v1.16155 [+5 other sequence copies] | K07874 | Ras-related protein Rab-1A |
| v1.09098 | K07875 | Ras-related protein Rab-1B |
| v1.13558; v1.08983 | K07877 | Ras-related protein Rab-2A |
| v1.14260 | K07878 | Ras-related protein Rab-2B |
| v1.07500; v1.20532; v1.07498 | K07884 | Ras-related protein Rab-3D |
| v1.21242; v1.07502 | K07880 | Ras-related protein Rab-4B |
| v1.01341; v1.05619 | K07888 | Ras-related protein Rab-5B |
| v1.07125 | K07889 | Ras-related protein Rab-5C |
| v1.09239 | K07893 | Ras-related protein Rab-6A |
| v1.10443; v1.13335 | K07897 | Ras-related protein Rab-7A |
| v1.03086; v1.17122; v1.07231 | K07916 | Ras-related protein Rab-7L1 |
| v1.02275 [+ 4 other sequence copies] | K07901 | Ras-related protein Rab-8A |
| v1.24612 | K07899 | Ras-related protein Rab-9A |
| v1.00411 | K07900 | Ras-related protein Rab-9B |
| v1.10697; v1.01515 | K07903 | Ras-related protein Rab-10 |
| v1.22278; v1.04408; v1.12528 | K07905 | Ras-related protein Rab-11B |
| v1.07033; v1.23028 | K07881 | Ras-related protein Rab-14 |
| v1.02275 | K07908 | Ras-related protein Rab-15 |
| v1.16455; v1.14911; v.1.4959 | K07910 | Ras-related protein Rab-18 |
| v1.04714 | K07911 | Ras-related protein Rab-20 |
| v1.01878; v1.12184 | K07890 | Ras-related protein Rab-21 |
| v1.09930 | K06234 | Ras-related protein Rab-23 |
| v1.13579; v1.12841 | K07912 | Ras-related protein Rab-24 |
| v1.10183 | K07913 | Ras-related protein Rab-26 |
| v1.08199 | K07885 | Ras-related protein Rab-27A |
| v1.13978; v1.18893 | K07917 | Ras-related protein Rab-30 |
| v1.03085; v1.06007; v1.07729 | K07918 | Ras-related protein Rab-32 |
| v.24721 | K07919 | Ras-related protein Rab-33A |
| v1.18892 | K07920 | Ras-related protein Rab-33B |
| Accession   | Description                                                                 |
|-------------|------------------------------------------------------------------------------|
| v1.16060    | K07876 Ras-related protein Rab-35                                             |
| v1.15894    | K07922 Ras-related protein Rab-36                                             |
| v1.03080    | K07923 Ras-related protein Rab-38                                             |
| v1.21391    | K07924 Ras-related protein Rab-39A                                            |
| v1.14786    | K07928 Ras-related protein Rab-40                                             |
| v1.05611    | K08502 Regulator of vacuolar morphogenesis (t-SNARE domain)                   |
| v1.18253    | K08520 SEC22 vesicle trafficking protein A/C                                   |
| v1.15499    | K13814 t-SNARE domain-containing protein 1                                    |
| v1.05749    | K08516 Synaptobrevin homologue YKT6                                          |
| v1.13229    | K12768 Synaptosomal complex protein ZIP1                                       |
| v1.16533; v1.17141 | K08508 Synaptosomal-associated protein, 23 kDa                             |
| v1.05301    | K08509 Synaptosomal-associated protein, 29 kDa                                |
| v1.19071    | K04560 Syntaxin 1A                                                            |
| v1.04614; v1.22747 | K08486 Syntaxin 1B/2/3                                                        |
| v1.16462    | K08490 Syntaxin 5                                                             |
| v1.20758; v1.21534 | K08498 Syntaxin 6                                                             |
| v1.22836; v1.15499 | K08488 Syntaxin 7                                                             |
| v1.01959; v1.24227 | K08501 Syntaxin 8                                                             |
| v1.02007; v1.06683; v1.12727 | K08491 Syntaxin 17                                                           |
| v1.21308; v1.11830; v0.01582 | K08492 Syntaxin 18                                                           |
| v1.22100; v0.09457 | K08518 Syntaxin binding protein 5 (tomosyn)                                   |
| v1.18555    | K08519 Syntaxin binding protein 6 (amisyn)                                    |
| v1.12938    | K08500 Syntaxin of plants SYP6                                                |
| v1.06575    | K08506 Syntaxin of plants SYP7                                                |
| v1.14699    | K08507 Unconventional SNARE in the endoplasmic reticulum protein 1            |
| v1.23782    | K08332 Vacuolar protein 8                                                     |
| v1.15282; v1.24603; v0.01672 | K12196 Vacuolar protein-sorting-associated protein 4                            |
| v1.17791    | K12479 Vacuolar protein sorting-associated protein 45                         |
| v1.20907    | K11664 Vacuolar protein sorting-associated protein 72                         |
| v1.15996    | K12199 Vacuolar protein sorting-associated protein VTA1                       |
| v1.15614    | K08510 Vesicle-associated membrane protein 1 (synaptobrevin)                  |
| v1.13353    | K13504 Vesicle-associated membrane protein 2 (synaptobrevin)                  |
| v1.12458; v0.07528 | K13505 Vesicle-associated membrane protein 3 (cellubrevin)                   |
| v1.19735; v1.21831; v0.07186 | K08513 Vesicle-associated membrane protein 4 (Golgi transport)            |
| v1.05299    | K08514 Vesicle-associated membrane protein 5 (exocytosis)                     |
| v1.13557; v1.24610 | K08515 Vesicle-associated membrane protein 7 (exocytosis)                  |
| v1.12279    | K08512 Vesicle-associated membrane protein 8 (endobrevin)                     |
| v1.00261; v1.08699; v0.04334 | K06096 Vesicle-associated membrane protein A                           |
| v1.20177    | K06027 Vesicle-fusing ATPase                                                  |
| v1.11431; v1.010487 | K08517 Vesicle transport protein SEC22                                       |
| v1.06393; v1.13003; v0.08735; v0.04261 | K08493 Vesicle transport interaction with t-SNAREs 1                             |
epigenetic reprogramming of the chromatin histone-DNA complex at different stages of development.

**Planula and early developmental proteins**

In this section we discuss predicted proteins encoded in the *A. digitifera* genome having functional homology to known proteins are specific to early embryonic development, planula larvae function and morphogenesis, which are given in Table 4. Annotation of the coral genome reveals a large set of homeobox proteins involved in the regulation of anatomical development during morphogenesis. The homeobox is a highly conserved DNA sequence (homeodomain) within genes that binds to DNA in a sequence-specific manner [81] often at the promoter region of their target gene to affect transcription in a sequence-specific manner [81] often at the promoter region of their target gene to affect transcription in a sequence-specific manner often at the promoter region of their target gene to affect transcription in a sequence-specific manner often at the promoter region of their target gene to affect transcription in a sequence-specific manner.

Table 2 Distribution of Rab homologues of *Aiptasia puchella*, *Hydra magnipapillata* and *Nematostella vectensis* in the predicted proteome of *A. digitifera*

| A. digitifera Rab protein | Cnidarian encoding Rab homologue |
|---------------------------|---------------------------------|
| Rab-like protein- 28, Rab-2B Rab-3D, Rab-4B, Rab-5B, Rab-26, Rab-32, Rab-38 | A. puchella, N. vectensis, H. magnipapillata |
| Rab-like protein-3, Rab-36 Rab-2A, Rab-23 | A. puchella, H. magnipapillata |
| Rab-like protein-6B, Rab-6A, Rab-7 L1, Rab-10, Rab11B, Rab-30, Rab-33B | A. puchella, N. vectensis |
| Rab effector protein-1, Rab11-interacting protein-3/4 | H. magnipapillata, N. vectensis |
| Rab-like protein-4, Rab-like protein-5, Rab-1A, Rab5C, Rab-7A, Rab-8A, Rab-9A, Rab-14, Rab-18, Rab-20, Rab-21, Rab-24, Rab-27A, Rab-35 | A. puchella, H. magnipapillata, N. vectensis |
| Rab-interacting lysosomal protein, Rab11-interacting protein-1/2/5, Rab-1B, Rab-9B, Rab-14, Rab-3A, Rab-39A, Rab-40 | No match |

Table 3 Synaptotagmin proteins in the predicted proteome of *A. digitifera*

| Gene sequence | GenBank Accession | Genome encoded homologue |
|---------------|------------------|--------------------------|
| v1.08623      | GI:268530614     | Caenorhabditis briggsae: XP_002630433 (worm) |
| v1.20682; v1.10560; v1.02080; v1.10015 | GI:150416761 | Platyneuris dumerilii: ABR68850 (worm) |
| v1.10269; v1.04412 | GI:288869516 | Nasonia vitripennis: NP_001165865 (wasp) |
| v1.01508 | GI:29378331 | Lymnaea stagnalis: AA093847 (snail) |
| v1.18613 | GI:391339919 | Metaseiulus occidentalis: XP_003744294 (mite) |
| v1.07402 | GI:260834895 | Branchiostoma floridae: XP_002612445 (lancelet) |
| v1.01542 | GI:149067023 | Rattus norvegicus: EDM16756 (rat) |
| v1.20683 | GI:383860584 | Megachile rotundata: XP_003705769 (bee) |
| v1.17688 | GI:48529130 | Oreochromis niloticus: XP_003452067 (fish) |
| v1.15777; v1.14902 | GI:269785031 | Saccoglossus kowalevskii: NP_001161667 (worm) |
| v1.17175; v1.11521 | GI:11559313 | Halocynthia roretzi: BAB18864 (ascidian) |
| v1.03344; v1.03345 | GI:12658419 | Manduca sexta: AF331039 (moth) |
| v1.16152 | GI:395729192 | Pongo abelii: XP_003780414 (orangutan) |
| v1.10268 | GI:327283049 | Anolis carolinensis: XP_003226254 (lizard) |
| v1.10277 | GI:125984480 | Drosophila pseudoobscura: XP_001356004.1 (fly) |
| v1.02083; v1.02777 | GI:226490194 | Schistosoma japonicum: CX693391.1 (fluke) |
| v1.04326 | GI:167744962 | Homo sapiens: 2883_A (human) |
| v1.14682; v1.04180 | GI:241704658 | Ixodes scapularis: XP_002411967 (tick) |
| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.09797; v1.11180; v1.08414 | K03776 | Aerotaxis receptor (oxygen sensing) |
| v1.07838 [+5 other sequence copies] | K07822 | Archaeal flagellar protein FlaC |
| v1.14039; v1.11310; v.1.1309 | K05502 | Bone morphogenetic protein 1 |
| v1.01025; v1.17008; v1.15796; v1.23658 | K04662 | Bone morphogenetic protein 2/4 |
| v1.02299; v1.07696; v1.10675 | K04663 | Bone morphogenetic protein 5/6/7/8 |
| v1.06335; v1.01763 | K04673 | Bone morphogenetic protein receptor type-1A |
| v1.13481 | K13578 | Bone morphogenetic protein receptor type-1B |
| v1.10550 [+4 other sequence copies] | K04671 | Bone morphogenetic protein receptor type-2 |
| v1.000912 [+4 other sequence copies] | K13579 | Bone morphogenetic protein receptor type-1, invertebrate |
| v1.19370 | K14624 | C-C motif chemokine 2 |
| v1.23163 | K12499 | C-C motif chemokine 5 |
| v1.08576 | K05511 | C-C motif chemokine 15/23 |
| v1.09229 | K05512 | C-C motif chemokine 19/21 |
| v1.09305 | K08373 | C-C chemokine receptor-like 2 |
| v1.04942 | K04179 | C-C chemokine receptor type 4 |
| v1.02658 | K04245 | Chemokine-like receptor 1 |
| v1.21300 | K12671 | C-X-C motif chemokine 10 |
| v1.16396; v1.21991 | K10035 | C-X-C motif chemokine 16 |
| v1.23712 | K11522 | Chemotaxis family two-component system response regulator PixG |
| v1.09435 | K13490 | Chemotaxis family, histidine kinase sensor response regulator (WspE-like) |
| v1.14142; v.1.05300 | K05874 | Chemotaxis protein I, serine sensor receptor (MCP family) |
| v1.07361 | K05877 | Chemotaxis protein IV, peptide sensor receptor (MCP family) |
| v1.17411 | K03414 | Chemotaxis protein CheZ |
| v1.16104 | K00575 | Chemotaxis protein methyltransferase CheR |
| v1.15537 [+ 7 other sequence copies] | K08482 | Circadian clock protein KaiC |
| v1.14925 [+ 4 other sequence copies] | K02223 | Circadian locomotor output cycles kaput protein |
| v1.06432 [+ 9 other sequence copies] | K04512 | Dishevelled associated activator of morphogenesis |
| v1.17637 [+ 70 other sequence copies] | K10408 | Dynein heavy chain, axonemal |
| v1.00202 [+5 other sequence copies] | K10409 | Dynein intermediate chain 1, axonemal |
| v1.04986; v1.09649; v1.23645 | K11413 | Dynein intermediate chain 2, axonemal |
| v1.08695; v1.09481; v1.23153 | K10411 | Dynein light chain 1, axonemal |
| v1.11684 | K10412 | Dynein light chain 4, axonemal |
| v1.23322; v1.01131; v.014207 | K10410 | Dynein light intermediate chain, axonemal |
| v1.14083 | K02401 | Flagellar biosynthetic protein FliH |
| v1.16997 | K02420 | Flagellar biosynthetic protein FliQ |
| v1.02867 | K02396 | Flagellar hook-associated protein 1 FlgK |
| v1.18101; v.1.13427 | K02408 | Flagellar hook-basal body complex protein FlIE |
| v1.04339; v1.07633 | K06603 | Flagellar protein FlaG |
| v1.17895 [+5 other sequence copies] | K02383 | Flagellar protein FlbB |
| v1.21111 | K02413 | Flagellar protein Flu |
| v1.17651 [+ 13 other sequence copies] | K02415 | Flagellar protein FlgL |
| v1.01971 [+ 6 other sequence copies] | K02418 | Flagellar protein Fio/FliZ |
| v1.14031 | K02423 | Flagellar protein FliT |
| Accession | Description |
|-----------|-------------|
| v1.08025  | v1.02396; v1.15777; v1.20693 | Flagellar P-ring protein precursor FlgI |
| v1.24732  | v1.41729; v1.19334; v1.04164 | Homeobox protein cut-like |
| v1.01801  | v1.09451 | Homeobox protein engrailed |
| v1.16835; v1.06323 | v1.09452 | Homeobox protein aristless-4 |
| v1.13604  | v1.09453 | Homeobox protein expressed in ES cells 1 |
| v1.06346; v1.08163 | v1.09454 | Homeobox protein goosecoid |
| v1.17295; v1.17294 | v1.09455 | Homeobox protein, BarH-like (vertebrate neurogenesis) |
| v1.07457  | v1.11157; v1.08573; v1.15250 | Homeobox protein EMX |
| v1.01800  | v1.10929; v1.06346; v1.05443; v1.07458 | Homeobox protein GSH |
| v1.13684; v1.24444 | v1.08025 | Homeobox protein HB9 |
| v1.16254; v1.16064 | v1.08024 | Homeobox protein HEX |
| v1.07458; v1.06706; v1.06705 | v1.09339 | Homeobox protein HLX1 |
| v1.06347; v1.06348; v1.17294 | v1.09302 | Homeobox protein HoxA/B2 |
| v1.06125  | v1.09306 | Homeobox protein HoxA/B/C6 |
| v1.19818  | v1.09304 | Homeobox protein HoxA/B/C/D4 |
| v1.06706  | v1.09301 | Homeobox protein HoxA/B/D1 |
| v1.02056  | v1.09353 | Homeobox protein LBX |
| v1.06347; v1.06348 | v1.09352 | Homeobox protein Unc-4 |
| v1.24342; v1.04552 | v1.09318 | Homeobox protein ventral anterior |
| v1.03823; v1.10070; v1.04435 | v1.09309 | Homeobox protein Nkx-1 |
| v1.12852  | v1.21630 | Homeobox protein Nkx-2.2 |
| v1.10625  | v1.09345 | Homeobox protein Nkx-2.5 |
| v1.10625; v1.13865; v1.05476 | v1.09347 | Homeobox protein Nkx-2.8 |
| v1.21628; v1.05475; v1.05477 | v1.09348 | Homeobox protein Nkx-3.1 |
| v1.06135; v1.10071 | v1.09995 | Homeobox protein Nkx-3.2 |
| v1.14702  | v1.09349 | Homeobox protein Nkx-5 |
| v1.14917; v1.11907 | v1.09303 | Homeobox protein Nkx-6.1 |
| v1.00777; v1.21453 | v1.09350 | Homeobox protein Nkx-6.2 |
| v1.00602 | v1.00602 | Homeobox protein Otx |
| v1.16722; v1.12785 | v1.09374 | LIM homeobox protein 3/4 |
| v1.11281; v1.05135 | v1.09375 | LIM homeobox protein 6/8 |
| v1.07988; v1.22037 | v1.09371 | LIM homeobox transcription factor 1 |
| v1.09328  | v1.09328 | LIM homeobox protein 6/8 |
| v1.15178; v1.10962; v1.16587; v1.01557 | v1.09330 | Methyl-accepting chemotaxis protein Pix (MCP family) |
| v1.23288; v1.13857 | v1.09332 | Nematode chemoreceptor |
| v1.00602 | v1.00602 | Paired mesoderm homeobox protein 2 |
| v1.10070; v1.04435 | v1.09309 | Period circadian protein |
| v1.00777; v1.21453 | v1.09322 | Pheromone a factor receptor |
morphogenesis 1 (Daam1) that initiates cytoskeleton formation via the control of actin assembly. Daam1 was found crucial for gastrulation in Xenopus [91], wherein Daam1 mutants of Drosophila exhibit trachea defects [92], and in mammals Daam1 is highly expressed in multiple developing organs and is deemed essential for cardiac morphogenesis [93]. Similar morphogenetic genes express regulatory proteins that are necessary for vacuole biogenesis in yeasts [94]. Others express bone morphogenetic proteins (and their BMP receptors), which are potent multi-functional growth activators that belong to the transforming growth factor beta (TGF-beta) cytokine superfamily of proteins that in humans have various functions during embryogenesis, skeletal formation, neurogenesis and haematopoiesis [95]. However, since many of the homeobox and morgenetic proteins (Table 4) are homologues of proteins with functions ascribed to higher organisms, their precise function in A. digitifera cannot be ascertained by KEGG orthology alone.

Another protein encoded in the A. digitifera genome is a retina and anterior neural fold homeobox-like (RAX) protein that may activate the development of primitive coral photoreceptors [96,97], including a blue-light-sensing, cryptochrome photoreceptor that in primitive coral photoreceptors [96,97], including a blue-light-sensing, cryptochrome photoreceptor that in A. millepora is implicated in the detection of light from the lunar cycle of night time illumination to signal synchronous coral spawning [98,99]. Photosensitive behaviours and the circadian rhythms of corals are well described, and diurnal cycles of gene transcription that regulate circadian biological processes in the coral A. millepora have been reported [100]. Such traits in A. millepora appear regulated by an endogenous biological clock entrained to daily cycles of solar illumination [101]. Annotation of the A. digitifera genome reveals a circadian timekeeper protein KaiC [102] that in cyanobacteria is activated during the diurnal phosphorylation rhythm [103,104]. In Synechococcus elongatus, KaiC regulates the rhythmic expression of all other proteins encoded in the genome [105], yet no homologue of any of the prokaryotic clustered circadian kiaABC genes has been identified in eukaryotes [106]. In Drosophila, KaiC together with a homologue of the eukaryotic period (Per) circadian protein drives circadian rhythms in eclosion (hatching) and locomotor activity [107]. Nevertheless, a circadian locomotor output cycles kaput (CLOCK) homologue (Table 4) was found in our annotation. Since CLOCK proteins serve as an essential activator of downstream elements in pathways critical to the regulation of circadian rhythms in eukaryotes [108], it would be worthy to examine how transcription of the RAX-like homeobox protein in this coral contributes to the development of circadian functions by activation of kaiC, per and Clock genes. Such a study might reveal that components of the animal circadian clock are more ancient than data previously suggested [109].

Broadcast-spawning corals, such as A. digitifera, release gametes, and the fertilised eggs develop into planula larvae within the water column until they have reached settlement competency, find a suitable hard substrate, attach and develop into the polyp on metamorphosis. Coral sperm and planula larvae achieve motility using flagella (sperm) or cilia (larvae) as their locomotor organelles. The eukaryotic axonemal proteins of cilia and flagella are composed of a dynein ATPase protein to provide mechanochemical energy transduction together with the principle structural proteins of the ciliary/flagellar microtubules [110]. The flagellar/ciliary microtubules consist of filaments composed of α- and β-tubulins, microtubule-stabilising tektins and kinesin motor proteins [111-113]. The coral genome encodes members of the dynein axonemal (flagella and cilia) proteins (Table 4) and many of the dynein cytoplasmic proteins (not tabulated), the latter being involved in intracellular organelle transport and centrosome assembly. The coral genome encodes α- and β-tubulins and members of the eukaryotic kinesin superfamily proteins (not tabulated). Amongst the many kinesin proteins encoded in the coral genome is the kinesin family member 3/17 protein, which is a direct homologue of the kinesin-II intraflagellar transport protein FLA10 essential for flagella assembly in the alga Chlamydomonas [114]. The microtubule-stabilising tektin protein, which is required for cilia and flagella assembly [113], is also encoded in the coral genome [note: there is no KEGG orthology identifier assigned to this protein]. It was a surprise, however, to find a large complement of

Table 4 Planula and early developmental proteins in the predicted proteome of A. digitifera (Continued)

| Accession | KEGG Orthology | Description |
|-----------|----------------|-------------|
| v1.22964; v1.17135 | K11213 | Pheromone alpha factor receptor |
| v1.05611 [+ 13 other sequence copies] | K08502 | Regulator of vacuolar morphogenesis |
| v1.04431 | K09333 | Retina and anterior neural fold homeobox-like protein |
| v1.17636 | K09331 | Short stature homeobox protein |
| v1.14704 | K09340 | T-cell leukemia homeobox protein |
| v1.11765 | Na^4 | Tektin |
| v1.04154 | K02669 | Twitching motility protein PiT |

^NA KEGG orthology designation not assigned.
prokaryotic flagellar proteins encoded in the coral genome consisting of archaeal flagellar (FlaC and FlaG), bacterial filament (FibB, FliE, FliF, FliJ, FliK, FlhO/FliZ, FliQ and FlIT) homologue components (Table 4). Included also are the prokaryote homologues FlgN and FliB that regulate transcriptional activation of flagellar assembly [115,116] and FlhB which controls the substrate specificity of the entire prokaryotic flagellar apparatus [117]. Encoded in the coral genome is a flagella-independent Type IV twitching mobility protein PilT that affords social gliding translocation in many prokaryotic organisms controlled by complex signal transduction systems that include two-component sensor regulators [118]. It is unlikely that these genes are derived from contamination from bacterial DNA. Such contamination would manifest itself by the random occurrence of bacterial genes from the whole genome including many housekeeping genes. In this case, the genes occur as members of groups with specialised functions, suggesting that multiple horizontal gene transfers between bacteria and the coral genome have occurred [119]. Their precise function in A. digitifera remains unknown; homologues of these prokaryotic genes have not been described previously in any other eukaryote genome.

Linked closely with flagellar/ciliary proteins are the sensory receptors that signal chemotaxis or avoidance to direct cellular motility. The coral genome reveals a variety of genes that encode chemoreceptor and chemotaxis proteins (Table 4). The chemoreceptor proteins of A. digitifera include an oxygen-sensing aerotaxis receptor that in bacteria invokes an avoidance response to anoxic micro-environments [120]. Encoded also are a nematode sensory chemoreceptor homologue [121], two homologous pheromone factor receptor proteins that in fungi activate a species-specific mating response [122], three chemotaxis protein sensor receptors belonging to the methyl-accepting chemotaxis family of proteins (MCPs) in bacteria and archaea [123], and two proteins (CheZ and CheR) and two regulators (PixG and WspE) of the two-component signal transduction (TCST) system for activation of gene expression. In bacteria and archaea, as well as some plants, fungi and protozoa [124], TCST systems mediate many cellular processes that respond to a broad range of environmental stimuli via activation of a specific histidine (or serine) kinase sensor and its cognate response regulator [125]. There are 77 sequence matches to various elements of the TCST family of proteins in the A digitifera genome (data not tabulated). Included also are genes encoding members of the chemotactic cytokine (chemokine) family of sensory proteins that on secretion directs chemotaxis in nearby responsive cells by stimulating target chemokine receptors; both chemokine and chemokine receptor proteins are encoded in the coral genome. Significantly, sensory chemokines/chemokine receptors are found in all vertebrates, some viruses and some groups of bacteria, but none have been described previously for invertebrates [126].

**Neural messengers, receptors and sensory proteins**

Corals and other cnidarians are the earliest extant group of organisms to have a primitive nervous system network [127] thought to be evolved from a eumetazoan ancestor prior to the divergence of Cnidaria and the Bilateria [128,129]. Unlike marine sponges (Porifera) that predate synaptic innovation [130], cnidarians possess a homogenous nerve net that, although lacking any form of cephalization, accommodates fundamental neurosensory transmission across the nerve net to end in a motoneural junction to coordinate tentacle movement required for feeding and predator avoidance [131]. The nervous systems of cnidarians consist of both ectodermal sensory cells and their effector cells and endodermal multipolar ganglia capable of neurotransmission [132]. At the functional level, synaptic transmission in cnidarians relies on fast neurotransmitters (glutamate, GABA, glycine) and slow neurotransmitters (catecholamine, serotonin, neuropeptides) for sensory-signal conduction [133]. At the ultrastructural level, many cnidarian neurons have multifunctional traits of sensory, neurosecretory and stimulatory attributes [134]. Significantly, the genome of A. digitifera encodes the expression of a ciliary neurotrophic factor, which is a polypeptide hormone and nerve growth factor that promotes neurotransmitter synthesis, neurite outgrowth and regeneration [135]. Additionally, the coral genome encodes nerve growth factor and neurotrophic kinase receptors, a survival motor neuron protein, a survival neuron splicing factor, the neural outgrowth protein neurotrophin, and a neurotrophin growth factor attributed to signalling neuron survival, differentiation and growth (Table 5). Encoded for neuron regulation and development are several neuron cation-gated channels, a neuronal guanine nucleotide exchange factor, a neurotransmitter Na⁺ symporter, several neurogenic differentiation proteins, a neuronal PAS domain transcription factor for activation of neurogenesis, the axon guidance protein neurophin-2, a neural crest protein of embryonic neural development, neural ELAV-like transcription proteins of neurogenesis, a Notch protein (79 sequence domain matches) and a neutralized protein subset of the Notch signalling pathway that promotes neuron proliferation in early neurogenic development. Structural elements of the coral nerve net include neurofilament polypeptides and neuronal adhesion proteins.

Cnidarians differentiate highly specialised sensory and mechanoreceptor cells involved in the capture of prey and for defence against predators. Their stinging cells, termed nematocysts or cnidocytes, are stimulated by adjacent chemosensory cells. Nematocysts trigger the release of a stinging barb (cnidae tubule) via ultra-fast
exocytosis on physical contact with ciliary mechano-receptors of the cnidocyte to deliver the discharge of its venom [136]. Despite considerable advances in the sensory biology of cnidarians, knowledge of the specific receptor genes that regulate cnidocyte function remains incomplete. In Hydra, and perhaps other cnidarians, cnidocyte discharge is controlled by an ancient light-activated, opsin-mediated phototransduction pathway [137] that precedes the evolution of cubozoan (box jellyfish) eyes [138]; cubozoans are the most basal of animals to have eyes containing a lens and ciliary-type visual cells similar to that of vertebrate eyes [139]. These G-coupled opsin photoreceptors of the retinylidene-forming protein family encoded in the genome of A. digitifera include rhodopsin, bacteriorhodopsin, c-opsin, r-opsin and G\textsubscript{0}-opsin (Table 5), but not the Gs-subfamily of opsin receptors reported to be present in sea anemones, hydra and jellyfish [140], that together with cyclic nucleotide-gated (CNG) ion channel proteins, arrestin (β-adrenergic receptor inhibitor) and other retino-protein receptors, are usual components of the bilaterian phototransduction cascade. Present also are genes to express rhodopsin kinase and β-adrenergic receptor kinase which are related members of the serine/threonine kinase family of proteins that specifically initiate deactivation of G-protein coupled receptors. Additional proteins of retinol metabolism of the phototransduction pathway encoded in the A. digitifera genome are retinol dehydrogenase, all-trans-retinol 13,14 reductase and phosphatidylcholine (lithchlin)-retinol O-acyltransferase, a neural retina-specific leucine zipper protein that is an intrinsic regulator of photoreceptor development and function, and a retina and anterior neural fold homeobox-like protein that modulates the expression of photoreceptor genes within the rhodopsin promoter. The genome of A. digitifera encodes also a blue light-sensing, cryptochrome photoreceptor thought to signal synchronous coral spawning by detecting illumination from the lunar cycle [98,99].

The A. digitifera genome reveals genes to express a broad array of neurotransmitter receptor proteins (Table 5), including glycine and glutamate neuroreceptors, adrenergic receptors that target non-dopamine catecholamines (i.e., epinephrine and norepinephrine), dopamine, muscarinic and nicotinic acetylcholine receptors, sensory G protein-coupled receptors and γ-aminobutyric acid (GABA) ligand-gated ion channel and G protein-coupled receptors (and inhibitors), several of which are encoded in high copy numbers. Cellular trafficking of neurotransmitters to presynaptic terminals is essential for neurotransmission, and significantly the genome of A. digitifera encodes a wide range of solute carrier neurotransmitter transporters, including a high affinity choline transporter and an acetylcholine-specific protein belonging to the major facilitator superfamily (MFS) of secondary transporters. Encoded also is dopamine β-monoxygenase that catalyses the conversion of dopamine to nor-epinephrine in the catecholamine biosynthetic pathway, which is necessary for cross-activation of adrenergic neuroreceptors [141]. Notably, the A. digitifera genome encodes acetylcholinesterase that is expressed at neuromuscular junctions and cholinergic synapses where its protease activity serves to terminate synaptic transmission. The primitive nervous networks of cnidarians are strongly peptidergic with at least 35 neuropeptides identified from different cnidian classes [142]. Our annotation of the sequenced A. digitifera genome, however, revealed only the neuropeptide FF-amide neurotransmitter, a RF amide related peptide, and its neuropeptide FF and Y receptors (Table 5). Neuropeptides are usually expressed as large precursor proteins which comprise multiple copies of “immature” neuropeptides. Our annotation did not readily reveal these precursor neuropeptide proteins, but we did find enzymes required for their processing, for example, a variety of carboxypeptidase enzymes (not tabulated) that remove propeptide carboxyl residues at basic peptidase sites, and the mature peptide neurotransmitters that are finished by consecutive modification by peptidylglycine (α-hydroxylating) monoxygenase (PHM) and peptidyl α-hydroxyglycine α-amidating lyase (PAL) enzymes, both of which are commonly expressed in mammals as a single bifunctional peptidylglycine monoxygenase (K00504/EC 1.14.17.3) [143]. Our extensive catalogue of animal-like neural and sensory proteins revealed by genome annotation is testament that essential neuro-biological features were developed in the primitive neural networks of early eumetazoa evolution.

**Calcification and Ca\textsuperscript{2+}-signalling proteins**

The massive structures of coral reefs evident today are a construction of aggregated calcium carbonate deposited over long geological time by scleractinian corals and other calcifying organisms, yet our understanding of the molecular processes that regulate the biological processes of coral calcification is limited [144]. Ca\textsuperscript{2+} transfer from seawater to the calcioelastic site of coral calcification occurs by passive diffusion through the gastrovascular cavity [145] and by active calcium transport [146]. Active entry of Ca\textsuperscript{2+} through the oral epithelial layer is regulated by voltage-dependent calcium channels, such as demonstrated by the L-type alpha protein cloned from the reef-building coral Stylophora pistillata [147]. Ca\textsuperscript{2+} transport across the calcioelastic ectoderm to the extracellular calcifying site is facilitated by the plasma-membrane ATP-dependent calcium pump that in S. pistillata resemble the Ca\textsuperscript{2+}-ATPase family of mammalian proteins [148]. By 2H\textsuperscript{+}/Ca\textsuperscript{2+}-exchange at the calcioelastic membrane, Ca\textsuperscript{2+}-ATPase removes H\textsuperscript{+} (from the net reaction Ca\textsuperscript{2+} + CO\textsubscript{2} + H\textsubscript{2}O ⇒ CaCO\textsubscript{3} + 2H\textsuperscript{+}) thereby
## Table 5 Neuronal and sensory proteins in the predicted proteome of *A. digitifera*

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.01918 [+ 5 other sequence copies] | K01049 | Acetylcholinesterase |
| v1.18087; v1.14516 | K04136 | Adrenergic receptor alpha-1B |
| v1.06394 | K04137 | Adrenergic receptor alpha-1D |
| v1.09628; v1.15688; v1.00966 | K04140 | Adrenergic receptor alpha-2C |
| v1.19831; v1.20450 | K04142 | Adrenergic receptor beta-2 |
| v1.17293 | K00910 | beta-Adrenergic-receptor kinase |
| v1.13740 [+ 5 other sequence copies] | K04828 | Amiloride-sensitive cation channel 1, neuronal (degenerin) |
| v1.23541 [+ 6 other sequence copies] | K04829 | Amiloride-sensitive cation channel 2, neuronal |
| v1.09323 [+ 4 other sequence copies] | K04439 | beta-Arrestin |
| v1.07723; v1.22465 | K04641 | Bacteriorhodopsin |
| v1.08062 | K05420 | Ciliary neurotrophic factor |
| v1.03288 [+ 5 other sequence copies] | K02295 | Cryptochrome |
| v1.20011; v1.20036; v1.20084; v1.18607 | K04948 | Cyclic nucleotide gated channel alpha 1 |
| v1.21470 | K04951 | Cyclic nucleotide gated channel alpha 4 |
| v1.21783; v1.01466; v1.01466 | K05326 | Cyclic nucleotide gated channel, invertebrate |
| v1.03645 | K05391 | Cyclic nucleotide gated channel, other eukaryote |
| v1.21256 | K08762 | Diazepam-binding inhibitor (GABA receptor, acyl-CoA-binding protein) |
| v1.22156 [+ 6 other sequence copies] | K00503 | Dopamine beta-monoxygenase |
| v1.21775; v1.15989 | K04148 | Dopamine D1-like receptor |
| v1.14160; v1.01697 | K04144 | Dopamine receptor D1 |
| v1.05089; v1.20018 | K04145 | Dopamine receptor D2 |
| v1.14030; v1.23273 | K04146 | Dopamine receptor D3 |
| v1.20536 | K13088 | ELAV-like protein 1 |
| v1.18658 [+ 5 other sequence copies] | K13208 | ELAV-like protein 2/3/4 |
| v1.05774 [+ 18 other sequence copies] | K08408 | G protein-coupled receptor 6 |
| v1.00572; v1.18152 | K04313 | G protein-coupled receptor 17 |
| v1.23842 | K04316 | G protein-coupled receptor 19 |
| v1.03948 | K08411 | G protein-coupled receptor 26 |
| v1.09271 | K08383 | G protein-coupled receptor 34 |
| v1.05595 | K04243 | G protein-coupled receptor 37 (endothelin receptor type B-like) |
| v1.04019 | K08409 | G protein-coupled receptor 45 |
| v1.19913; v.09821; v1.04291 | K08450 | G protein-coupled receptor 56 |
| v.05404 | K04321 | G protein-coupled receptor 63 |
| v1.02179; v1.10397 | K08451 | G protein-coupled receptor 64 |
| v1.23269 [+ 5 other sequence copies] | K08408 | G protein-coupled receptor 68 |
| v1.21091 | K08421 | G protein-coupled receptor 84 |
| v1.11008 | K04302 | G protein-coupled receptor 85 |
| v1.21884; v1.01951 | K08452 | G protein-coupled receptor 97 |
| v1.03243 [+ 13 other sequence copies] | K08378 | G protein-coupled receptor 103 |
| v1.13790; v1.18939 | K08453 | G protein-coupled receptor 110 |
| v1.09442; v1.14019 | K08455 | G protein-coupled receptor 112 |
| v1.24009 | K08456 | G protein-coupled receptor 113 |
| v1.04290 | K08459 | G protein-coupled receptor 114 |
Table 5 Neuronal and sensory proteins in the predicted proteome of *A. digitifera* (Continued)

| v1.06608; v1.24223 | K08457 | G protein-coupled receptor 115 |
| v1.10800 [+ 6 other sequence copies] | K08458 | G protein-coupled receptor 116 |
| v1.07662 [+ 6 other sequence copies] | K08462 | G protein-coupled receptor 125 |
| v1.09663; v1.08981 | K08463 | G protein-coupled receptor 126 |
| v1.24252 | K08464 | G protein-coupled receptor 128 |
| v1.02750 [+ 26 other sequence copies] | K08465 | G protein-coupled receptor 133 |
| v1.05774 [+ 11 other sequence copies] | K08466 | G protein-coupled receptor 144 |
| v1.05497; v1.13272; v1.01323 | K08436 | G protein-coupled receptor 152 |
| v1.08653 [+ 5 other sequence copies] | K08467 | G protein-coupled receptor 157 |
| v1.11807; v1.10392; v1.10394 | K08469 | G protein-coupled receptor 158 |
| v1.07294; v1.00247 | K08439 | G protein-coupled receptor 161 |
| v1.05167 | K08442 | G protein-coupled receptor 176 |
| v1.08677; v1.23465; v1.19865; v1.06986 | K12762 | G protein-coupled receptor GPR1 |
| v1.13395 | K08291 | G protein-coupled receptor kinase |
| v1.18529; v1.07599; v1.0558 | K12487 | G protein-coupled receptor kinase interactor 2 |
| v1.02481 | K04619 | G protein-coupled receptor family C group 5 member B |
| v1.22242 | K04622 | G protein-coupled receptor family C group 6 member A |
| v1.08625; v1.13650; v1.13048; v1.18694 | K04599 | G protein-coupled receptor Mth (Methuselah protein) |
| v1.07465; v1.10540 | K08341 | GABA(A) receptor-associated protein (autophagy-related protein B) |
| v1.09831 [+ 30 other sequence copies] | K05270 | Gamma-aminobutyric acid (GABA) receptor, invertebrate |
| v1.18702; v1.11701 | K05183 | Gamma-aminobutyric acid (GABA) A receptor beta-3 |
| v1.04252 [+ 6 other sequence copies] | K05185 | Gamma-aminobutyric acid (GABA) A receptor epsilon |
| v1.06325 | K05186 | Gamma-aminobutyric acid (GABA) A receptor gamma-1 |
| v1.00048 | K05188 | Gamma-aminobutyric acid (GABA) A receptor gamma-3 |
| v1.07506 [+ 6 other sequence copies] | K04615 | Gamma-aminobutyric acid (GABA) B receptor 1 |
| v1.07506 [+ 24 other sequence copies] | K04616 | Gamma-aminobutyric acid (GABA) B receptor 2 |
| v1.06426; v1.10563; v1.01138 | K05192 | Gamma-aminobutyric acid (GABA) receptor theta |
| v1.15485 | K05198 | Glutamate receptor, ionotropic, AMPA 2 |
| v1.09807 | K05200 | Glutamate receptor, ionotropic, AMPA 4 |
| v1.04764 | K05207 | Glutamate receptor, ionotropic, delta 2 |
| v1.15247 [+ 12 other sequence copies] | K05313 | Glutamate receptor, ionotropic, invertebrate |
| v1.15247 [+ 7 other sequence copies] | K05202 | Glutamate receptor, ionotropic, kainate 2 |
| v1.00617 | K05203 | Glutamate receptor, ionotropic, kainate 3 |
| v1.09688 [+ 6 other sequence copies] | K05208 | Glutamate receptor, ionotropic, N-methyl-D-aspartate 1 |
| v1.21204 [+ 4 other sequence copies] | K05212 | Glutamate receptor, ionotropic, N-methyl-D-aspartate 2D |
| v1.01622 | K05214 | Glutamate receptor, ionotropic, N-methyl-D-aspartate 3B |
| v1.01418 [+ 5 other sequence copies] | K05387 | Glutamate receptor, ionotropic, other eukaryote |
| v1.04275 | K05194 | Glycine receptor alpha-2 |
| v1.10737; v1.06885 | K05195 | Glycine receptor alpha-3 |
| v1.05498 | K05271 | Glycine receptor alpha-4 |
| v1.08900; v1.06885 | K05196 | Glycine receptor beta |
| v1.18634 | K05397 | Glycine receptor, invertebrate |
| v1.14569; v1.14570 | K09071 | Heart-and neural crest derivatives-expressed protein |
| v1.16783 [+ 4 other sequence copies] | K02168 | High-affinity choline transport protein |
| v1.13837 | K07608 | Internexin neuronal intermediate filament protein, alpha |
Table 5 Neuronal and sensory proteins in the predicted proteome of *A. digitifera* (Continued)

| v1.1671 | K04309 | Leucine-rich repeat-containing G protein-coupled receptor 4 |
| v1.09480; v1.05605 | K04308 | Leucine-rich repeat-containing G protein-coupled receptor 5 |
| v1.15300 [+ 8 other sequence copies] | K08399 | Leucine-rich repeat-containing G protein-coupled receptor 6 |
| v1.17524 [+ 14 other sequence copies] | K04306 | Leucine-rich repeat-containing G protein-coupled receptor 7 |
| v1.21700; v1.03578; v1.17196 | K04307 | Leucine-rich repeat-containing G protein-coupled receptor 8 |
| v1.16104 | K08396 | Mas-related G protein-coupled receptor member X |
| v1.08718; v1.02042; v1.02042 | K04604 | Metabotropic glutamate receptor 1/5 |
| v1.22794 [+ 7 other sequence copies] | K04605 | Metabotropic glutamate receptor 2/3 |
| v1.15331 | K04607 | Metabotropic glutamate receptor 4 |
| v1.01418 | K04608 | Metabotropic glutamate receptor 6/7/8 |
| v1.21698; v1.04544; v1.21739 | K14636 | MFS transporter, solute carrier family 18 (acetylcholine transporter) 3 |
| v1.05751; v1.19720; v1.22165; v1.02336 | K04134 | Muscarinic acetylcholine receptor |
| v1.11550 | K04129 | Muscarinic acetylcholine receptor M1 |
| v1.01913 [+ 4 other sequence copies] | K04131 | Muscarinic acetylcholine receptor M3 |
| v1.18723 | K04132 | Muscarinic acetylcholine receptor M4 |
| v1.08171 | K04133 | Muscarinic acetylcholine receptor M5 |
| v1.07408 [+ 34 other sequence copies] | K02583 | Nerve growth factor receptor (TNFR superfamily member 16) |
| v1.15265 [+ 91 other sequence copies] | K06491 | Neural cell adhesion molecule |
| v1.13789; v1.24010; v1.03980 | K09038 | Neural retina-specific leucine zipper protein |
| v1.24586; v1.16386; v1.16387 | K08052 | Neurofibrin 1 |
| v1.05520; v1.15407; v1.07590 | K04572 | Neurofilament light polypeptide |
| v1.19724 | K04573 | Neurofilament medium polypeptide (neurofilament 3) |
| v1.15787 [+ 4 other sequence copies] | K09081 | Neurogenin 1 (neurogenic differentiation protein) |
| v1.00345; v1.05338; v1.10997 | K08033 | Neurogenic differentiation factor 1 |
| v1.07355; v1.14517 | K09078 | Neurogenic differentiation factor 2 |
| v1.08832 | K09079 | Neurogenic differentiation factor 4 |
| v1.06678; v1.06677 | K01393 | Neurolysin |
| v1.16238 [+ 19 other sequence copies] | K06756 | Neuronal cell adhesion molecule |
| v1.20460; v1.16967 | K06757 | Neurofascin NFASC (cell adhesion molecule CAMs) |
| v1.22060; v1.03561 | K07525 | Neuronal guanine nucleotide exchange factor |
| v1.03908 | K09098 | Neuronal PAS domain-containing protein 1/3 |
| v1.00089 | K05247 | Neuropeptide FF-amide peptide |
| v1.21565 | K08375 | Neuropeptide FF receptor 2 |
| v1.06392 [+ 11 other sequence copies] | K04209 | Neuropeptide Y receptor, invertebrate |
| v1.08609 [+ 31 other sequence copies] | K06819 | Neurexlin 2 |
| v1.11492 [+ 5 other sequence copies] | K03308 | NeurotransmitterNa+ symporter, NSS family |
| v1.16744 [+ 8 other sequence copies] | K06774 | Neurotrophin |
| v1.05353 | K03176 | Neurotrophic tyrosine kinase receptor type 1 |
| v1.20055 | K04360 | Neurotrophic tyrosine kinase receptor type 2 |
| v1.03803 | K04356 | Neurotrophin 3 |
| v1.09523 | K04803 | Nicotinic acetylcholine receptor alpha-1 (muscle) |
| v1.11940 | K04806 | Nicotinic acetylcholine receptor alpha-4 |
| v1.01548 | K04808 | Nicotinic acetylcholine receptor alpha-6 |
| v1.05056; v1.12097 | K04809 | Nicotinic acetylcholine receptor alpha-7 |
| v1.07222; v1.11069 | K04810 | Nicotinic acetylcholine receptor alpha-9 |
increasing the saturation state of CaCO3 to sustain calcium precipitation [146]. Importantly, located also at the calicoblastic membrane is carbonic anhydrase [149] which is required to catalyse the intermediate step of calcification by the reversible hydration of carbon dioxide (CO2 + H2O \( \Rightarrow \) HCO3\(^{-}\) + H\(^{+}\)). In coral phototrophic symbiosis, despite numerous studies describing the well-known phenomenon of light-enhanced calcification, the relationship linking symbiont photosynthesis to coral calcification has been elusive [150,151]. Nonetheless, efforts to better understand the calcifying response of scleractinian corals to environmental change and ocean acidification are gaining traction [149,152,153].

Voltage-gated calcium channels (VGCCs) have been examined extensively in mammalian physiology for converting membrane potential into intracellular Ca\(^{2+}\) transients for signalling transduction pathways (reviewed in [154]). VGCC signalling affects cellular processes to include muscle contraction, neuronal excitation, gene transcription, fertilisation, cell differentiation and development, proliferation, hormone release, activation of calcium-dependent protein kinases, cell death via necrosis and apoptosis pathways, phagocytosis and endo/exocytosis. Remarkably, annotation of the genome of A. digitifera reveals sequences encoding homologues of all the VGCC (\(\alpha\), \(\alpha\delta\), \(\beta\), and \(\gamma\)) subunits of the molecular (L,
N, P/Q and R) phenotypes expressed in mammalian physiology (Table 6). There are multiple sequences encoding three variants of Ca\(^{2+}\)-transporting ATPase, of which at least one is necessary for coral calcification. There is only one sequence match for expressing carbonic anhydrase in the genome of A. digitifera, which may reflect the high catalytic efficiency of this catalyzing enzyme [155], although a BLAST search of ZophyteBase does reveal scaffolds with low e-values which on future experimental inspection might uncover multiple copies of this enzyme essential for calcification. There are multiple sequences that express solute carrier Na\(^+\)/Ca\(^{2+}\) and Na\(^+\)/K\(^+\)/Ca\(^{2+}\)-exchange families of transport proteins with expression of the coral Ca\(^{2+}\)/H\(^+\)-antiporter may regulate cellular pH and Ca\(^{2+}\) homeostasis.

Implicit to coral calcification is Ca\(^{2+}\) regulation that affects signalling of other vital cellular functions. Cellular Ca\(^{2+}\) is mediated by the calcium-sensing receptor calmodulin (18 sequence matches) and other messenger calcium-binding effectors (Table 6), including the calcium-binding protein CML (40 protein domain sequence matches). Calcium/calmodulin-protein kinase proteins are arguably key to Ca\(^{2+}\)-signalling in coral symbiosis but, with the exception of activation of sperm flagellar motility [156], their precise role has not been elaborated.

**Plant-derived proteins**

Endosymbiosis has contributed greatly to eukaryotic evolution, most notably to the genesis of plastids and mitochondria derived from prokaryotic antecedents. Genetic integration by endosymbiotic-to-host transfer (EGT) or replacement (EGR) has been a significant force in early metazoa innovation, whereby nuclear transferred genes may even adopt novel functions in the host cell or replace existing versions of the protein that they encode [157]. Prokaryote-to-eukaryotic gene transfer has been widespread in evolution, but examples of genetic exchange between unrelated eukaryotes, such as between algal symbionts and their multicellular eukaryote host, are considered rare (reviewed by [158,159]). One such example is aroB (3-dehydroquinate synthase) transferred to the genome of the sea anemone N. vectensis, which sequence best fits that of the dinoflagellate Oxyrrhis marina [119]. Close inspection of the amino acid sequence of the aroB gene product, as reported by Shinzato et al. [45], clearly shows this protein to be 2-epi-5-epi-valioline synthase (EVS), a sugar phosphate cyclase orthologue that catalyses the conversion of sedoheptulose 7-phosphate to 2-epi-5-epi-valioline found to be a precursor of the mycosporine-like amino acid (MAA) sunscreen shionorine in the cyanobacterium Anabaena variabilis [160]. Additionally, the EVS gene of N. vectensis has a distinctive O-methyltransferase fusion that is identical in O. marina [161]. The shikimate pathway is essential to apicomplexan parasites of the genera Plasmodium, Toxoplasma and Cryptosporidium and of Tetrahymena ciliates to express a pentafunctional aroM gene similar to that of Ascomycetes, which is thought to have been conveyed by fungal gene transfer to a common ancestral progenitor [162]. In a separate example, H. viridis expresses a plant-like ascorbate peroxidase gene (HvAPX1) during oogenesis in both symbiotic and asposymbiotic individuals [163], whereby peroxidase activity is coincident with oogenesis and embryo genesis that in Hydra acts as a ROS scavenger to protect the oocyte from apoptotic degradation [164]. The sacoglossan (sea slug) molluscs Elysia chlorotica and E. viridis (Plakobranchiidae) acquire plastids on ingestion of the siphonaceous alga Voucherea litorea (termed “kleptoplasty”) and, by maintaining sequestered plastids in an active photosynthetic state, has emerged as a model organism for the transfer of nuclear-encoded plant genes from algal symbiont to its animal host [165]. In this symbiosis, the family of light-harvesting genes psbO, prk (phosphoribokinase) and chlorophyll synthase (chlG) are entrained in the genome of Elysia chlorotica (reviewed in [166,167]), although there is debate whether these genes are transcriptionally expressed (compare [168] and [169]). Also, phylogenomic analysis of the predicted proteins of the asposymbiotic unicellular chaoflagellate Monosiga brevicollis, considered to be a stem progenitor of the animal kingdom [170,171], reveals 103 genes having strong algal affiliations arising from multiple phototrophic donors [172]. Such notable examples illustrate the transfer of algal genes to animal recipients.

KEGG orthology-based annotation of the predicted proteome of A. digitifera reveals a plethora of sequences presumed to be of algal origin (Table 7). Like E. chlorotica, the coral genome has encoded the photosystem II (PSII) protein PsbO of the oxygen-evolving complex of photosynthesis, as well as the PSII light-harvesting complex protein PsbL that is important in protecting PSII from photo-inactivation [173]. Encoded also are the photosystem I subunit proteins Psal and Psao. Additionally encoded are the photosystem P840 reaction center cytochrome c551 (PscC) protein and the photosynthetic reaction center M subunit protein, the light-harvesting proteins complex 1 alpha (PufA), the complex II chlorophyll a/b binding protein 6 (LHCB6), the cyanobacterial phycobilisome proteins AcpF and AcpG, the phycocyanin-associated antenna protein CpcD, the phycocyanobilin laise protein CpcF and the phycoerythrin-associated linker protein CpeS. Like E. chlorotica, the coral genome encodes chlorophyll synthase (ChlG), a chlorophyll transporter protein PucC, a light-independent nitrogenase-like protoclorophyllide reductase enzyme that is sensitive to oxygen [174] and a red chlorophyll reductase essential to the detoxification of photodynamic chlorophyll catabolites arising from plant/algal senescence [175]. Three
chlorosome proteins of the photosynthetic antenna complex of green sulphur bacteria, a bacteriochlorophyll methyltransferase involved in BChl c biosynthesis [176] and the retinylidene bacteriorhodopsin of phototrophic Archaea are also encoded in the coral genome. Present are genes encoding subunit 6 of the cytochrome B6f complex that links PSII and PSI via the plastoquinone pool, together with chloroplast ferredoxin-like NapH and NapG proteins and their 2Fe-2S cluster protein. The coral genome, however, encodes sequences for NAD+-ferredoxin

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|----------------------------|
| v1.06452; v1.06451; v1.24424; v1.16923 | K07300 | Ca2+-H+ antiporter |
| v1.01669 [+ 9 other sequence copies] | K01537 | Ca2+-transporting ATPase |
| v1.22367; v1.22366; v1.22265 | K05850 | Ca2+ transporting ATPase, plasma membrane |
| v1.19074 | K05853 | Ca2+ transporting ATPase, sarcoplasmic/endoplasmic reticulum |
| v1.22416; v1.22417; v1.15682; v1.00750 | K14757 | Calbindin D28 |
| v1.24568 [+ 9 other sequence copies] | K01672 | Carbonic anhydrase |
| v1.09241 | K08272 | Calcium binding protein 39 |
| v1.02323 [+ 39 other sequence copies] | K13448 | Calcium-binding protein CML |
| v1.05162 [+ 21 other sequence copies] | K13412 | Calcium-dependent protein kinase |
| v1.09352 | K07359 | Calcium/calcmodulin-dependent protein kinase kinase |
| v1.06475; v1.07555; v1.00945; v1.00159; v1.21122 | K08794 | Calcium/calcmodulin-dependent protein kinase I |
| v1.06475; v1.01061; v1.21150; v1.22443 | K04515 | Calcium/calcmodulin-dependent protein kinase II |
| v1.00159 | K05869 | Calcium/calcmodulin-dependent protein kinase IV |
| v1.21927; v1.01218; v1.22226; v1.06623; v1.13703 | K06103 | Calcium/calcmodulin-dependent serine protein kinase |
| v1.13460 | K08284 | Calcium channel MID1 |
| v1.20738; v1.01401 | K12841 | Calcium homeostasis endoplasmic reticulum protein |
| v1.22794 [+ 11 other sequence copies] | K04612 | Calcium-sensing receptor |
| v1.10079 [+ 17 other sequence copies] | K02183 | Calmodulin |
| v1.10994 | K14734 | S100 calcium binding protein G |
| v1.02488 [+ 14 other sequence copies] | K05849 | Solute carrier family 8 (sodium/calcium exchanger) |
| v1.23153 [+ 9 other sequence copies] | K13749 | Solute carrier family 24 (sodium/potassium/calcium exchanger) |
| v1.14863 | K12304 | Soluble calcium-activated nucleotidase 1 |
| v1.18656 [+ 13 other sequence copies] | K04858 | Voltage-dependent calcium channel alpha-2/delta-1 |
| v1.13222 | K04860 | Voltage-dependent calcium channel alpha-2/delta-3 |
| v1.08078 [+ 9 other sequence copies] | K05515 | Voltage-dependent calcium channel alpha 1, invertebrate |
| v1.03896 [+ 6 other sequence copies] | K05316 | Voltage-dependent calcium channel alpha-2/delta, invertebrate |
| v1.04798 | K05317 | Voltage-dependent calcium channel beta, invertebrate |
| v1.22788 | K04863 | Voltage-dependent calcium channel beta-2 |
| v1.09999 | K04872 | Voltage-dependent calcium channel gamma-7 |
| v1.02505 | K04873 | Voltage-dependent calcium channel gamma-8 |
| v1.03648 [+ 6 other sequence copies] | K04850 | Voltage-dependent calcium channel L type alpha-1C |
| v1.03648; v1.117267 | K04851 | Voltage-dependent calcium channel L type alpha-1D |
| v1.03648; v1.13219; v1.21895 | K04857 | Voltage-dependent calcium channel L type alpha-1S |
| v1.06313; v1.01656; v1.23096 | K04344 | Voltage-dependent calcium channel P/Q type alpha-1A |
| v1.08078 [+ 10 other sequence copies] | K04849 | Voltage-dependent calcium channel N type alpha-1B |
| v1.07968 | K04852 | Voltage-dependent calcium channel R type alpha-1E |
| v1.01364; v1.13467; v1.08705 | K04854 | Voltage-dependent calcium channel T type alpha-1G |
| v1.15414; v1.14241; v1.09595 | K04855 | Voltage-dependent calcium channel T type alpha-1H |

Dunlap et al. BMC Genomics 2013, 14:509
http://www.biomedcentral.com/1471-2164/14/509
Table 7 Plant-derived proteins in the predicted proteome of A. digitifera

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.14452      | K09843         | (+)-Abscisic acid 8′-hydroxylase |
| v1.18658      | K14496         | Abscisic acid receptor PYR/PYL family (PYL) |
| v1.21983; v1.05890 | K03342         | p-Aminobenzoate synthetase / 4-amino-4-deoxychorismate lyase (FabBC) |
| v1.15436      | K02822         | Ascorbate-specific IIB component, PTS system (PTS-Ula-EiiB) |
| v1.11187; v1.13966 | K00423         | L-Ascorbate oxidase |
| v1.20081; v1.22465 | K13604         | Bacteriochlorophyll C20 methyltransferase (BchU) |
| v1.07723      | K04641         | Bacteriorhodopsin (BoP) |
| v1.21858      | K04040         | Chlorophyll synthase (ChlG) |
| v1.01742      | K08945         | Chlorosome envelope protein A (CsmA) |
| v1.04797; v1.14208 | K08946         | Chlorosome envelope protein B (CsmB) |
| v1.18698      | K08948         | Chlorosome envelope protein D (CamD) |
| v1.18637      | K02642         | Cytochrome b6f complex subunit 6 (PetL) |
| v1.21101; v1.14192; v1.14548 | K01735         | 3-Dehydroquinate synthase (AroB) |
| v1.05796      | K10210         | 4,4′-Diaponeurosporene oxidase (carotenoid biosynthesis; CrtP) |
| v1.11730      | K04755         | Ferredoxin, 2Fe-2S (FdX) |
| v1.19154; v1.00014 | K00532         | Ferredoxin hydrogenase |
| v1.17698; v1.06031; v1.16647 | K02574         | Ferredoxin-type protein (NapH) |
| v1.23058      | K02573         | Ferredoxin-type protein (NapG) |
| v1.08414      | K08926         | Light-harvesting complex 1 alpha chain (PufA) |
| v1.21458      | K08917         | Light-harvesting complex II chlorophyll a/b binding protein 6 (LHCB6) |
| v1.03743      | K08226         | MFS transporter, BCD family, chlorophyll transporter (PucC) |
| v1.13030; v1.08678 | K13413         | Mitogen-activated protein kinase kinase 4/5, plant ((MKK4_5P) |
| v1.02429; v1.10744; v1.03340 | K08929         | Photosynthetic reaction center M subunit (PufM) |
| v1.03631      | K02696         | Photosystem I subunit VIII (PsaI) |
| v1.11432      | K14332         | Photosystem I subunit (PsaO) |
| v1.17422      | K02713         | Photosystem II protein (PsbL) |
| v1.18303      | K02716         | Photosystem II oxygen-evolving enhancer protein 1 (PsbO) |
| v1.12300; v1.21136 | K08942         | Photosystem P840 reaction center cytochrome c551 (PscC) |
| v1.00280      | K02097         | Phycobilisome core component 9 (AcpF) |
| v1.10967      | K02290         | Phycobilisome rod-core linker protein (AcpG) |
| v1.02166      | K02287         | Phycocyanin-associated, rod protein (CpeD) |
| v1.19642; v1.07305; v1.19572; v1.01248 | K02289         | Phycocyanobilin lyase beta subunit (CpcF) |
| v1.10441      | K05382         | Phycoerythrin-associated linker protein (CpeS) |
| v1.13406      | K10027         | Phytoene dehydrogenase (desaturase; CrtI) |
| v1.18809; v1.06199 | K02291         | Phytoene synthase (CrtB) |
| v1.20411; v1.102037; v1.14064; v1.21095 | K09060         | Plant G-box-binding factor (GBF) |
| v1.10035      | K00218         | Protochlorophyllide reductase [NfEN-like; Por] |
| v1.21846      | K05358         | Quinate dehydrogenase (QuiA) |
| v1.03127      | K13545         | Red chlorophyll catabolite reductase (ACD2) |
| v1.05899      | K00891         | Shikimate kinase (AroK, AroL) |
| v1.21101; v1.14192; v1.05899 | K13829         | Shikimate kinase / 3-dehydroquinate synthase (AroKB) |
| v1.12938      | K08500         | Syntaxin of plants (SYP6) |
reductase (HcaD; not tabulated), rather than the required NADP+-ferredoxin reductase of photosynthesis. Annotation of the *A. digitifera* genome revealed genes unexpectedly encoding ferredoxin hydrogenase [EC:1.12.7.2] and that of its small subunit protein (Table 7) involved in light-dependent production of molecular hydrogen having its [Fe-Fe]-cluster coupled to the photosynthetic transport chain via a charge-transfer complex with ferredoxin (see [177]).

Like *N. vectensis* and the dinoflagellate *Oxyrrhis marina*, the genome of *A. digitifera* encodes an O-methyltransferase which is immediately downstream of EVS, but the two genes are not fused. Using a ZoophyteBase BlastP search, the O-methyltransferase showed little sequence homology with the corresponding protein of *A. variabilis* (e-value of 6.972E-2 and Bit score of 34.27), whereas the EVS protein shared 87% absolute sequence identity to the *A. variabilis* EVS protein. What role, if any, these two genes play in mycosporine-like amino acid (MAA) biosynthesis in *A. digitifera* has yet to be determined, although it has been suggested from the transcriptome of *Acropora microphthalma* that MAA biosynthesis proceeds from a branch point at 3-dehydroquinate of the shikimic acid pathway as a shared metabolic adaptation between the coral host and its symbiotic zooxanthellae [40]. The 3-dehydroquinate synthase enzyme of the shikimic acid pathway, thought to be a key intermediate in an alternative MAA biosynthetic pathway in *A. variabilis* [178], is instead encoded by the fused *aroKB* gene of *A. digitifera* (Table 7). Additional shikimate proteins of the predicted proteome, although not limited to phototrophs, are shikimate kinase (*Arok*), quinate dehydrogenase (*QuiA*) and the conjoined *p*-aminobenzoate synthase and 4-amino-4-deoxychlorismate lysate (*PabBC*) enzyme necessary for folate biosynthesis [179]. Other plant-related gene homologues include the phytohormone abscisic acid receptor protein (*PabBC*) and its cytochrome P450 monooxygenase abscisic acid 8'-hydroxylase, L-ascorbate oxidase and PTS system degrading enzymes, the unique SYP6 and SYP7 syntaxes of plant vesicular transport, tocopherol cyclase and a tocopherol O-methyltransferase enzyme that converts γ-tocopherol to α-tocopherol. Essential for carotene biosynthesis are phytoene synthase (*CrtB*) and phytoene dehydrogenase (*CrtI*) enzymes. Significantly, encoded within the coral genome is zeaxanthin epoxidase that is essential for abscisic acid biosynthesis and is a key enzyme in the xanthophyll cycle of plants and algae to impart oxidative stress tolerance.

Given that viruses often mediate gene transfer processes, it is intriguing that certain bacteriophages of marine *Synechococcus* and *Prochlorococcus* cyanobacteria are reported to carry genes encoding the photosynthesis D1 (*psbA*), and D2 (*psbD*) proteins, a high-light inducible protein (HLIP) [180,181] and the photosynthetic electron transport plastocyanin (*petE*) and ferredoxin (*petF*) proteins thought to enhance the photosynthetic fitness of their host [182-184]. Accordingly, it has been suggested that the transfer of *psbA* by viruses associated with *Symbiodinium* could lessen the severity of thermal impairment to PSII and the response of corals to thermal bleaching [185]. It is yet unknown if phages or dinoflagellate-infecting viruses [186], particularly those of *Symbiodinium* [187], may affect gene transfer leading to complementary (or “shared”) metabolic adaptations of symbiosis [119,188].

**Proteins of nitrogen metabolism**

It is well accepted that intracellular *Symbiodinium* spp. provide reduced carbon for coral heterotrophic metabolism by photosynthetic carbon fixation. Because of this metabolic relationship, light is a critical feature in the bioenergetics of coral symbiosis [189]. The algal photosynthetic electron transport to corals, however, is deficient in nitrogen at levels necessary to sustain autotrophic growth. While corals can assimilate fixed nitrogen from their surroundings seawater [190], “recycled” nitrogen within the symbiosis may account for as much as 90% of the photosynthetic nitrogen demand [191]. It would not be surprising then that light would have a strong influence on the uptake and retention of ammonium by symbiotic corals. Consequently, corals excrete excess ammonium in darkness [192], and in light excretion is induced by treatment with the photosynthetic electron transport inhibitor 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) [193]. Since ammonia is the product of nitrogen fixation, these observations suggest that the coral holobiont may fix nitrogen in the dark, or when photosynthesis is repressed, during which coral tissues are hypoxic [194], and nitrogenase activity is not inactivated by molecular oxygen [195].

Tropical coral reefs are typically surrounded by low-nutrient oceanic waters of low productivity but, paradoxically, the waters of coral reefs often have elevated levels of inorganic nitrogen [196,197] attributed to high rates of nitrogen fixation. While nitrogen fixation from diazotrophic
epiphytes of the coral reef substrata and sediments [197,198] and diazotrophic bacterioplankton of the coral reef lagoon [199] provide substantial quantities of fixed nitrogen for assimilation by the coral reef, mass-balance estimates show this input to be less than the community's annual nitrogen demand [200]. Endolithic nitrogen-fixing bacteria are abundant in the skeleton of living corals where they benefit from organic carbon excreted by overlaying coral tissues to provide a ready source of energy for dinitrogen reduction [201]. Additionally, intracelluar nitrogen-fixing cyanobacteria are reported to coexist with dinoflagellate symbionts in the tissues of *Monastraea cavernosa* and to functionally express nitrogenase activity [202]. Corals also harbour a diverse assemblage of heterotrophic microorganisms in their skeleton, tissues and lipid-rich mucus (reviewed in [203]), and these communities include large populations of diazotrophic bacteria [204,205], and archaea [206]. Apart from nitrogen fixation, the coral microbiota contributes to other nitrogen-cycling processes, such as nitrification, ammonification and denitrification [207,208]. We were surprised to find several nitrogen fixation and cycling proteins encoded in the genome of *A. digitifera* (Table 8), notably a nitrogen fixation NifU-like protein, the Nif-specific regulatory protein (NifA), the regulatory NAD(+)dinitrogen-reductase ADP-D-ribosyltransferase protein, a nitrifying ammonia monoxygenase enzyme and nitrate reductase, which are usually expressed only by prokaryotic microorganisms.

The presence of genes encoding proteins involved in nitrogen fixation raises speculation that corals may contribute directly to, or perhaps co-regulate, certain processes that catalyse the reduction of dinitrogen (N₂) to ammonia (NH₃) by the enzyme nitrogenase reductase (NiFH). The functional NiFH enzyme is a binary protein composed of a molybdenum-iron (MoFe) protein (NiFB/ NiFDK), or its NiFEN homologue, fused with a FeMo-cofactor (FeMoco) protein [209]. While genes encoding NiFB, NiFDK (or NiFEN) and their FeMo-cofactor do not appear in the genome of *A. digitifera*, a gene encoding the NiFEN-like protein protoclorophyllide oxidoreductase (POR) is present (Table 8). POR has all three subunits with high similarity to the assembled MoFe nitrogenase [210], but this homologue is unlikely to be effective in nitrogen reduction [211,212] since its activity is light dependent [213] when tissues are highly oxic [193]. The NiFU protein encoded in the coral genome preassembles the metalloccatalytic Fe-S clusters for maturation of nitrogenase [214], but its assemblage without NifS, a cysteine desulfurase needed for [Fe-S] cluster assembly [215], would be incomplete, and its pre-nitrogenase receptor is also missing. Yet, the coral does have the *nif* gene that encodes pyruvate:flavodoxin oxidoreductase required for electron transport in nitrogenase reduction [216]. The regulatory NifA protein encoded in the coral genome might activate, on stimulation by the integration host factor (INF), transcription of nitrogen fixation (*nif*) operons of RNA polymerase [217], and both of these proteins are encoded in the coral genome. Additional to this transcriptional control, post-translational nitrogenase activity is controlled by reversible ADP-ribosylation of a specific arginine residue in the nitrogenase complex [218]. NAD(+)dinitrogen-reductase ADP-D-ribosyltransferase (DraT) inactivates the nitrogenase complex while ADP-ribosylglycohydroxolase (DraG) removes the ADP-ribose moiety to restore nitrogenase activity, and both of these enzymes are encoded in the coral genome. Given that genes encoding essential constituent proteins of nitrogenase assembly appear incomplete, corals are unlikely to fix nitrogen *per se*, but co-opted elements of the coral genome to regulate processes of nitrogen fixation by its diazotrophic consortia is a prospect worthy of exploration [219].

Nitrofying/nitrifying bacteria and archaea express the enzyme ammonia monooxygenase that converts fixed ammonia to nitrite (via hydroxylamine) and the enzyme nitrite (oxido)reductase completes the oxidation of nitrite to nitrate, and both of these enzymes are entrained in the genome of *A. digitifera* (Table 8). The ammonia monooxygenase subunit A (*amoA*) of archaeal consort has been described in nine species of coral from four reef locations [220], but the presence of *amoA* in the coral genome, together with encoded ammonium transport proteins, was not anticipated. Another protein of prokaryotic origin encoded in the coral genome is nitrate reductase (periplasmic, assimilatory and respiratory), the latter being required for anaerobic respiration by bacteria [221], and unlike the nitrate reductase family of sulphite oxidase enzymes in eukaryotes, the nitrate reductases of prokaryotes (*K00363*) belong to the DMSO reductase family of enzymes. Also encoded in the coral genome are a nitrite transporter (NirC) and a formate-dependent nitrite reductase (*NrfA*) required for nitrite ammonification [222]. In addition to nitrite reduction, *NrfA* reduces nitric oxide, hydroxylamine, nitrous oxide and sulphite, the last providing a metabolic link between nitrogen and sulphur cycling in coral metabolism. Other enzymes of nitrogen metabolism encoded in the coral genome are the carbamoyl-phosphate synthase family of enzymes [223] that catalyses the ATP-dependent synthesis of carbamoyl phosphate used for the production of urea (ornithine cycle) to provide a ready store of fixed-N in the urea-nitrogen metabolism of corals [224]. Another nitrogen source comes from glutamate dehydrogenase (GDH) that reversibly converts glutamate to α-ketoglutarate with liberation of ammonia, and as expected [225], this enzyme is encoded in the coral genome, together with the prokaryotic nitrogen regulatory protein PII of glutamine synthase, which in bacteria is activated in response to nitrogen availability. Encoded also is
histidine ammonia-lyase (histidase) that liberates ammonia (and urocanic acid) from cytosolic stores of histidine. It is now accepted that uric acid deposits accumulated by symbiotic algae provide a significant store of nitrogen for the coral holobiont [226], so it is noteworthy that the coral genome encodes urate oxidase (uricase) to catalyse uric acid oxidation to allantoin from which urea and ureidoglycolate are produced in a reaction catalysed by allantoicase (allantoate amidohydrolase), both of which known isoforms are present in the coral genome. Encoded in the coral genome is also urease to catalyse the hydrolysis of urea, presumably excreted by its algal symbionts, with the release of carbon dioxide and ammonia to meet the nitrogen demand of the coral holobiont during periods of low nitrogen availability. Similarly, xanthine dehydrogenase (xanthine: NAD⁺-oxidoreductase) acts by oxidation on a variety of purines.
including hypoxanthine, to yield urate for the recycling of nitrogen in coral nutrition. Many of the aforementioned proteins of nitrogen metabolism, including Nif proteins, have been detected in the proteome of an endosymbiont-enriched fraction of the coral *S. pistillata* [39].

Notwithstanding consideration of the rapid diffusion rate of nitric oxide (NO) or its apparent short biological half-life [227], there is debate about the provenance of endogenously produced NO in signalling the bleaching of corals in response to environmental stress. Elevated nitric oxide synthase (NOS) activity and NO production in algal symbionts has been attributed to the thermal stress response of corals [228,229], whereas the host is ascribed to be the major source of NO during exposure to elevated temperature [230,231]. While our annotation may not resolve this dispute, we show (Table 8) that nitric oxide synthase enzymes (Nor D, Nor E, Nor Q and an invertebrate NOS protein) are encoded in the genome of *A. digitifera*, together with a nitric oxide-interacting protein (NOIP) that in higher animals regulates neuronal NOS activity [232]. Nitric oxide is an intermediate of nitrite reduction catalysed by nitrite reductase (NIR), which by further reduction produces ammonia. The coral genome also encodes nitric oxide dioxygenase (NOD) that converts nitric oxide to nitrate. Accordingly, enhanced expression of NIR (NO reduction) or NOD (NO oxidation) could ameliorate the NO-signalling response of coral bleaching presumed activated by environmental stress.

**DNA repair**

Cellular DNA is prone to damage caused by the products of normal metabolism and by exogenous agents. Damage to DNA from metabolic processes include the oxidation of nucleobases and strand interruptions by the production of reactive oxygen species (ROS), from alkylation of nucleotide bases, from the hydrolysis of bases causing deamination, depurination and depyrimidination, and from the mismatch of base pairs from errors in DNA replication. Damage affected by external agents include exposure to UV light causing pyrimidine dimerization and free radical-induced damage, exposure to ionising radiation causing DNA strand breaks, thermal disruption causing hydrolytic depurination and single-strand breaks, and by xenobiotic contamination to cause DNA adduct formation, nucleobase oxidation and DNA crosslinking. Most of these lesions affect structural changes to DNA that alter or prevent replication and gene transcription at the site of DNA damage. Thus, recognition and repair of DNA abnormalities are vital processes essential to maintain the genetic integrity of the coral genome. Since there are multiple pathways causing DNA damage at diverse molecular sites, there are likewise diverse and overlapping processes available to repair cellular DNA damage. Of the many nuclear repair processes, photoreactivation (photolyase), base excision repair and nucleotide excision repair are the main elements for the repair of cellular DNA damage.

Exposure to sunlight is an absolute requirement for phototrophic symbiosis, but excessive exposure of corals to solar ultraviolet radiation can inflict direct damage to DNA by pyrimidine dimerization and 6-4 photoduct formation and cause indirect damage by the production of ROS to initiate free-radical damage. While there have been abundant studies on the sensitivity of corals to solar ultraviolet radiation, only a few have examined the effects of solar UV to cause DNA damage. Photoreactivation has been shown to be an important repair pathway for reversing UV-activated DNA damage in adult coral [233] and coral planulae [234]. UV damage to DNA was first demonstrated by the detection of unrepaired cyclobutane pyrimidine dimers (CPDs) in the host tissues and algal symbionts of the coral *Porites porites*, in which CPDs had increased in a UV dose-dependent manner [235], whereas CPDs and 6-4 pyrimidine-pyrimidone photodadducts in the coral *Montipora verrucosa* holobiont were correlated inversely with levels of coral “sunscreen” protection [236]. The effects of solar UV radiation causing DNA lesions in coral have been determined by use of the comet assay [237], and UV-induced DNA damage and repair has been examined in the symbiotic anemone *Aiptasia pallida* [238]. The comet assay showed also that DNA lesions in coral planulae had increased on acquiring algal symbionts, presumably from greater ROS production resulting as a by-product of photosynthesis [239]. Iron-induced oxidative stress was found likewise to enhance DNA damage in the coral *Pocillopora damicornis* as determined by the occurrence of DNA apurinic/apyrimidinic sites caused by hydrolytic lesions [240]. Significantly, DNA damage in the host and algal symbionts of the coral *Montastraea faveolata* was found to occur simultaneously during thermal “bleaching” stress, and DNA damage is further enhanced on exposure to greater irradiances of solar radiation [241]. Nevertheless, despite the serious risk of unrepaired DNA damage to coral survival, the DNA repair processes of corals to mitigate the detrimental effects of environmental stress have not been adequately characterised at the transcriptome level of expression [29,242].

Our annotation of the sequenced genome of *A. digitifera* has revealed genes encoding a large repertoire of DNA repairing enzymes and their adaptor proteins (Table 9). Given strong evidence for DNA photoreactivation in corals having been reported [233,234], it was surprising to find only one gene in single copy that encodes a sole photolyase enzyme for reversing pyrimidine dimer and 6-4 photoduct formation. Notably, we found genes encoding 6 members of the ERCC family of nucleotide excision repair enzymes, together with the UV excision repair protein RAD23, for the repair of UV-induced DNA damage.
| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.02961; v1.13402 | K03575 | A/G-specific adenine glycosylase (MutY) |
| v1.11766 | K03919 | Alkylated DNA repair protein |
| v1.04821 | K10765 | Alkylated DNA repair protein alkB homologue 1 |
| v1.02479 | K10766 | Alkylated DNA repair protein alkB homologue 4 |
| v1.20302 | K10767 | Alkylated DNA repair protein alkB homologue 5 |
| v1.24450 | K10768 | Alkylated DNA repair protein alkB homologue 6 |
| v1.02766; v1.09413 | K10770 | Alkylated DNA repair protein alkB homologue 8 |
| v.1.01590 [+ 4 other sequence copies] | K10884 | ATP-dependent DNA helicase 2 subunit 1 |
| v1.18810; v1.03166; v1.08449 | K10885 | ATP-dependent DNA helicase 2 subunit 2 |
| v1.08013 | K03722 | ATP-dependent DNA helicase DinG |
| v1.03542 | K14635 | ATP-dependent DNA helicase MPH1 |
| v1.06737 [+ 5 other sequence copies] | K15255 | ATP-dependent DNA helicase PIF1 |
| v1.17360; v1.21235 | K10899 | ATP-dependent DNA helicase Q1 |
| v1.01081 [+ 8 other sequence copies] | K10730 | ATP-dependent DNA helicase Q4 |
| v1.16859 | K10902 | ATP-dependent DNA helicase Q5 |
| v1.11661 [+ 19 other sequence copies] | K03654 | ATP-dependent DNA helicase RecQ |
| v1.20397 | K03656 | ATP-dependent DNA helicase Rep |
| v1.18049; v1.07731; v1.05830 | K10905 | ATR interacting protein |
| v1.01679 | K01669 | Deoxyribodipyrimidine photo-lyase |
| v1.03410; v1.12968; v1.00865; v1.16876 | K10887 | DNA cross-link repair 1C protein |
| v1.07474; v1.07473; v1.01809 | K10610 | DNA damage-binding protein 1 |
| v1.13116; v1.03378; v1.16328 | K10140 | DNA damage-binding protein 2 |
| v1.17099 [+ 5 other sequence copies] | K11885 | DNA damage-inducible protein 1 |
| v1.05469 | K06663 | DNA damage checkpoint protein |
| v1.02859; v1.14719; v1.21030; v1.10920 | K04452 | DNA damage-inducible transcript 3 |
| v1.02191 | K10844 | DNA excision repair protein ERCC-2 |
| v1.19108 [+ 5 other sequence copies] | K10843 | DNA excision repair protein ERCC-3 |
| v1.22267 [+ 4 other sequence copies] | K10848 | DNA excision repair protein ERCC-4 |
| v1.15137 [+ 5 other sequence copies] | K10846 | DNA excision repair protein ERCC-5 |
| v1.18550; v1.02606; v1.14935; v1.08831 | K10841 | DNA excision repair protein ERCC-6 |
| v1.20045; v1.01844; v1.11724; v1.03203 | K10570 | DNA excision repair protein ERCC-8 |
| v1.15430; v1.03058 | K03658 | DNA helicase IV |
| v1.00228 [+ 4 other sequence copies] | K11665 | DNA helicase INO80 |
| v1.00136; v1.0678; v1.21529 | K10776 | DNA ligase 3 |
| v1.23293; v1.19418; v1.23430; v1.15721 | K10777 | DNA ligase 4 |
| v1.19248 | K07458 | DNA mismatch endonuclease, patch repair protein |
| v1.19011 | K08739 | DNA mismatch repair protein MLH3 |
| v1.11513; v1.11449 | K08735 | DNA mismatch repair protein MSH2 |
| v1.14781 | K08736 | DNA mismatch repair protein MSH3 |
| v1.05696; v1.22444; v1.19162 | K08740 | DNA mismatch repair protein MSH4 |
| v1.04904 | K08741 | DNA mismatch repair protein MSH5 |
| v1.15360; v1.19426; v1.08585 | K08737 | DNA mismatch repair protein MSH6 |
| v1.02429 [+ 8 other sequence copies] | K03572 | DNA mismatch repair protein MutL |
More abundant are the DNA mismatch repair enzymes from the MLH, MSH, Mut and PMS protein families and related glycosylase/lyase proteins for repairing erroneous insertion, deletion and mis-incorporation of bases to arise during DNA replication and recombination. There is additionally a specific gene that encodes a 3′-endonuclease protein that has a preference to correct mispaired nucleotide sequences. Abundant also are other members of the RAD-family of DNA repair proteins, including 28 sequence copies of a gene encoding the RAD50 protein for DNA double-strand break repair that, together with members of the MRE, Rec, REV, Swi5/Sae3, XRCC and XRS families of recombination and polymerase proteins, have complementary roles in DNA repair. Apparent also in the genome are the DNA helicase proteins, including RuvB-like proteins, which are primarily involved in DNA replication and transcription, but assist also in the repair of DNA damage by separating double strands at affected sites of DNA damage.

Table 9 DNA repair proteins in the predicted proteome of A. digitifera (Continued)

| Accession | Description |
|-----------|-------------|
| v1.03990  | K03555 DNA mismatch repair protein MutS |
| v1.14015  | K07456 DNA mismatch repair protein MutS2 |
| v1.08443  | K10864 DNA mismatch repair protein PMS1 |
| v1.15229  | K10858 DNA mismatch repair protein PMS2 |
| v1.108658; v1.14152; v1.01681 | K15082 DNA repair protein RAD7 |
| v1.16407 [+ 27 other sequence copies] | K10866 DNA repair protein RAD50 |
| v1.22193  | K04482 DNA repair protein RAD51 |
| v1.02646; v1.22076 | K10938 DNA repair protein RAD57 |
| v1.15671 [+ 4 other sequence copies] | K04483 DNA repair protein RadA |
| v1.16193; v1.19033 | K04485 DNA repair protein RadA/SmS |
| v1.16079; v1.07685 | K04484 DNA repair protein RadB |
| v1.21363; v1.22360; v1.02900 | K03584 DNA repair protein RecO (recombination protein O) |
| v1.18390  | K03515 DNA repair protein Rev1 |
| v1.04705  | K10991 DNA repair protein Swi5/Sae3 |
| v1.13920; v1.03800; v1.16133 | K10803 DNA repair protein XRCC1 |
| v1.15052  | K10879 DNA repair protein XRCC2 |
| v1.09315 [+ 4 other sequence copies] | K10886 DNA repair protein XRCC4 |
| v1.02733; v1.24592 | K10868 DNA repair protein XRS2 |
| v1.14551; v1.23176 | K10873 DNA repair and recombination protein RAD52 |
| v1.20503 [+ 4 other sequence copies] | K10875 DNA repair and recombination protein RAD54 |
| v1.23173; v1.16050 | K10877 DNA repair and recombination protein RAD54B |
| v1.07227; v1.08907; v1.09439; v1.02644 | K10847 DNA repair protein complementing XP-A cells |
| v1.11534 [+ 5 other sequence copies] | K10865 Double-strand break repair protein MRE11 |
| v1.07939  | K03660 N-glycosylase/DNA lyase |
| v1.16163  | K03652 3-Methyladenine DNA glycosylase |
| v1.07231  | K10726 Replicative DNA helicase Mcm |
| v1.05482  | K04499 RuVB-like protein 1 (pontin 52) |
| v1.19813  | K11338 RuVB-like protein 2 |
| v1.06890  | K15080 Single-strand annealing weakened protein 1 |
| v1.17193; v1.14087 | K03111 Single-strand DNA-binding protein |
| v1.15575  | K10800 Single-strand monofunctional uracil DNA glycosylase |
| v1.07134  | K10992 Swi5-dependent recombination DNA repair protein 1 |
| v1.13860  | K03649 TDG/mug DNA glycosylase family protein |
| v1.14423; v1.14399; v1.05070 | K03648 Uracil-DNA glycosylase |
| v1.23838  | K10791 Three prime repair exonuclease 2 |
| v1.19522  | K10839 UV excision repair protein RAD23 |
to facilitate repair. Of the multiple families of ATP-dependent DNA helicase proteins encoded in the coral genome, RecQ and helicase Q predominate. Encoded in the coral genome are 5 homologues of the DNA repair alkB proteins that reverse damage to DNA from alkylations caused by chemical agents by removing methyl groups from 1-methyl adenine and 3-methyl cytosine products in single-strand DNA. Annotated also are genes encoding DNA ligase 3 for repairing single-strand breaks, DNA ligase 4 to repair double-strand breaks, and a DNA cross-link repair 1C protein with single-strand specific endonuclease activity that may serve in a proofreading function for DNA polymerase. Taken together, expressing this arsenal of DNA protection may provide corals with limited ability to transcribe gene-encoded adaptation to a changing global environment.

**Stress response proteins**

Annotation of the *A. digitifera* genome reveals a wide assortment of thermal shock proteins, molecular chaperones and other stress response elements that are given in (Table 10), excluding antioxidant and redox-protective proteins which are described in the next section. Heat shock proteins 70 kDa, 90 kDa, 110 kDa, HspQ and HspX (the last two proteins being homologues of the bacterial heat shock factor sigma32 and α-crystallin, respectively) are encoded in the coral genome, together with several HSP gene transcription factors. HSPs play a role in various cellular functioning such as protein folding, intracellular protein trafficking and resistance to protein denaturation. HSP expression is usually increased on exposure to elevated temperatures and other conditions of biotic and abiotic stress that include infection, inflammation, metabolic hyperactivity, exposure to environmental toxicants, ultraviolet light exposure, starvation, hypoxia and desiccation [243]. HSPs and chaperones are transcriptionally regulated and are induced by heat shock transcription factors [244], of which there are several encoded in the coral genome. Since HSPs are found in virtually all living organisms, it is not surprising that cnidian *hsp* transcription and protein expression (HSP60, HSP70 and HSP90) have been profiled as a stress determinant [245-250] and early warning indicator of coral bleaching [251-254]. The coral genome reveals also a cold shock protein encoded by the *cspA* gene family, but profiling its expression with other stress response proteins activated by sub-optimum cold temperatures [255] has not been reported. Additionally, the coral genome encodes transcription of a homologue of the universal stress protein A (UspA), a member of an ancient and conserved group of stress-response proteins [256,257], which have been studied mostly in bacteria [258] but have been described also in several plants [259] and animals, including members of the Cnidaria [260]. UsP transcripts have been quantified in the thermal stress response of the coral *Montastraea faveolata* [261] and its aposymbiotic embryos [262]. Another gene product of potential interest is a homologue of the oxidative-stress responsive protein 1 (OXSR1) that belongs to the Ser/Thr kinase family of proteins, as do other mitogen-stress activated protein kinases (MAPKs), that regulate downstream kinases in response to environmental stress [263] by interacting with the Hsp70 subfamily of proteins [264]. Another significant response protein encoded in the coral genome (Table 10) is a homologue of the stress-induced phosphoprotein 1 (30 domain sequence alignments), known also as the Hsp70-Hsp90 organising protein (HOP) belonging to the stress inducible (STI1) family of proteins, which is a principle adaptor protein that mediates the functional cooperation of molecular chaperones Hsp70 and Hsp90 [265,266]. It is yet to be determined if *Hop1* transcription may serve as a primary indicator of environmental stress in corals.

Molecular chaperones are a diverse family of proteins expressed by both prokaryotic and eukaryotic organisms that serve to maintain correct protein folding in a 3-dimensional functional state, assist in multiprotein complex assembly and protect proteins from irreversible aggregation at synthesis and during conditions of cellular stress [267]. Additionally, heat shock proteins and their co-chaperones may regulate cell death pathways by inhibition of apoptosis [268]. The coral genome encodes a large number of DnaJ subfamily (J-domain) chaperones (Hsp40) that with co-chaperones GrpE (Table 10) regulates the ATPase activity of Hsp70 (DnaK in bacteria) to enable correct protein folding [269]. The coral genome encodes homologues of the molecular chaperones HscA (specialised Hsp70), the redox-regulated chaperone Hsp33, HtpG (high temperature protein G), members of the calnexin/calreticulin chaperone system of the endoplasmic reticulum, a mitochondrial chaperone BCS1 protein necessary for the assembly of the respiratory chain complex III and a specific chaperone of trimethyl N-oxide reductase (TorA). The coral genome also encodes hypoxia-inducible factors (HIFs) that moderate the deleterious effects of hypoxia on cellular metabolism (reviewed in [270]). In the HIF signalling cascade, the alpha subunits of HIF are hydroxylated at conserved proline residues by HIF prolyl-hydroxylases allowing their recognition for proteasomal degradation, which occurs during normoxic conditions but is repressed by oxygen depletion. Hypoxia-stabilised HIF1 upregulates the expression of enzymes principally of the oxygen-independent glycolysis pathway, and in higher animals promotes vascularisation, whereas the mammalian HIF2 paralogue regulates erythropoietin control of hepatic erythrocyte production in response to hypoxic stress [271]. The roles of HIF1 and HIF2 homologues in corals have been established, with HIF1 regulation of glycolysis critical to metabolic function during the dark diurnal anoxic state of coral respiration [193,272].
| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.04616; v1.06277 | K03694 | ATP-dependent Clp protease subunit ClpA |
| v1.04617; v1.23486; v1.23484; v1.10207 | K03695 | ATP-dependent Clp protease subunit ClpB |
| v1.13464 | K03697 | ATP-dependent Clp protease subunit ClpE |
| v1.06903; v1.11461 | K06891 | ATP-dependent Clp protease adapter protein ClpS |
| v1.12577; v1.09531; v1.17184 | K03544 | ATP-dependent Clp protease subunit ClpX |
| v1.09407 | K08054 | Calnexin (protein-folding chaperone) |
| v1.16781 | K08057 | Calreticulin (Ca\(^{2+}\)-binding chaperone) |
| v1.04005 | K10098 | Calreticulin 3 (Ca\(^{2+}\)-binding chaperone) |
| v1.02702 [+ 5 other sequence copies] | K03704 | Cold shock protein (beta-ribbon, CspA family) |
| v1.01907; v1.18998 | K07213 | Copper chaperone |
| v1.23457; v1.01713; v1.19228 | K04569 | Copper chaperone for superoxide dismutase |
| v1.08719; v1.19128 | K09502 | DnaJ homologue subfamily A member 1 |
| v1.08719; v1.18432 | K09503 | DnaJ homologue subfamily A member 2 |
| v1.16210; v1.22054 | K09504 | DnaJ homologue subfamily A member 3 |
| v1.19128 | K09505 | DnaJ homologue subfamily A member 4 |
| v1.04818 [+ 6 other sequence copies] | K09506 | DnaJ homologue subfamily A member 5 |
| v1.02841; v1.02842 | K09507 | DnaJ homologue subfamily B member 1 |
| v1.00368; v1.13308; v1.16977; v1.03340 | K09508 | DnaJ homologue subfamily B member 2 |
| v1.11537; v1.09205; v1.08628; v1.02840 | K09511 | DnaJ homologue subfamily B member 5 |
| v1.24549 [+ 9 other sequence copies] | K09512 | DnaJ homologue subfamily B member 6 |
| v1.01573 | K09513 | DnaJ homologue subfamily B member 7 |
| v1.00352; v1.09196; v1.06645 | K09514 | DnaJ homologue subfamily B member 8 |
| v1.18536 [+ 4 other sequence copies] | K09515 | DnaJ homologue subfamily B member 9 |
| v1.14710 | K09517 | DnaJ homologue subfamily B member 11 |
| v1.14959 | K09518 | DnaJ homologue subfamily B member 12 |
| v1.09205 | K09519 | DnaJ homologue subfamily B member 13 |
| v1.16242 | K09520 | DnaJ homologue subfamily B member 14 |
| v1.20109; v1.03468 | K09521 | DnaJ homologue subfamily C member 1 |
| v1.07111 [+ 5 other sequence copies] | K09522 | DnaJ homologue subfamily C member 2 |
| v1.21077 [+ 13 other sequence copies] | K09523 | DnaJ homologue subfamily C member 3 |
| v1.07739; v1.22910 | K09524 | DnaJ homologue subfamily C member 4 |
| v1.01239 [+ 13 other sequence copies] | K09525 | DnaJ homologue subfamily C member 5 |
| v1.17629 [+ 29 other sequence copies] | K09527 | DnaJ homologue subfamily C member 7 |
| v1.18619; v1.08300; v1.23789 | K09528 | DnaJ homologue subfamily C member 8 |
| v1.13575; v1.04213 | K09529 | DnaJ homologue subfamily C member 9 |
| v1.05956; v1.05955; v1.21265; v1.21205 | K09530 | DnaJ homologue subfamily C member 10 |
| v1.13525; v1.04120 | K09531 | DnaJ homologue subfamily C member 11 |
| v1.09496 [+ 4 other sequence copies] | K09533 | DnaJ homologue subfamily C member 13 |
| v1.24546 | K09534 | DnaJ homologue subfamily C member 14 |
| v1.05866 | K09536 | DnaJ homologue subfamily C member 16 |
| v1.16151; v1.08307; v1.14980 | K09537 | DnaJ homologue subfamily C member 17 |
| v1.16309 | K09539 | DnaJ homologue subfamily C member 19 |
| v1.05241; v1.22599; v1.17372 | K14258 | Facilitated trehalose transporter (anhydrobiosis) |
Heat shock proteins that repair unfolded or misfolded protein have a complementary function to the ubiquitin-proteasome system (ubiquitins not tabulated) that selects damaged protein for degradation [273], such that HSP chaperones and the proteasome act jointly to preserve cellular proteostasis [274,275]. Thus, several proteasome chaperones and assembly chaperones are encoded in the *A. digitifera* genome (Table 10). While proteasome chaperones serve to target aberrant proteins for ubiquination, the proteasome chaperones facilitates 20S assembly for

**Table 10** Stress response proteins in the predicted proteome of *A. digitifera* (Continued)

| Accession | Description |
|-----------|-------------|
| v1.12967; v1.19789 | K14590 | FtsJ methyltransferase [heat shock protein] |
| v1.02247 | K09414 | Heat shock transcription factor 1 |
| v1.24112 | K09416 | Heat shock transcription factor 3 |
| v1.05839 | K09419 | Heat shock transcription factor, other eukaryote |
| v1.12890 [+ 10 other sequence copies] | K03283 | Heat shock 70 kDa protein 1/8 |
| v1.07996 | K09489 | Heat shock 70 kDa protein 4 |
| v1.02854; v1.07452; v1.01623 | K09490 | Heat shock 70 kDa protein 5 |
| v1.14149; v1.14150 | K09487 | Heat shock protein 90 kDa beta |
| v1.07995; v1.07996; v1.16399; v1.11283 | K09485 | Heat shock protein 110 kDa |
| v1.08943; v1.05577 | K11940 | Heat shock protein HspQ |
| v1.00537; v1.00043 | K03799 | Heat shock protein HtpX |
| v1.01623 | K04046 | Hypothetical chaperone protein |
| v1.16216 | K08268 | Hypoxia-inducible factor 1 alpha |
| v1.08869; v1.15120 | K09097 | Hypoxia-inducible factor 1 beta |
| v1.22724 | K09095 | Hypoxia-inducible factor 2 alpha |
| v1.23608 [+ 16 other sequence copies] | K06711 | Hypoxia-inducible factor prolyl 4-hydroxylase |
| v1.16737; v1.22345 | K09486 | Hypoxia up-regulated 1 (heat shock protein 70 family) |
| v1.10188 | K08900 | Mitochondrial chaperone BCS1 |
| v1.17197; v1.04394 | K04445 | Mitogen-stress activated protein kinases |
| v1.16301; v1.21224; v1.19344 | K04043 | Molecular chaperone DnaK |
| v1.09682; v1.16748; v1.07471; v1.13624 | K03687 | Molecular chaperone GrpE |
| v1.01621; v1.04945; v1.15919 | K04044 | Molecular chaperone HscA |
| v1.18210 | K04083 | Molecular chaperone Hsp33 |
| v1.17478; v1.16977; v1.10289; v1.19007 | K04079 | Molecular chaperone HtpG |
| v1.08895; v1.18099 | K11416 | Mono-ADP-ribosyltransferase sirtuin 6 |
| v1.02024 | K11411 | NAD-dependent deacetylase sirtuin 1 |
| v1.04813 | K11412 | NAD-dependent deacetylase sirtuin 2 |
| v1.22049; v1.22211; v1.02221 | K11413 | NAD-dependent deacetylase sirtuin 3 |
| v1.11849; v1.02221 | K11414 | NAD-dependent deacetylase sirtuin 4 |
| v1.05495 | K11415 | NAD-dependent deacetylase sirtuin 5 |
| v1.04868 | K11417 | NAD-dependent deacetylase sirtuin 7 |
| v1.15070 [+ 4 other sequence copies] | K08835 | Oxidative-stress responsive protein 1 (OXSR1) |
| v1.04503 | K11875 | Proteasome assembly chaperone 1 |
| v1.01531 | K11878 | Proteasome assembly chaperone 4 |
| v1.01210 | K11879 | Proteasome chaperone 1 |
| v1.18611 | K11880 | Proteasome chaperone 2 |
| v1.00599 [+ 29 other sequence copies] | K09553 | Stress-induced-phosphoprotein 1 (HOP1) |
| v1.108830 | K13057 | Trehalose synthase (anhydrobiosis) |
| v1.22042 | K03533 | TorA specific chaperone |
| v1.16986 [+ 7 other sequence copies] | K06149 | Universal stress protein A |
biogenesis of the multiunit 26S proteasome that is activated in response to stress [276,277], possibly by FtsJ (aka Rrm1), a well-conserved heat shock protein having novel ribosomal methyltransferase activity that targets methylation of 26S rRNA under heat shock control [278,279]. The HspQ protein encoded in the coral genome, although studied almost exclusively in bacteria, is known to stimulate degradation of denatured proteins caused by hyperthermal stress, particularly DnaA that initiates DNA replication in prokaryotes [280]. Specifically, HspQ (heat shock factor sigma32) regulates the expression of Clp ATPase-dependent protease family enzymes [281,282], of which ClpA, ClpB, ClpE, the protease adaptor protein ClpS [283] and the unfoldase ClpX protein [284] are encoded in the coral genome (Table 10). HspX is a small 16 kDa α-crystallin chaperone (Acr) protein belonging to the Hsp20 family of proteins [285] that suppresses thermal denaturation and aggregation of proteins [286]. Significantly, Acr proteins are known to bind with carbonic anhydrase [287] and be more important in moderating stress-induced loss of calcium deposition. Thus, HspX/Acr expression may account for differences in the thermal sensitivity of corals to calcification that varies among genera [287]. In a different context, HspX is attracting considerable attention for its potential to elicit long-term memory [287]. In a different context, HspX is attracting considerable attention for its potential to elicit long-term memory [287].

The coral genome encodes complete membership of the human sirtuin (SIRT1-7) family of NAD(+)-dependent protein deacetylases and ADP-ribosyltransferases. Mammalian SIRT1 (a homologue of yeast Sir2) is an important regulator of metabolism, cell differentiation, stress response transcription and pathways of cellular senescence (reviewed in [291]). SIRT proteins regulate chromatin function through deacetylation of histones that promote subsequent alterations in the methylation of histones and DNA to affect, via deactivation of nuclear transcription factors and co-regulators, epigenetic control of nuclear transcription. As NAD(+)-dependent enzymes, SIRT1 can regulate gene expression in response to cellular NAD+/NADH redox status providing a metabolic template for epigenetic transcriptome reprogramming [292,293]. In the human genome repertoire, SIRT1 modulates cellular responses to hypoxia by deacetylation of HIF1α [294] and inhibits nitric oxide synthesis by suppression of the nuclear factor-kappaB (NF-κB) signalling pathway [295], SIRT2 promotes oxidative stress resistance by deacetylation of forkhead box O (FOXO) proteins [296], SIRT3 decreases ROS production in adipocytes [297], SIRT4 regulates fatty acid metabolism and stress-response elements of mitochondrial gene expression [298], SIRT5 is a protein lysine desuccinylase and demalonylase of unknown function [299], SIRT6 activates base-excision repair [300] and SIRT7 inhibits apoptosis induced by oxidative stress by deacetylation of p53 [301,302]. The significance of coral SIRT proteins, by analogy, to exert stress tolerance is yet to be examined.

Metallochaperones are an important class of enzymes that transport co-factor metal ions to specific proteins [303]. The copper chaperone protein ATX1 (human ATOX1) delivers cytosolic copper to Cu-ATPase proteins and serves as a metal homeostasis factor to prevent Fenton-type production of highly reactive hydroxyl radicals. ATX1, which is strongly induced by molecular oxygen, functions additionally as an antioxidant to protect cells against the toxicity of both the superoxide anion and hydrogen peroxide [304]. Encoded also is a specific copper chaperone essential to the activation of Cu/Zn superoxide dismutase [305,306] that is enhanced by photooxidative stress in scleractonian corals [307], although reported to be less pronounced in the host than in symbiotic algae [308]. In addition to high light exposure, reef-building corals of shallow reef flats are occasionally exposed to the atmosphere for periods that can last several hours during extreme low tides. Hence, species that are adapted to withstand acute desiccation (anhydrobiosis) have a better chance of surviving such conditions. The disaccharide trehalose is an osmolyte that in some plants and animals allows them to survive prolonged periods of desiccation [309]. The hydrated sugar has high water retention that forms a gel phase when cells dehydrate, which on rehydration allows normal cellular activity to resume without damage that would otherwise follow a dehydration/rehydration cycle. Furthermore, trehalose is highly effective in protecting enzymes in their native state from inactivation from thermal denaturation [310]. Given that A. digitifera is endemic on shallow reef flats prone to exposure at low tides [311], it is not surprising that the coral genome encodes trehalose synthase and a facilitated trehalose transporter for protection against dehydration.

Antioxidant and redox-protective proteins
Oxygen is vital for life, but it can also cause damage to cells, particularly at elevated levels. In coral symbiosis, the photosynthetic endosymbionts of corals typically produce more oxygen than the holobiont is able to consume by respiration, so that coral tissues are hyperoxic with tissue pO2 levels often exceeding 250% of air saturation during daylight illumination [193]. Furthermore, because algal symbionts reside within the endodermal cells of their host, coral tissues must be transparent to facilitate the penetration of downwelling light required for photosynthesis by their algal consorts. In clear shallow waters this entails concurrent exposure to vulnerable...
molecular sites of both partners to damaging wavelengths of ultraviolet radiation. The synergistic effects of tissue hyperoxia and UV exposure can cause oxidative damage to the symbiosis via the photochemical production of cytotoxic oxygen species [312] that are produced also during normal mitochondrial function [313]. Consequently, protective proteins (antioxidant enzymes) are expressed to maintain the fine balance between oxygen metabolism and the production of potentially toxic reactive oxygen species (ROS). If this balance is not maintained by regulation of oxidative and reductive processes (redox regulation), oxidative stress occurs by the generation of excess ROS, causing damage to DNA, proteins, and lipids. Corals elaborate a variety of molecular defences that including the production of UV-protective sunscreens, (MAAS), antioxidants, antioxidant enzymes, chaperones and heat shock proteins, which are often inducible under conditions of enhanced oxidative stress [307], including conditions that elicit coral bleaching [314,315]. An excellent review on the formation of ROS and the role of antioxidants and antioxidant enzymes in the field of redox biology is given by Halliwell [316].

Annotation of the A. digitifera genome reveals sequences encoding two isoforms of the antioxidant enzyme superoxide dismutase (SOD) from both the Cu/Zn and Fe/Mn families of SOD (Table 1). These metalloprotein enzymes catalyse the dismutation of superoxide to yield molecular oxygen and hydrogen peroxide, the latter being less harmful than superoxide. Superoxide can oxidize proteins, denature enzymes, oxidize lipids and fragment DNA. By removing superoxide, SOD protects also against the production of reactive peroxynitrite formed by the combination of superoxide and nitric oxide, which is a precursor reactant for production of the super-reactive hydroxyl radical. Hydrogen peroxide per se is a mild oxidant, but it readily oxidises free cellular ferrous iron to ferric iron with production of hydroxyl radicals via the Fenton reaction. Accordingly, both the removal of hydrogen peroxide and the expression of proteins, such as transferrin, (bacterio)ferritins and metallothioneins, that bind reactive (transition) metal ions is important to protect cellular components from acute oxidative damage. Oddly, only a metallothionein expression activator was found encoded in the coral genome without finding a sequence to activate transcription of the actual metallothionein protein gene.

As expected from the foregoing, the genome of A. digitifera encodes the antioxidant enzyme catalase (CAT) that is highly efficient in decomposing hydrogen peroxide to yield molecular oxygen and water. Two isoforms of CAT are encoded at multiple sites. One is a peroxisomal eukaryotic CAT enzyme that targets the removal of hydrogen peroxide formed as a by-product of oxidase enzymes, and the other is a related catalase domain-containing protein presumed also to decompose hydrogen peroxide. Glutathione peroxidise (GPx) reduces both hydrogen peroxide and lipid hydroperoxides, the latter of which are formed by radical-induced lipid autoxidation. Photo-trophic organisms, including higher plants, utilise ascorbate peroxidase (APx) as a primary catalyst for the reduction of hydrogen peroxide and lipid hydroperoxides. However, unlike the freshwater cnidarian H. viridis [164], there is no evidence for transfer of APx-encoding genes to A. digitifera. The antioxidant enzymes SOD, CAT, GPx and APx are well characterised in the algal and animal partners of coral symbiosis (reviewed in [317]). Additionally, the coral genome has sequences encoding alkyl hydroperoxide reductase, hydroperoxide lyase, phospholipid-hydroperoxide glutathione peroxidase, thiol peroxidase and multiple isoforms of peroxiredoxin, all of which function in the detoxification of organo-hydroperoxides that are produced as a by-product of aerobic metabolism. Additionally, sulfiredoxin (Table 11) repairs peroxiredoxins when these enzymes are inhibited by over-oxidation [318].

Thioredoxins and glutaredoxins have important secondary roles in regulating multiple pathways in many biological processes, including redox signalling of apoptotic pathways, which have been attributed to processes involved in coral bleaching [56]. Other enzymes that regulate cellular thiol-disulfide homeostasis in this coral are monothiol glutaredoxin and protein-disulfide reductase. The coral genome encodes the ubiquitous thioredoxin system of antioxidant proteins (Table 11) that act as electron donors to peroxidases and ribonucleotide reductase (the latter not tabulated). By cysteine thiol-disulfide exchange, thioredoxins function as a protein thiol-disulfide oxidoreductase [319]. In the thioredoxin system, thioredoxins are maintained in their reduced state by NADPH-dependent, flavoenzyme thioredoxin reductase [320]. Peptide-methionine (R)-S-oxide reductase can additionally rescue thioredoxin from oxidative inactivation by disulfide reduction. Related glutaredoxins share many of the functions of thioredoxins but are reduced directly by glutathione, rather than by a specific reducing enzyme, while in turn glutathione is kept in its native state by NADPH: glutathione reductase.

In recent years there has been a particular focus on the role of ROS in coral bleaching, fuelled by dire prediction of future catastrophic episodes caused by environmental change affected by global warming [321]. Early predictions of coral bleaching were based principally on physical environmental parameters, rather than on the determination of the physiological state of coral populations to such conditions. While gene expression markers are being developed to monitor sub-bleaching levels of stress in situ (e.g., [261]), Kenkel et al. [322] opined that the current challenge for implementing expression-based methods lies in identifying coral genes demonstrating the most pronounced and consistent
stress response, preferably with a large dynamic range to enable reliable quantification. To this end, we offer in Table 11 the annotation of novel redox-related genes for examination as potential candidate biomarkers to monitor the physiological response of *A. digitifera* to environmental stress.

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|----------------------------|
| v1.10918      | K04756         | Alkyl hydroperoxide reductase subunit D |
| v1.11551      | K03387         | Alkyl hydroperoxide reductase subunit F |
| v1.07812      | K03594         | Bacterioferritin |
| v1.21362 [+ 4 other sequence copies] | K00429 | Catalase (bacterial) |
| v1.17525 [+ 4 other sequence copies] | K03781 | Catalase (peroxisomal) |
| v1.23457; v1.01713; v1.19228 | K04569 | Copper chaperone for superoxide dismutase |
| v1.20153; v1.20154 | K10558 | Hydroperoxide lyase |
| v1.19687; v1.19688; v1.18796; v1.18795 | K00522 | Ferritin heavy chain |
| v1.06441      | K03674         | Glutaredoxin 1 |
| v1.19449      | K03675         | Glutaredoxin 2 |
| v1.14929 [+ 5 other sequence copies] | K03676 | Glutaredoxin 3 |
| v1.13285; v1.03722; v1.03688; v1.10496 | K00432 | Glutathione peroxidase |
| v1.13174; v1.13775; v1.05473 | K00383 | Glutathione reductase (NADPH) |
| v1.14344; v1.19399; v1.01421 | K01920 | Glutathione synthase |
| v1.02173      | K02938         | Metallothionein expression activator |
| v1.09719; v1.16134; v1.18608 | K07390 | Monothiol glutaredoxin |
| v1.14890; v1.17685 | K07305 | Peptide-methionine (R)-S-oxide reductase |
| v1.14909      | K00435         | Peroxiredoxin |
| v1.14106      | K13279         | Peroxiredoxin 1 |
| v1.08691      | K11187         | Peroxiredoxin 5, atypical 2-Cys peroxiredoxin |
| v1.01410      | K11188         | Peroxiredoxin 6, 1-Cys peroxiredoxin |
| v1.03688      | K05361         | Phospholipid-hydroperoxide glutathione peroxidase |
| v1.05148      | K05905         | Protein-disulfide reductase |
| v1.02922; v1.22772; v1.24164 | K05360 | Protein-disulfide reductase (glutathione) |
| v1.06810      | K12260         | Sulfiredoxin |
| v1.01713 [+ 4 other sequence copies] | K04565 | Superoxide dismutase, Cu/Zn family |
| v1.09974; v1.20324 | K04564 | Superoxide dismutase, Fe/Mn family |
| v1.02378      | K11065         | Thioldioxidase, atypical 2-Cys peroxiredoxin |
| v1.22324 [+ 7 other sequence copies] | K03671 | Thioredoxin 1 |
| v1.05148; v1.03230; v1.20699 | K03672 | Thioredoxin 2 |
| v1.17881 [+ 5 other sequence copies] | K13984 | Thioredoxin domain-containing protein 5 |
| v1.04532; v1.24501 | K09585 | Thioredoxin domain-containing protein 10 |
| v1.11551; v1.19049 | K00384 | Thioredoxin reductase (NADPH) |
| v1.10930      | K14736         | Transferrin |

Table 11 Antioxidant and redox-protective proteins in the predicted proteome of *A. digitifera*

Proteins of cellular apoptosis

Apoptosis is the signalling of programmed cell death (PCD) that occurs in multicellular organisms in response to cellular injury. A key feature of apoptosis is the activation of endogenous endonucleases causing nuclear fragmentation, chromatin condensation and chromosomal DNA fragmentation, which typically presents in affected cells by the morphological appearance of plasma membrane blebbing and cell shrinkage. Caspases and related family member proteases are described as “executioners” of apoptosis that on post-translational activation degrade the regulatory proteins that prevent DNA degradation. Fragmentation of nuclear DNA is one of the
hallmarks of apoptotic cell death that occurs by PCD stimuli in a wide variety of proliferating cells. NF-κB is a protein complex that controls the transcription of DNA that can induce the expression of nitric oxide synthase (NOS) to produce NO that is a well-known promoter of the of the pro-apoptotic transcription factor p53 cell-cycle gatekeeper of the caspase cascade. In contrast to necrosis, which is the outcome of PCD, apoptosis mediates the fragmentation of damaged cells, which by phagocytosis are removed or degraded in phagolysosomes to spare surviving cells from the uncontrolled release of cytotoxic agents. Proteins of the caspase-mediated apoptotic cascade are regarded as products of constituent housekeeping genes that are necessary to maintain healthy multicellular function [323]. In the progression of cnidarian bleaching, apoptotic pathways are activated [322-325], but not all corals that suffer bleaching are destined to die [326,327]. Coral survival has been attributed to having a high level of apoptotic protection at the onset of coral bleaching [328] and during post-bleaching recovery [329] by specific activation of anti-apoptotic Bcl-2 proteins in surviving cells [330].

Cnidarians have a complex apoptotic protein network that has exceptional ancestral complexity and is comparable to that of higher vertebrates [331,332]. Cnidarian metamorphosis is tightly coupled with caspase-dependent apoptosis [333] and subsequent host-symbiont selection by post-phagocytic winnowing of *Symbiodinium* genotypes during the establishment of coral-dinoflagellate mutualism [334]. As expected, the coral genome of *A. digitifera* encodes multiple isoforms of genes that transcribe the caspase family of apoptotic effectors (Table 12). Included in this signalling pathway are the pro- and anti-apoptotic Bax/Bcl regulators and Bcl-2 anathogene (DNA-binding) activators of apoptosis. Notable in our annotation dataset are multiple genes that encode the protein domains of the apoptotic protease-activating factor (Apaf) that triggers assembly of the apoptosome leading to caspase activation [335]. Additional to this arsenal of cell cycle regulators are the death associated protein-6 (DAXX), a Fas-binding adaptor of c-Jun N-terminal kinase (JNK) activation [336], death-associated protein kinase (DAPK), a mediator of calcium/calmodulin-regulated Ser/Thr kinase [337], and the programmed cell death 6-interacting protein (PDCD6IP), which binds to PDCD-6 for execution of apoptosis via the caspase-3 pathway [338]. PDCD6IP activation of apoptosis is an enigma since PDCD-6 is not encoded in the coral genome, nor is caspase-3. Other cell cycle regulators are the p53 binding and p53-associated parkin-like proteins, and the activating TP53 regulating kinase protein and TP53 apoptosis effector of *TP53* gene expression.

Our genome annotation reveals 73 sequence matches for expressing the Apaf protein domain that, in conjunction with a high copy number for expressing caspase-8 (28 protein sequence matches), may enhance coral survival during embryogenesis by suppressing receptor-induced protein kinase (45 sequence matches) during early development [339]. The most conserved function of the CAPS2/RIPK adaptor (45 sequence matches) encoded in the coral genome is its essential regulation of apoptosis [340]. We find a wide repertoire of genes that additionally encode proteins that mediate apoptosis (Table 12). Amongst these are the calpain Ca²⁺-sensing family of proteins that initiate the signalling of apoptotic pathways [341]. There are 79 matches to sequences that encode the tumor necrosis Fas superfamily member 6 (TNFRSF6) receptor, which coupled with the death domain (FADD) protein is a cell signalling mediator for recruitment of caspase-8 that activates the apoptotic cysteine protease cascade. Coincident in the genome are 67 sequences encoding the leucine-rich repeat and death domain-containing (LRDD) adaptor that, by interacting with other p53-inducible death domain-containing (PIDD) proteins such as FADD, induces the caspase-2 pathway of apoptosis in response to DNA damage [342]. Elements of the NF-κB signalling pathway of cnidarians are highly conserved traits [343], which includes the caspase cascade and the pro-apoptotic and anti-apoptotic Bcl-2 family of proteins [344]. The coral genome of *A. digitifera* encodes the pleiotropic nuclear factor NF-κB p105 subunit, and astonishingly there are 212 sequence matches to the NF-κB inhibitor-like protein 2 domain with fewer matches to the NF-κB inhibitor-like protein 1 and NF-κB family inhibitors alpha, delta and epsilon. Evident in our genome annotation is the tumor necrosis factor-alpha induced protein 3 (TNFAIP3), a cytokine produced by activated (inflammatory) macrophages. Although TNF cytokines are a major extrinsic mediator of cellular apoptotic pathways, the precise function of the superfamily members of TNF ligands and receptors (Table 12) remains elusive in coral symbiology.

**Microbial symbiosis and pathogenicity**

It is well established that corals associate with a vast consortia of microbes, including phototrophic symbionts (*Symbiodinium spp.*) and other eukaryotic microorganisms, cyanophytes, heterotrophic bacteria, archaea and viruses [345]. Corals harbour diverse and abundant prokaryotic communities with distinct populations residing in separate habitats of the host skeleton, tissues and surface mucus layer (reviewed in [203]). Microbial populations are dominated by a few coral-specific taxonomic traits [346], but the majority of the population comprises a high number of taxonomically diverse, low-abundance ribotypes [347] with much of the diversity within the coral microbiome belonging to the “rare” biosphere [348,349]. The coral microbiome is vital to the nutrition and health of the holobiont [350] and contributes
Table 12 Proteins of cellular apoptosis in the predicted proteome of *A. digitifera*

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.17521; v1.02505; v1.20702; v1.05077 | K02159 | Apoptosis regulator BAX (BCL2-associated) |
| v1.05086; v1.20659 | K02161 | Apoptosis regulator BCL-2 |
| v1.17522; v1.00181; v1.10817; v1.20703 | K02163 | Apoptosis regulator BCL-W |
| v1.05147 [+ 6 other sequence copies] | K12875 | Apoptotic chromatin condensation inducer |
| v1.22264 [+ 72 other sequence copies] | K02084 | Apoptotic protease-activating factor (Apaf) |
| v1.17326; v1.20305; v1.11586 | K09555 | BCL2-associated athanogene 1 |
| v1.08601 | K09558 | BCL2-associated athanogene 4 |
| v1.02839 | K09559 | BCL2-associated athanogene 5 |
| v1.01518 | K13087 | BCL2-associated transcription factor 1 |
| v1.20278; v1.00172; v1.07858 | K14021 | BCL-2 homologous antagonist/killer |
| v1.09624 | K02561 | BCL2-related (ovarian) killer protein |
| v1.17749 | K08573 | Calpain-3 |
| v1.00595; v1.14671; v1.00040 | K08574 | Calpain-5 |
| v1.00040 | K08575 | Calpain-6 |
| v1.19153; v.17749 | K08576 | Calpain-7 |
| v1.15226 | K04740 | Calpain-12 |
| v1.02951 | K08582 | Calpain-15 |
| v1.11167; v1.06681; v1.20230; v1.01376 | K08585 | Calpain, invertebrate |
| v1.0312 7 [+ 6 other sequence copies] | K08583 | Calpain, small subunit 1 |
| v1.17229; v1.00023; v1.09976 | K02186 | Caspase 2 |
| v1.11989 [+ 5 other sequence copies] | K04397 | Caspase 7 |
| v1.02756 [+ 27 other sequence copies] | K04398 | Caspase 8 |
| v1.01818 | K04399 | Caspase 9 |
| v1.00817 [+ 4 other sequence copies] | K04400 | Caspase 10 |
| v1.02005 | K04741 | Caspase 12 |
| v1.00818 [+ 11 other sequence copies] | K04489 | Caspase apoptosis-related cysteine protease |
| v1.13260 | K07367 | Caspase recruitment domain-containing protein 11 |
| v1.06297 [+ 44 other sequence copies] | K02832 | CASP2 and RIPK1 adaptor with death domain |
| v1.21531 | K02308 | Death-associated protein 6 (DAXX) |
| v1.09448; v1.15529; v1.20164 | K08803 | Death-associated protein kinase (DAPK) |
| v1.23110; v1.14222; v1.03658 | K12366 | Engulfment and motility protein 1 (phagocytosis/apoptosis) |
| v1.18448 [+ 78 other sequence copies] | K02373 | Fas (TNFRSF6)-associated via death domain (FADD) |
| v1.24288 [+ 66 other sequence copies] | K10130 | Leucine-rich repeats and death domain-containing protein |
| v1.20620 | K04734 | NF-kappa-B inhibitor alpha |
| v1.01706 | K14214 | NF-kappa-B inhibitor delta |
| v1.10378; v1.10729; 1.05609; v1.05609 | K05872 | NF-kappa-B inhibitor epsilon |
| v1.17893; v1.22419; v1.00700; v.08415 | K09256 | NF-kappa-B inhibitor-like protein 1 |
| v1.04158 [+ 211 other sequence copies] | K09257 | NF-kappa-B inhibitor-like protein 2 |
| v1.05320; v1.06979; v1.04467; v1.21371 | K02580 | Nuclear factor NF-kappa-B p105 subunit |
| v1.20334; v1.22743 | K11970 | p53-Associated parkin-like cytoplasmic protein |
| v1.14920; v1.11864; v1.15271; v1.11865 | K06643 | p53-Binding protein |
| v1.04289 | K06708 | Programmed cell death 1 ligand 2 |
| v1.05882 [+ 7 other sequence copies] | K12200 | Programmed cell death 6-interacting protein (PDCD6IP) |
significantly to the protection of coral reef ecosystems against the detrimental effects of organic enrichment [351,352]. One emerging threat to coral reefs is the outbreak of infectious diseases (reviewed in [353]). Although highly subjective and with little experimental evidence to date, the coral probiotic hypothesis [354] suggests that the coral prokaryotic microbiome can adapt to changing environmental conditions by selective microbial reorganisation to impart greater resistance to disease and pathogen-mediated bleaching [355]. Whether the coral microbiome can respond to changing environmental conditions more rapidly than by host genetic mutation and selection based on contemporary phenotypic evolution on ecological time-scales [356], is a topic of current debate [357].

Corals, like other invertebrates, have an innate immune system based on self-histocompatibility recognition (reviewed in [358]), but to date few adaptive components have been identified [359]. Corals do not produce antibodies and thus lack a true adaptive immune system. Nonetheless, corals once susceptible to infection and bleaching caused by a specific bacterial agent can become immune to the invading pathogen by a phenomenon termed “experience-mediated tolerance”, a precept of the hologenome theory of evolution [360], although how this process occurs is largely unknown. In our annotation of the genome sequence of *A. digitifera* we uncovered genes encoding the expression of disease resistance proteins (Table 13), two of which match the plant RPM1 and RPS2 pathogen resistance proteins that guard against disease by binding with pathogen avirulence receptors [360,361]. Significant also is a gene to express the pathogenesis-related protein PR-1 (29 sequence domain matches) that is inducible in plants for systemic acquired resistance to pathogenic invasion [362]. We uncovered also multiple genes encoding the expression of myeloperoxidase (MPO) enzymes. MPOs produce hypochlorous acid from hydrogen peroxide and chloride ion (requiring heme as a cofactor), and it oxidizes tyrosine to the tyrosyl radical using hydrogen peroxide as an oxidizing agent. Hypochlorous acid and tyrosyl radicals are strong cytotoxic agents that in higher organisms are used as a primary defence by neutrophils to protect against invading pathogens. Phenoloxidase (tyrosinase) activity is reported to contribute to the innate defence system of *A. millepora* and *Porites* sp.
via activation of the melanin-signalling pathway that is induced in response to coral bleaching and localised disease [364,365]. Three genes of A. digitifera encode tyrosinase enzymes (data not tabulated) to account for the phenoloxidase activity reported in corals.

The genome of A. digitifera also reveals homologues of genes that promote bacterial pathogenicity (Table 13), including virulence factors that are expressed and excreted by invading pathogens (bacteria, viruses, fungi and protozoa) to inhibit certain protective functions of the host. Such are the bacterial Type III cytotoxic effector protein and multiple Type IV Cag pathogenicity island proteins encoded in the coral genome. Many Gram-negative bacteria utilize Type III secretion proteins, which are regulated by quorum sensing, to deliver cytotoxic effector proteins into eukaryote host cells during infection. Cag (cytotoxin-associated) pathogenicity island (PAI) proteins are encoded by mobile genetic elements of the Type IV system secreting both proteins and large nucleoprotein complexes [366] that may be transferred between prokaryotes to enhance selected traits of virulence [367]. Our annotation reveals genes encoding six pathogenicity island proteins (Table 13) with similarity to the Cag PAI proteins of the human Helicobacter pylori, an infectious bacterium causing peptic ulcers that may lead to the development of stomach cancer. While many properties of Type III and IV secretion system proteins have been well characterized in bacteria, the functional purpose of homologous genes in A. digitifera, if expressed, are unknown.

The genome of A. digitifera contains genes of bacterial origin that encode the motility quorum-sensing regulator of the GCU-specific mRNA interferase toxin and acyl homoserine lactone synthesis used for the communication

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|----------------------------|
| v1.06126      | K13061         | Acyl homoserine lactone synthase |
| v1.19990      | K01372         | Bleomycin hydrolase         |
| v1.00209; v1.06178 | K03587     | Cell division protein FtsI (penicillin-binding protein 3) |
| v1.18860      | K13458         | Disease resistance protein  |
| v1.16231; v1.00374; v0.08191 | K13457       | Disease resistance protein RPM1 |
| v1.13482 [+ 4 other sequence copies] | K13459       | Disease resistance protein RPS2 |
| v1.07889      | K12090         | Cag pathogenicity island protein 5 |
| v1.24345      | K12091         | Cag pathogenicity island protein 6 |
| v1.18924; v1.17622 | K12093       | Cag pathogenicity island protein 8 |
| v1.05278      | K12096         | Cag pathogenicity island protein 11 |
| v1.02083      | K12104         | Cag pathogenicity island protein 19 |
| v1.12907      | K12109         | Cag pathogenicity island protein 24 |
| v1.00209; v1.06178 | K03587     | Cell division protein FtsI (penicillin-binding protein 3) |
| v1.13874      | K07259         | Carboxy/endopeptidase (penicillin-binding protein 4) |
| v1.12514; v.09758 | K04127       | Isopenicillin-N epimerase    |
| v1.21332      | K04126         | Isopenicillin-N synthase     |
| v1.07742      | K02547         | Methicillin resistance protein |
| v1.17478; v1.16977; v1.10289; v1.19907 | K04079       | Molecular chaperone HtpG (anti-bacterial) |
| v1.08255      | K13651         | Motility quorum-sensing regulator, GCU-specific toxin |
| v1.14792 [+ 7 other sequence copies] | K10789       | Myeloperoxidase               |
| v1.02333 [+ 26 other sequence copies] | K13449       | Pathogenesis-related protein 1 |
| v1.05017      | K03693         | Penicillin-binding protein   |
| v1.17507      | K12556         | Penicillin-binding protein 2X |
| v1.13874      | K07259         | Penicillin-binding protein 4 |
| v1.16655      | K02171         | Penicillinase repressor      |
| v1.14688      | K15126         | Type III secretion system cytotoxic effector protein |
| v1.20647      | K03980         | Virulence factor, integral membrane protein |
| v1.18964      | K03810         | Virulence factor, oxidoreductase domain |
of quorum sensing between bacteria to enable the coordination of group behaviour based on collective population density. Apparent in our annotation (Table 13) is a wide array of microbial penicillin-binding proteins (PBP)s that have an affinity for β-lactam antibiotics that by binding to PBPs prevent bacteria from constructing a cell wall. There are genes also to enhance antibiotic resistance, including potential expression of a penicillinase repressor, a methicillin resistance protein and bleomycin hydrolase (cysteine peptidase). Additionally, isopenicillin-N synthase and an isopenicillin-N epimerase, both of which catalyse key steps in the biosynthesis of penicillin and cephalosporin antibiotics, are encoded in the coral genome. Taken as a whole, we demonstrate an extensive presence of ancient non-metazoan genes that are maintained in the genome of A. digitifera, as is reported in the genomes of A. millepora and the anemone N. vectensis [368]. Recent thought on genome evolution places these ancestral conserved domains as ‘orphan’ or ‘taxonomically restricted’ genes [352,369,370], rather than acquired later by horizontal gene transfer. There is, of course, little knowledge of how or when, if at all, these non-metazoan genes are expressed or even their function to mediate pathogenicity in the coral holobiont.

Proteins of viral pathogenicity
Marine viruses were of minor interest until 1989, when it was realised that virus-like particles (VLPs) are the most abundant biological entities to occupy aquatic environments with variable numbers reaching ~10^8 VLPs ml^-1 [371]. Typically, VLPs surpass the number of marine bacteria by an order of magnitude in coastal waters [372]; their diversity is extremely high and many are specific to the marine environment [373,374]. Significant VLP numbers are reported from the surrounding waters of oceanic coral reef atolls [375], in waters flowing across the reef substratum [376] and in samples taken within the close vicinity of coral colonies [377,378]. The viral load within the surface microlayer of scleractinian corals is enumerated as being 10^7-10^8 VLPs ml^-1 [379] and, based on VLP morphological diversity, is attributed to infecting various microbial hosts (bacteria, archaea, cyanobacteria, fungi and algae) residing within the coral mucus [380]. VLPs have been observed in the epidermal and gastrodermal tissues of corals and occasionally occur in the mesogloea [381]. Latent viruses were found to infect Symbiodinium isolated from several scleractinian corals [382-384] with a preponderance of eukaryotic algae-infecting phycodnaviruses suggested [385]. A wide range of bacteriophage and eukaryotic virus families have been identified within scleractinians using metagenomic analyses [207,386-388], with bacteriophages being by far the most abundant entities (Wood-Charlson EM, Weynberg KD, Suttle CA, Roux S, van Oppen MJH: Methodological biases in coral viromics, submitted).

The importance of the coral-virus interactome in bleaching and disease (reviewed in [185,389]) is founded on reports showing that VLP abundances are higher in the seawater immediately surrounding diseased compared to that of healthy corals [378], that latent viruses are induced by heat stress in symbiotic dinoflagellates of the sea anemone Anemonia viridis [382] and the coral Pavona danai [383], and that UV exposure induces a latent virus-like infection in cultured Symbiodinium [187]. Quantitative 454 pyrosequence analysis of the coral Porites compressa on exposure to reduced pH, elevated nutrients or thermal stress showed that the abundance of its viral consortia varied across treatments, but notably a novel herpes-like virus increased by up to 6 orders of magnitude on exposure to abiotic stress [387], although some caution may be warranted in assessing the reliability of such determinations [Wood-Charlson et al., submitted]. Unexpectedly, the proteome of an endosymbiont-enriched fraction of the coral Stylophora pistillata showed a significant 114-fold increase in a viral replication protein on thermal bleaching [39], which is consistent with the finding of VLP induction in P. compressa by similar treatment [387].

General aspects of histocompatibility [390-393] and the genetic structure of innate immune receptors of the Cnidaria [363,394-401], including the immune response effected by coral disease and bleaching [364,402], have been examined extensively, hence further elaboration here is unnecessary. Instead, we focus on proteins that directly regulate the pathogenicity of coral-associated microbes and viruses. The A. digitifera genome encodes protein homologues having either putative antiviral and virus-promoting activities (Table 14). These homologues include the antiviral “superkiller” helicase SKI2 protein that acts by blocking viral mRNA translation [403] and, together with the superkiller proteins SKI3 (69 sequence alignments) and SKI8 of the exosome complex, function in a 3’-mRNA degradation pathway [404]. The coral genome encodes also three exoribonuclease (RNase) enzymes (XRN, XRN2 and RNB) with antiviral RNA-degrading properties [405,406]. Annotation of the coral genome reveals homologues to four interferon proteins (IFNB, IFNG, IFNW1 and IFNT1). Interferons are potent and selective antiviral cytokines [407], which are induced by viral infection or by sensing dsRNA, a by-product of viral replication, leading to the transcription of interferon-stimulated genes whose products have antiviral activities and others having antimicrobial, antiproliferative/antitumor or immunomodulatory effects [408,409]. Included in the coral antivirus defence system are three members of the interferon regulatory transcription factor (IRF1, IRF2 and IRF8) family proteins. IRF1 and IRF2 are transcriptional activators of cytokines and other target genes [410]; IRF1 is known to trans-activate the tumor suppressor protein p53 [411] while IRF2 regulates post-
transcriptional induction of NO synthase [412]. Conversely, IRF8 is an interferon consensus sequence-binding protein that is a negative (interference) regulator of enhancer elements common to interferon-inducible genes [413]. The coral genome additionally includes an interferon-stimulated 20 kDa protein (ISG20) RNase specific to deactivation of singled-stranded RNA viruses [414]. The coral genome encodes several interferon-inducible proteins, notably interferon gamma induced GTPase (IGTP) that accumulates in response to IFNB [415], the interferon-induced GTP-binding protein Mx1 that is a key element of host antiviral defence [416], the interferon-induced helicase C domain-containing protein 1 (aka MDA-5), which is an immune receptor that senses viral dsRNA to activate the interferon antiviral-response cascade [417] and the interferon-induced transmembrane protein (IFITM1) that suppresses cell growth [418]. The coral genome encodes the interferon-gamma

![Table 14 Regulatory and related proteins of viral pathogenicity in the predicted proteome of A. digitifera](image-url)

**Table 14 Regulatory and related proteins of viral pathogenicity in the predicted proteome of A. digitifera**

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.20647; v1.06188; v1.21287 | K12599 | Antiviral helicase SK2 |
| v1.18443 [+ 40 other sequence copies] | K12807 | Baculoviral IAP repeat-containing protein 1 (BIRC1) |
| v1.06263 [+ 6 other sequence copies] | K04725 | Baculoviral IAP repeat-containing protein 2/3/4 (BIRC2/3/4) |
| v1.14355 | K08731 | Baculoviral IAP repeat-containing protein 5 (BIRC5) |
| v1.04171 [+ 7 other sequence copies] | K10586 | Baculoviral IAP repeat-containing protein 6 (BIRC6) |
| v1.12348; v1.01945; v1.16612 | K06731 | Bone marrow stromal cell antigen 2 (antiviral BST2) |
| v1.01539 [+ 7 other sequence copies] | K04012 | Complement component receptor 2 (CR2) |
| v1.17305 | K04462 | Ecotropic virus integration site 1 protein (EV11) |
| v1.1406 [+ 4 other sequence copies] | K12618 | 5′-3′ Exoribonuclease 1 (antiviral XRN1) |
| v1.22746; v1.19002; v1.12850; v1.21216 | K12619 | 5′-3′ Exoribonuclease 2 (antiviral XRN2) |
| v1.09005 | K01147 | Exoribonuclease II (antiviral RNB) |
| v1.22793; v1.12978; v1.19008; v.20838 | K09239 | HIV virus type I enhancer-binding protein (HIVEP) |
| v1.02776 [+ 7 other sequence copies] | K15046 | Interferon virus NS1A-binding protein (NS1A-BP) |
| v1.09829; v1.13077 | K05415 | Interferon beta (IFNB) |
| v1.11946; v1.21512; v1.11221; v1.11927 | K04687 | Interferon gamma (IFNG) |
| v.121512 | K14140 | Interferon gamma induced GTPase (ITGP) |
| v.11946 | K05133 | Interferon gamma receptor 2 (IFNGR2) |
| v.01539 [+ 4 other sequence copies] | K04012 | Interferon-induced GTP-binding protein Mx1 |
| v.10782; v.23797; v.17119; v.03221 | K12647 | Interferon-induced helicase C domain-containing protein 1 |
| v.06274; v.1.15849; v.05943 | K06566 | Interferon induced transmembrane protein (IFITM1) |
| v.21327; v.24081 | K05440 | Interferon, omega 1 (IFNW1) |
| v.11817 | K09444 | Interferon regulatory factor 1 (IRF1) |
| v.11816; v.07639 | K10153 | Interferon regulatory factor 2 (IRF2) |
| v.11421 | K10155 | Interferon regulatory factor 8 (IRF8) |
| v.02158 | K12579 | Interferon-stimulated gene 20 kDa protein (ISG20) |
| v.15947 | K05442 | Interferon tau-1 (IFNT1) |
| v.122825; v.08034; v.08520 | K05788 | Integration host factor subunit beta (IHB) |
| v.14899 | K08220 | MFS transporter, FLVCR family virus subgroup C receptor |
| v.04514; v.04513; v.16929 | K12648 | Mitochondrial antiviral-signalling protein (MAVS) |
| v.17718; v.08002; v.08001; v.22382 | K06801 | Poliovirus receptor-related protein 1 (PVRL1) |
| v.21413; v.06637 | K06531 | Poliovirus receptor-related protein 2 (PVRL2) |
| v.11740; v.21467; v.11410; v.17135 | K06592 | Poliovirus receptor-related protein 3 (PVRL3) |
| v.15077 | K06593 | Poliovirus receptor-related protein 4 (PVRL4) |
| v.04158 [+ 68 other sequence copies] | K12600 | Superkiller protein 3 (antiviral SHI3) |
| v.18238 [+ 4 other sequence copies] | K12601 | Superkiller protein 8 (antiviral SHI8) |
receptor 2 (IFNγR2) transmembrane protein that activates downstream signal transduction cascades that control cell proliferation and apoptosis [419]. Encoded also is a homologue of the human bone marrow stromal cell antigen 2 (BST2) that inhibits retrovirus infection by preventing VLP release from infected cells [420]. Additionally encoded is a mitochondrial antiviral-signalling protein (MAVS) that triggers the host immune response by activation of the nuclear transcription factor NF-κB and the interferon regulatory transcription factor IRF3 which coordinates the expression of type-1 interferons such as IFNβ [421].

The coral genome encodes a full set of baculoviral IAP repeat-containing proteins BIRC 1-6 (Table 14). The IAP (inhibitor of apoptosis) family proteins were first identified secreted by baculovirus to protect infected cells from death in the progression of viral replication [422]. Expressed by most eukaryotic organisms (reviewed in [423]), their IAP function is presumably conserved in corals. The coral genome encodes a full set of poxvirus receptor-related proteins (PVRL1-4) of the immunoglobulin superfamily, which bind and transport herpesvirus-related proteins (PVRL1-4) of the immune system. The coral genome encodes a full set of baculoviral IAP repeat-containing proteins BIRC 1-6 (Table 14). The IAP (inhibitor of apoptosis) family proteins were first identified secreted by baculovirus to protect infected cells from death in the progression of viral replication [422]. Expressed by most eukaryotic organisms (reviewed in [423]), their IAP function is presumably conserved in corals. The coral genome encodes a full set of poxvirus receptor-related proteins (PVRL1-4) of the immunoglobulin superfamily, which bind and transport herpesvirus-related proteins (PVRL1-4) of the immune system. The coral genome encodes a full set of baculoviral IAP repeat-containing proteins BIRC 1-6 (Table 14). The IAP (inhibitor of apoptosis) family proteins were first identified secreted by baculovirus to protect infected cells from death in the progression of viral replication [422]. Expressed by most eukaryotic organisms (reviewed in [423]), their IAP function is presumably conserved in corals. The coral genome encodes a full set of poxvirus receptor-related proteins (PVRL1-4) of the immunoglobulin superfamily, which bind and transport herpesvirus-related proteins (PVRL1-4) of the immune system.

Toxins and venom
A review of protein sequences deposited in the UniProt database in October 2012 shows that there are 150 known cnidarian toxins. These toxins have diverse biological activities (neurotoxins, pore-forming cytolytins and venom phospholipases) used to capture prey and for protection against predators [432] that are best characterised in sea anemones (Actiniaria) with 141 sequences deposited [433,434]. The cytotoxin MCTx-1 isolated from the Net Fire Coral Millepora dichotoma is the only toxin from a coral deposited in Uniprot (accession number ABQZJ5). However, our initial examination of the predicted proteome of A. digitifera shows 18 proteins with similarity to bacterial toxins and associated regulatory proteins (Table 15). Unlike reports from proteomic examination of the coral S. pistillata [39] and nematocysts (stinging organelles) of the jellyfish Chrysaora quinquecirrha [435], Tamoya haplonema, Chironemus quadrumanus, Chrysaora lactea (PF Long et al., pers comm), by sea anemones [434] and by the highly dangerous box jellyfish Chironex fleckeri [436,437], no venoms typical of higher animals were found in the A. digitifera genome. This was because our annotation was carried out using the KEGG database (release v58 [53]) to relate A. digitifera protein sequences to KEGG orthologues. The KEGG database is a collection of proteins from well characterised and ubiquitous biochemical pathways. Animal venoms, however, are highly specialised proteins for which this release of the KEGG database does not contain any described orthologues.

KEGG orthology-based annotation of the A. digitifera genome reveals genes encoding protein homologues of 10 bacterial toxins, 7 regulatory toxin proteins and a botulinum protein substrate (Table 15). Of the 9 toxin homologues, one with similarity to anthrax edema factor (EF) adenylate cyclase (CyA) is one of three proteins that comprise the anthrax toxin of Bacillus anthracis, the other two being a protective antigen (PA) and lethal factor (LF). Without the LF protein, anthrax CyA has no known toxic effects in animals [438], although the EF protein does play an important role in disabling
cellular functions vital for microbial host defences [439]. The *A. digitifera* genome encodes a secretion virulence factor exotoxin A-like protein produced by *Pseudomonas aeruginosa*, which for this bacterium affects local tissue damage, bacterial invasion and immunosuppression within their eukaryote host [440] with pathogenicity similar to that of the diphtheria toxin [441]. Another encoded protein is a murine-like toxin (Ymt) produced by the enterobacterium *Yersinia pestis*, which is the causative agent responsible for transmission of the notorious bubonic plague [442]. Additionally, two hemolytic enterotoxins similar to NheA and NheBC produced by *Bacillus cereus* [443], an enterotoxin (EntA) similar to that of *Staphylococcus aureus* [444], a Shiga-like enterotoxin (StxB) produced by *Shigella dysenteria*, the diarrhea-causing toxin A/B (TcdAB) such as that secreted by *Clostridium difficile* [445], and a protein similar to the zonula occludens (tight junction) enterotoxin (Zot) secreted by *Vibrio cholera* [446] are encoded in the *A. digitifera* genome. Within the predicted proteome is also a homologue of the vacuolating cytotoxin (VacA) produced by *Helicobacter pylori* that colonises the gastric mucosa of the human stomach epithelium [447].

Although a direct homologue of the cholera toxin (CT) was not found encoded in the *A. digitifera* genome (Table 15), a protein similar to its transcriptional activator ToxR was. ToxR not only controls the expression of CT in *Vibrio cholera* [448], but also a co-regulated pilin (TcpA) protein that is under control of the ToxR regulon cascade [449]. Bacterial TcpA protein is assembled into toxin-coregulated pili that induce the transfer of DNA by horizontal exchange of genetic material during conjugation [450]. TcpA and two toxin co-regulated biosynthetic proteins (TcpI and TcpS) of the bacterial virulence-associated pilus appendage [451] are encoded in the coral genome. Entrained also are the motility quorum-sensing interference regulator MsqR and its transcriptional regulator MsqA that in *Escherichia coli* controls biofilm formation by inhibiting quorum-sensing motility, and together the MsqR/MsqA complex represses the lethal cold shock-like protein cspD gene [452] that on expression impairs DNA replication [453]. The *A. digitifera* genome likewise encodes a Type III secretion system T3SS cytotoxic effector (RbeA) protein [454] that in Gram-negative invasive bacteria is translocated into host cells to suppress innate immunity to enhance virulence [455,456]. However, the ecophysiological significance of these toxigenic proteins and allied regulators, if indeed expressed by the coral genome, is unknown.

In addition to using the KEGG database, we undertook a BLAST search of the predicted proteome of *A. digitifera* against peptide sequences for all animal venoms using the annotated UniProtKB/Swiss-Prot Tox-Prot program [457]. This search revealed a large number of accession hits from the predicted proteome, although these are unlikely to be true multiple copies given that the genome sequence has yet to be completely assembled. However, just taking a single accession number from each annotation reveals a complex array of 83 toxins that represents the predicted venom of *A. digitifera* (Table 16); UniProt BLAST E-values are given in Additional file 1: Table S16b. These venoms are highly diverse and are significantly homologous to toxins from a wide variety of venomous marine and terrestrial creatures such as fish, reptiles, other cnidarians, cone-snails, stinging insects and even a venomous mammal (Shrew), covering the complete range of pharmacological properties known in venoms, including cytolytic, neurotoxic, haemotoxic, phospholipase, proteinase and proteinase inhibitor activities. Both the number of toxins predicted in the venom of *A. digitifera* and the degree of homology to such widely divergent phyla is remarkable. Accordingly, cnidarian venoms may possess unique biological properties that might generate new

| Gene sequence | KEGG Orthology | Encoded protein description |
|---------------|----------------|-----------------------------|
| v1.20214      | K11029         | Anthrax edema toxin adenylate cyclase (CyaA) |
| v1.17686      | K10921         | Cholera toxin transcriptional activator (ToxT) |
| v1.13017      | K11020         | Exotoxin A (ToxA) |
| v1.23507      | K13655         | HTH-type transcriptional regulator (MsqA) antitoxin for MqsR |
| v1.21184      | K11009         | Murine toxin (Ymt) |
| v1.04313      | K11033         | Non-hemolytic enterotoxin A (NheA) |
| v1.08011      | K11034         | Non-hemolytic enterotoxin B/C (NheBC) |
| v1.08255      | K13651         | Motility quorum-sensing regulator (MqsR) interferase toxin |
| v1.15986      | K11059         | Probable enterotoxin A (EntA) |
| v1.13046      | K04392         | Ras-related C3 botulinum toxin substrate 1 (Rac1) |
| v1.13966      | K11007         | Shiga toxin subunit B (StxB) |
| v1.23958      | K11063         | Toxin A/B (TcdAB) |
| v1.21174      | K10930         | Toxin co-regulated pilin (TCP) |
| v1.05802      | K10961         | Toxin co-regulated pilus biosynthesis protein (Tcpl) |
| v1.21783      | K10964         | Toxin co-regulated pilus biosynthesis protein 5 (TcpS) |
| v1.14688      | K15126         | Type III secretion system cytotoxic effector protein (RbeA) |
| v1.05520      | K11028         | Vacuolating cytotoxin (VacA) |
| v1.06590      | K10954         | Zona occludens toxin (Zot) |
| Gene sequence | UniProt toxin accession | Animal with closest homology |
|---------------|-------------------------|------------------------------|
| v1.01916      | Q92035; Acetylcholinesterase | Bungarus fasciatus (Banded Krait) |
| v1.06761; v1.08075; v1.09940; v1.20523 | Q9IAM1; Agkisacutacin (subunit anticoagulant protease) | Deinagkistrodon acutus (Sharp-nosed Viper) |
| v1.04809      | A8QL52; L-Amino acid oxidase | Bungarus fasciatus (Banded Krait) |
| v1.06380      | Q4JHE1; L-Amino acid oxidase | Pseudechis australis (Mulga Snake) |
| v1.10291      | P81383; L-Amino acid oxidase | Ophiophagus hannah (King Cobra) |
| v1.14411      | A6MFL0; L-Amino acid oxidase | Demansia vestigata (Lesser Black Whipsnake) |
| v1.16469      | P81383; L-Amino acid oxidase | Ophiophagus hannah (King Cobra) |
| v1.23477      | P81382; L-Amino acid oxidase | Calloselasma rhodostoma (Malayan Pit Viper) |
| v1.16440      | CSN5L2; Bandaporin (haemolysin) | Anthropleura asiatica (Sea Anemone) |
| v1.16571      | Q76B45; Blarina toxin (vasoactive protease) | Blarina brevicauda (Northern Short-Tailed Shrew) |
| v1.06055      | Q593B6; Coagulation factor V | Pseudonaja textilis (Eastern Brown Snake) |
| v1.07831; v1.10094; v1.20732 | P14530; Coagulation factor IX | Protobothrops flavoviridis (Okinawa Habu Snake) |
| v1.01708      | Q4JXT9; Coagulation factor X | Tropidechis carinatus (Rough-Scaled Snake) |
| v1.09601; v1.10410 | Q93109; Equinatoxin-5 (cytolsin) | Actinia equina (Beadlet Anemone) |
| v1.06621      | Q08169; Hyaluronidase | Apis mellifera (European Honey Bee) |
| v1.08924      | I0CM7; Hyaluronidase, Conohyal-Cn1 | Conus consors (Singed Cone) |
| v1.06199      | Q2XZO; α-Latrocruztotoxin Lt1a (neurotoxin) | Latrodectus tredecimguttatus (Mediterranean Black Widow Spider) |
| v1.02942      | GOLV8; α-Latrocruztotoxin Lh1a (neurotoxin) | Latrodectus hasseltii (Australian Redback Spider) |
| v1.06444      | Q2S338; Δ-Latrocruztectotoxin Lt1a (neurotoxin) | Latrodectus tredecimguttatus (Mediterranean Black Widow Spider) |
| v1.07446      | A7X6X3; Lectin, Lectoxin Enh4 (platelet binding) | Enhydris polyplepis (Macleay’s Water Snake) |
| v1.20653      | A7X3Y6; Lectin, Lectoxin Enh7 (platelet binding) | Enhydris polyplepis (Macleay’s Water Snake) |
| v1.02561; v1.11493; v1.16681 | A7X3Z4; Lectin, Lectoxin Lio1 (platelet binding) | Liophis poecilogyrus (Water Snake) |
| v1.13597; v1.08696; v1.10757; v1.20554 | A7X3Z7; Lectin, Lectoxin Lio2 (platelet binding) | Liophis poecilogyrus (Water Snake) |
| v1.18386; v1.15479 | A7X413; Lectin, Lectoxin Lioc3 (platelet binding) | Liophis poecilogyrus (Water Snake) |
| v1.06094      | A7X406; Lectin, Lectoxin Phi1 (platelet binding) | Philodryas oiffersii (Green Cobra) |
| v1.06416; v1.16248; v1.23712 | A7X3Z0; Lectin, Lectoxin Thr1 (platelet binding) | Thrasops jacksonii (Black Tree Snake) |
| v1.17681      | Q6TPG9; Lectin, Mucroctein (platelet binding) | Protobothrops mucrosquamatus (Brown Spotted Pit Viper) |
| v1.12241; v1.02332; v1.12298 | Q71RQ1; Lectin, Stejaggregin-A (platelet binding) | Trimeresurus stejnegeri (Bamboo Viper) |
| v1.02245      | A0FKN6; Metalloprotease, Astacin-like toxin | Loxosceles intermedia (Recluse Spider) |
| v1.03638; v1.14772 | Q90391; Metalloprotease, Atrolysin | Crotalus atrox (Western Diamondback Rattlesnake) |
| v1.13106      | D3TTC2; Metalloproteinase, Atragen | Naja atra (Chinese Cobra) |
| v1.11132      | Q7T1T4; Metalloproteinase, BjussuMP-2 | Bothrops jararacussu (Jararacussu Pit Viper) |
| v1.02168      | O73795; Metalloproteinase, Disintegrin | Gloydius brevicaudus (Chinese Mambushi Snake) |
| v1.06910      | Q7SZE0; Metalloproteinase, Disintegrin | Gloydius saxatilis (Rock Mamushi Snake) |
| v1.22282      | P14530; Metalloproteinase, Disintegrin | Protobothrops flavoviridis (Okinawa Habu Snake) |
| UniProt ID | Gene Name | Protein Name | Species | Function |
|-----------|-----------|--------------|---------|----------|
| Q2UKQ5    | v1.03804  | Metalloproteinase, EoVMP2 | Echis ocellatus | (West African Carpet Viper) |
| Q91511    | v1.02016  | Mucrofibrase-5, Hypotensive serine protease | Protobothrops mucrosquamatus | (Brown Spotted Pit Viper) |
| Q7ZZN8    | v1.09026  | Natrin-2 (neurotoxin) | Naja atrocyclus | (Chinese Cobra) |
| A0ZSK3    | v1.04153  | Neoverrucotoxin (haemolysin) | Synanceia verrucosa | (Reef Stone Fish) |
| Q06478    | v1.12300  | Phospholipase A1 | Dolichovespula maculata | (Bald-Faced Hornet) |
| P0CH47    | v1.04157  | Magnin | Vespa velutina | (Giant Wasp) |
| P53357    | v1.11796  | Basic | Polybia paulistens | (Southern Wasp) |
| Q06478    | v1.09028  | Phospholipase A1 | Polybia paulistens | (Southern Wasp) |
| P0CH47    | v1.12433  | Basic | Polybia paulistens | (Southern Wasp) |
| Q7SSN0    | v1.09855  | Prothrombin activator Pseutarin-C | Pseuderpes textilis | (Eastern Brown Snake) |
| P82370    | v1.10405  | Prothrombin activator Hopsarin-D | Hoplocephalus stephensii | (Stephen’s Branded Snake) |
| Q58L94    | v1.00387  | Prothrombin activator Notecarin D2 | Notelaps scutatus | (Tiger Snake) |
| Q58L93    | v1.00618  | Prothrombin activator Porpharin D | Pseudechis porphyriacus | (Red-Bellied Black Snake) |
| Q58L93    | v1.00618  | Prothrombin activator Porpharin D | Pseudechis porphyriacus | (Red-Bellied Black Snake) |
| P81428    | v1.09896  | Prothrombin activator Trocarin D | Trimeresurus gramineus | (Bamboo Viper) |
| A6MF87    | v1.13726  | Prothrombin activator Vestarin D1 | Demansia vesparia | (Lesser Black Whip Snake) |
| Q6T269    | v1.02128  | Protease inhibitor, Bitisilin-3 (neurotoxin) | Bitis gabonica | (Gaboorn Viper) |
| Q58L93    | v1.02137  | Prothrombin activator Porpharin D | Pseudeschis porphyriacus | (Red-Bellied Black Snake) |
| Q58L93    | v1.02137  | Prothrombin activator Porpharin D | Pseudeschis porphyriacus | (Red-Bellied Black Snake) |
| P82370    | v1.09855  | Prothrombin activator Hopsarin-D | Hoplocephalus stephensii | (Stephen’s Branded Snake) |
| Q58L93    | v1.00618  | Prothrombin activator Porpharin D | Pseudeschis porphyriacus | (Red-Bellied Black Snake) |
| Q58L93    | v1.00618  | Prothrombin activator Porpharin D | Pseudeschis porphyriacus | (Red-Bellied Black Snake) |
| P82370    | v1.09855  | Prothrombin activator Hopsarin-D | Hoplocephalus stephensii | (Stephen’s Branded Snake) |
| Q58L93    | v1.00618  | Prothrombin activator Porpharin D | Pseudeschis porphyriacus | (Red-Bellied Black Snake) |
leads in the discovery of novel pharmacologically active drugs. Gene duplication followed by mutation and natural selection is widely held as the key mechanism whereby the large diversity of toxins found within a single venom could have evolved [458,459]. Conversely, primary mRNA splicing patterns have been shown to account for the diversity of metallopeptases in the pit viper Bothrops neuwiedi [460]. Variations in peptide processing have also been shown by proteomics and transcriptomics to explain how a limited set of genes transcripts could generate thousands of toxins in a single species of cone snail [461]. Despite these various processes that could account for the evolution of toxin diversity, it has never been demonstrated how gene duplications or variations in transcript or peptide processing could have radiated across the very different poisonous creatures found on Earth. Our data (Table 16) reveal that the predicted toxins of A. digitifera venom are orthologues to all of the most important superfamilies of peptide/protein venoms found in diverse taxa. We posit that the origins of toxins in the venoms of higher organisms may have arisen from deep eumetazoan innovations and that the molecular evolution of these venom super gene families can now be addressed using an integrated venomics approach using Cnidaria such as the jellyfish as model systems [462].

Detoxification proteins of the chemical defensome

There have been considerable advancements made to better understand the effects of pollution on coral reef habitats. The three main categories of environmental pollutants from anthropogenic sources are nutrient enrichment (eutrophication), hydrocarbon pollution and heavy metal contamination. Eutrophication from terrestrial inputs are a significant threat to coral reefs stemming from the discharge of treated sewage, the runoff of agricultural fertilizers (plus herbicides and pesticides), and by sedimentation caused by the erosion of organic-rich soils [463]. Notwithstanding that eutrophication can shift coral reef communities towards macroalgae domination [19], nitrogen and phosphorus enrichment can diminish coral growth and affect the photosynthetic performance of their algal symbionts [464]. Nutrient enhancement alters multiple pathways of primary metabolism that in coral is complicated by the photosynthetic demands of its symbiotic partners. While corals respond to hypertrophic levels of nutrients by activating general stress-response proteins [465], there are no specific proteins known to mitigate the cellular effects of nutrient enrichment on corals per se, and we have not attempted to identify such in this study.

Gene families and their regulators that defend against chemical stressors comprise the chemical defensome encoding a network of detoxifying proteins that allows an organism to sense, transform and eliminate potentially toxic endogenous metabolites and xenobiotic contaminants [466]. Expressed proteins of the chemical defensome include the biotransformation cytochrome P450 (CYP) family of enzymes, conjugating enzymes, efflux transporters, heavy metal membrane pump exporters and their transcriptional activators. Annotation of the genome of A. digitifera reveals multiple genes encoding 20 hemoproteins belonging to the Phase II cytochrome P450 superfamily of monooxygenase enzymes that catalyse the oxidation of diverse organic substances (Table 17). The substrates of CYP enzymes include intermediates of lipid metabolism and steroid/steroid biosynthesis, and include the detoxification of exogenous xenobiotics. Of significance are the CYP1A-type (aryl hydrocarbon hydroxylase) enzymes that have been studied widely in the hepatic response of fishes to polycyclic aromatic hydrocarbon (PAH) contamination (from crude or fuel oil) and exposure to polychlorinated biphenyl and

Table 16 UniProt-predicted homologues of animal venom proteins in the predicted proteome of A. digitifera (Continued)

| UniProt ID | Protein Name | Species Name |
|------------|--------------|--------------|
| v1.05361   | B6RLX2       | Serine protease inhibitor, TCI (neurotoxin) | Ophiophagus hannah (King Cobra) |
| v1.0994    | B7S4N9       | Serine protease inhibitor, Taicatoxin (neurotoxin) | Oxyuranus scutellatus (Coastal Taipan) |
| v1.11218   | Q00WAO      | Serine protease inhibitor, Textilinin-2 (thrombin inhibitor) | Pseudonaja textilis (Eastern Brown Snake) |
| v1.17856   | Q8T375      | Serine protease inhibitor, U1-aranetoxin-Av1a (neurotoxin) | Araneus ventricosus (Devil Spider) |
| v1.04154   | [- 4 other sequence copies] Q98989; Stonustoxin (haemostasis inhibitor) | Synanceia horrida (Estuarine Stonefish) |
| v1.09427, v1.16619, v1.19446 | Q76DT2, Toxin AvTX-60A (cytolyisin) | Actinera villosa (Okinawan Sea Anemone) |
| v1.12311   | Q9GV72      | Toxin CrTX-A (haemolysin) | Carybdea rastonii (Jimble Jellyfish) |
| v1.07546   | [- 5 other sequence copies] P58911; Toxin PsTX-60 (haemolysin) | Phyllodiscus semenoi (Night Anemone) |
| v1.11270, v1.14265 | E2IYB3; Veficolin-1 (complement activator) | Vananus komodoensis (Komodo Dragon) |
| v1.02115   | Q98993      | Verrucotoxin (cytolyisin) | Synanceia verrucosa (Reef Stonefish) |
### Table 17 Proteins of the chemical defensome in the predicted proteome of A. digitifera

| Gene sequence | KEGG Orthology | Encoded protein description                                      |
|---------------|----------------|-----------------------------------------------------------------|
| v1.06127; v1.06128 | K01015         | Alcohol sulfotransferase                                         |
| v1.09267          | K00537         | Arsenate reductase                                                |
| v1.24496; v1.24495; v1.03953 | K03893         | Arsenical pump membrane protein                                   |
| v1.10691          | K07755         | Arsenite methyltransferase                                        |
| v1.20443          | K11811         | Arsenical resistance protein ArsH                                 |
| v1.14972          | K01551         | Arsenite-transporting ATPase                                      |
| v1.17644; v1.00480; v1.08150; v1.22865 | K07408         | Cytochrome P450, family 1, subfamily A, polypeptide 1            |
| v1.21535; v1.11835; v1.02456 | K01014         | Alcohol sulfotransferase                                          |
| v1.03485; v1.21926; v1.05686 | K01534         | Cd$^{2+}$/Zn$^{2+}$-exporting ATPase                              |
| v1.10691          | K07755         | Arsenite methyltransferase                                        |
| v1.14972          | K01551         | Arsenite-transporting ATPase                                      |
| v1.22646 [+ 8 other sequence copies] | K07421         | Cytochrome P450, family 2, subfamily T                            |
| v1.10544; v1.02314; v1.17490 | K07422         | Cytochrome P450, family 2, subfamily U                            |
| v1.23039 [+ 13 other sequence copies] | K07425         | Cytochrome P450, family 4, subfamily A                            |
| v1.07750          | K07426         | Cytochrome P450, family 4, subfamily B                            |
| v1.22798; v1.23000 | K07427         | Cytochrome P450, family 4, subfamily V                            |
| v1.02020 [+ 4 other sequence copies] | K07428         | Cytochrome P450, family 4, subfamily X                            |
| v1.19495          | K15002         | Cytochrome P450, family 6                                         |
| v1.15382          | K07430         | Cytochrome P450, family 7, subfamily B                            |
| v1.16427          | K00498         | Cytochrome P450, family 11, subfamily A                           |
| v1.17631          | K15004         | Cytochrome P450, family 12                                         |
| v1.08074 [+ 4 other sequence copies] | K00512         | Cytochrome P450, family 17, subfamily A                           |
| v1.02478 [+ 5 other sequence copies] | K07435         | Cytochrome P450, family 20, subfamily A                           |
| v1.06713          | K07436         | Cytochrome P450, family 24, subfamily A                           |
| v1.22414 [+ 5 other sequence copies] | K12665         | Cytochrome P450, family 26, subfamily C                           |
| v1.20153          | K00488         | Cytochrome P450, family 27, subfamily A                           |
| v1.08074 [+ 6 other sequence copies] | K07439         | Cytochrome P450, family 39, subfamily A                           |
| v1.06537          | K07440         | Cytochrome P450, family 46, subfamily A                           |
| v1.22302 [+ 5 other sequence copies] | K09832         | Cytochrome P450, family 710, subfamily A                          |
| v1.16335          | K01016         | Estrone sulfotransferase                                          |
| v1.18439; v1.02594; v1.02593 | K00699         | Glucuronosyltransferase                                          |
| v1.07758 [+ 5 other sequence copies] | K13299         | Glutathione S-transferase kappa 1                                 |
| v1.00764          | K0799          | Glutathione S-transferase                                         |
| v1.17188          | K07239         | Heavy-metal exporter, HME family                                  |
| v1.04140          | K00481         | p-Hydroxybenzoate 3-monoxygenase                                  |
| v1.01818          | K08365         | MerR family transcriptional regulator, mercuric resistance        |
| v1.16748; v1.07471 | K13638         | MerR family transcriptional regulator, Zn(II)-responsive          |
| v1.04382; v1.24424 | K08363         | Mercuric ion transport protein                                    |
| v1.12760          | K03284         | Metal ion transporter, MIT family                                 |
| v1.04179; v1.01891; v1.00145 | K01253         | Microsomal epoxide hydrolase                                     |
| v1.21500 [+ 5 other sequence copies] | K08970         | Nickel/cobalt exporter                                           |
| v1.08005          | K08364         | Periplasmic mercuric ion binding protein                          |
| v1.03484          | K07245         | Putative copper resistance protein D                             |
dibenzo-dioxin toxicants (reviewed in [467]). CYP450 activity has been detected in the corals *Favia fragum* [468], *Siderastrea siderea* [469], *Montastraea faveolata* [470] and *Pocillopora damicornis*, [471]. Furthermore, CYP encoding sequences have been extracted from the genome of *N. vectensis* [472] and the transcriptome of *A. millepora* [29]. As well as providing chemical defence, mixed-function CYPs perform multiple endogenous tasks that are often taxon-specific. Hence, the orthology and substrate specificity of coral CYP enzymes cannot be predicted solely on homology to CYPs of known function assigned to higher metazoans. Similar to the function of CPY enzymes, there are genes encoding p-hydroxybenzoate 3-monoxygenase, an oxidoreductase catalyzing aryl oxidation and the soluble and microsomal forms of epoxide hydrolase that converts epoxides, formed by the degradation of aromatic compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Conjugating enzymes to eliminate hydroxylated compounds, to trans-diols that by conjugation are readily excreted. Paralogs of bacteria [491]. Enzymes specific for arsenic detoxification are an arsenate oxidoreductase for conversion of arsenate to arsenite [492] and arsenite methyltransferase for conversion of arsenite to the less toxic dimethylarsenite that is amenable to excretion [493]. Such processes may enhance the resilience of corals exposed to natural [494] and site-affected [495] levels of arsenic contamination. In contrast, there were no (organo)cyanide detoxification genes apparent in the *A. digitifera* genome, but one sequence (v1.01601; K10814) encodes for hydrogen cyanide synthase of unknown metabolic purpose (data not tabulated). Ancillary evidence suggests that the expression of HCN synthase could be linked to quorum sensing [496] for regulating microbial densities of the coral holobiont community.

### Epigenetic and DNA-remodelling proteins

In all Kingdoms of life, DNA methylation and chromatin remodelling is pivotal to the regulation of gene transcription independent of underlying allelic variation. One such process mediated by epigenetic changes in eukaryotic biology is the all-important cellular differentiation during morphogenetic development. Epigenetic modifications cause the activation, regulation or silencing of certain genes without changing the basic DNA code. Changes in epigenetic regulation can persist during cell division and across multiple generations [497]. In addition, cytosine methylation may be associated with a higher mutation rate, because deamination of the methylated base produces thymine resulting in C/T mutations, which on reproduction may be transmitted by the germline to subsequent generations in selective processes of evolution [498]. On the other hand, environmentally induced destabilisation of the epigenome can produce epigenetic gene variants (epialleles) that activate transcription and mobilization of DNA transposable elements, which may subsequently lead to stable heritable traits of environmental adaptation, as does occur by genetic imprinting in plants [499]. Transposition has thus the potential to direct increased frequencies of permanent genetic mutations for selective adaptation.
One way by which genes are regulated at the epigenome is through the remodelling of the chromatin histone-DNA complex (the nucleosome), which by post-translational modification changes the template structure of DNA associated histone proteins. These modifications are affected by histone-lysine (and histone-arginine) N-methyltransferase enzymes (Table 18) by which these proteins may be further modified by acetylation, ADP-ribosylation, ubiquination, and phosphorylation (annotation not tabulated).

The methylation pattern of histone lysine residues is highly predictive of the gene expression states of transcriptional activation and repression [500]. Necessary epigenomic re-programming of histone modification at different stages of cell development is affected by the activation of histone and lysine-specific demethylase enzymes (Table 18). Determinants for recognition of the histone code are being revealed by a growing body of experimental data providing valuable information on the molecular tractability of

---

**Table 18 Epigenetic and DNA-remodelling proteins in the predicted proteome of A. digitifera**

| Gene sequence   | KEGG Orthology | Encoded protein description                                                                 |
|-----------------|----------------|---------------------------------------------------------------------------------------------|
| v.04242; v.02042| K02528         | 16S rRNA (adenine1518-N6/1519-N6)-dimethyltransferase                                         |
| v.22358; v.00249| K14191         | 18S rRNA (adenine1779-N6/1780-N6)-dimethyltransferase                                         |
| v.14000; v.04238| K00561         | 23S rRNA (adenine2085-N6)-dimethyltransferase                                                |
| v.05107; v.05242| K01488         | Adenosine deaminase                                                                          |
| v.04152; v.09790| K14857         | AdoMet-dependent rRNA methyltransferase SPB1                                                  |
| v.00197          | K13530         | AraC family transcriptional regulator DNA methyltransferase                                   |
| v.12967; v.19789; v.07763 | K14589 | Cap-specific mRNA (nucleoside-2’-O-)-methyltransferase 1                                      |
| v.24281          | K01489         | Cytidine deaminase                                                                           |
| v.16211; v.14952; v.01094; v.06093 | K00558 | DNA (cytosine-5)-methyltransferase                                                           |
| v.19683; v.05688; v.04223 | K11324 | DNA methyltransferase 1-associated protein 1                                                   |
| v.14033; v.19860; v.19081; v.04188 | K11420 | Euchromatic histone-lysine N-methyltransferase                                                   |
| v.02068          | K01487         | Guanine deaminase                                                                           |
| v.02920          | K05931         | Histone-arginine methyltransferase CARM1                                                      |
| v.17589 [+ 7 other sequence copies] | K11446 | Histone demethylase JARID1                                                                   |
| v.07640          | K06101         | Histone-lysine N-methyltransferase ASH1                                                        |
| v.13515; v.18577; v.20187; v.19182 | K09186 | Histone-lysine N-methyltransferase MLL1                                                        |
| v.08381          | K09187         | Histone-lysine N-methyltransferase MLL2                                                        |
| v.24258; v.19182 | K09188         | Histone-lysine N-methyltransferase MLL3                                                         |
| v.07992; v.10302; v.13829 | K09189 | Histone-lysine N-methyltransferase MLL5                                                        |
| v.06939; v.15255; v.15254 | K11424 | Histone-lysine N-methyltransferase NSD1/2                                                       |
| v.05552          | K11422         | Histone-lysine N-methyltransferase SETD1                                                       |
| v.07744          | K11423         | Histone-lysine N-methyltransferase SETD2                                                        |
| v.03190          | K11431         | Histone-lysine N-methyltransferase SETD7                                                        |
| v.21867          | K11428         | Histone-lysine N-methyltransferase SETD8                                                        |
| v.18700 [+ 8 other sequence copies] | K11421 | Histone-lysine N-methyltransferase SETD8                                                        |
| v.07557; v.11409 | K11419         | Histone-lysine N-methyltransferase SLV39H                                                        |
| v.24733; v.13497 | K11429         | Histone-lysine N-methyltransferase SLV420H                                                        |
| v.15405; v.10291; v.17601; v.02845; v.08629 | K11450 | Lysine-specific histone demethylase 1                                                              |
| v.23155; v.09394; v.17624; v.05370 | K14835 | Ribosomal RNA methyltransferase Nop2                                                                  |
| v.18460 [+ 6 other sequence copies] | K03500 | Ribosomal RNA small subunit methyltransferase B                                                                 |
| v.07407; v.03110 | K08316         | Ribosomal RNA small subunit methyltransferase D                                                                 |
| v.12193          | K02427         | Ribosomal RNA large subunit methyltransferase E                                                                 |
| v.11499          | K11392         | Ribosomal RNA small subunit methyltransferase F                                                                 |
| v.16053; v.12676 | K03437         | RNA methyltransferase, TrmH family                                                                 |

---
Deaminated bases are subsequently removed by DNA demethylases that are encoded in the cytosine, adenosine and guanine deaminase enzymes [507]. Hypoxanthylated nucleobases are recognized for deamination by the cytosine by C/T mismatch repair enzymes. These methylation reactions require base-specific deamination by the AID/Apobec family of deaminase enzymes that, for example, converts 5-meC to thymine that is replaced subsequently by thymine. Another pathway for DNA demethylation, although there is encoded a 5-methylcytosine-specific restriction enzyme [506] for excision of 3-meG, but there are no such restriction enzyme. Another pathway for DNA demethylation occurs by several base-repair pathways involving DNA glycosylases, DNA mismatch repair enzymes and proteins linking symbiont photosynthesis to coral calcification. An extensive catalogue of mammalian-like proteins essential to neural function and venoms related to calcification. An extensive catalogue of mammalian-like proteins essential to neural function and venoms related to calcification. An extensive catalogue of mammalian-like proteins essential to neural function and venoms related to calcification.

**Conclusions**

We offer ZoophyteBase as an unprecedented foundation to interrogate the molecular structure of the predicted *A. digitifera* proteome. Some key findings include proteins with relevance to host-symbiont function, dysfunction and recovery including those that direct vacuolar trafficking and proteins linking symbiont photosynthesis to coral calcification. An extensive catalogue of mammalian-like proteins essential to neural function and venoms related to distant animal phyla suggests their origins lie deep in early eumetazoan evolution. Homologues of prokaryotic genes that have not been described previously in any eukaryote
The genome such as flagella proteins, proteins essential for nitrogen fixation and photosynthesis point towards lateral gene transfer, perhaps mediated by viruses, that may lead to “shared” metabolic adaptations of symbiosis, and provide corals with limited ability for gene-encoded adaptation to a changing global environment. It is anticipated that understanding how the genome of a coral hosts interacts with that of its vast array of symbionts, and how it may regulate its metabolic quotient, for example through biochemical or epigenetic modification, will rapidly accelerate our ability to predict the fate of coral reefs.

### Availability and requirements

ZoophyteBase was constructed using the Metagenome/Genome Annotated Sequence Natural Language Search Engine (MEGGASENSE). This is a general system for the annotation of sequence collections and presentation of the results in a database that can be searched using biologically intuitive search terms. In this implementation, the predicted proteome of *A. digitifera* (genome assembly v. 1.0 [48]) was used as the source of protein sequences. The annotation was carried out using the KEGG database (release v58 [51]) to relate *A. digitifera* protein sequences to KEGG orthologues. The homologous protein sequences were used to construct hidden Markov model (HMM) profiles using the HMMER3 package [49]. The predicted proteome sequences of *A. digitifera* were searched with the HMM profiles to link proteins to appropriate KEGG orthologues [50,512]. A web interface was developed with various tools. The search platform Lucene/Solr [52] was used to implement natural language search. Protein sequences provided by the user can be used for BLAST [50] searches against the coral proteome. Selected sequences of the coral proteome can be analysed with third party software (e.g. [53]) to interrogate conserved domains. ZoophyteBase is deployed using Apache-Tomcat (version 7.0.28 for Linux ×64 [513]) on the Ubuntu Linux server of the Section of Bioinformatics at the Faculty of Food Technology and Biotechnology, University of Zagreb, Croatia and is accessible at our published web address [47].

### Additional file

**Additional file 1: Table S16b.** Predicted (UniProt) homologues of animal toxins encoded in the genome of *A. digitifera*.

### Competing interests

The authors declare no competing interests exist.

### Authors’ contributions

WCD and PFL conceived the study, and participated in its design, coordination and drafted the manuscript. AS carried out the annotation. DB, JD, JZ and RG participated in the database design and testing. WCD, MJvO, AS and PFL performed data analysis. DH and JC participated in and coordinated the annotation, database design and testing. All authors have read and approved the final manuscript.

### Acknowledgments

The authors are indebted to Professor Noriyuki Satoh and his research team at the Marine Genomics Unit, Okinawa Institute of Science and Technology, Japan for making the protein sequence of *A. digitifera* publically available without restriction of use. We thank the following scientists from the Australian Institute of Marine Science for helpful discussions, Dr David Roume, Dr Andrew Negri, Dr Elitra Wood-Charlton and Dr Karen Weyenberg. We are also grateful to Dr Bill Leggat of the ARC Centre of Excellence for Coral Reef Studies and School of Pharmacy and Molecular Sciences, James Cook University, Townsville, Australia for critically reviewing this manuscript. Financial support for this work has come from the Biotechnology and Biological Sciences Research Council of the United Kingdom (BB/SRC grant BB/H01000/2 to WCD and PFL); from the Australian Institute of Marine Science (WCD and MJvO); from a cooperation grant of the German Academic Exchange Service (DAAD) and Ministry of Science, Education and Sports, Republic of Croatia (to JC and DH) and from the Croatian Science Foundation (grant 09/5 to DH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Initial work (in progress) was presented at the 12th International Coral Reef Symposium, Cairns, Queensland, Australia, 9-13 July 2012.

### Author details

1. Centre for Marine Microbiology and Genetics, Australian Institute of Marine Science, PAB No. 3 Townsville MC, Townsville 4810, Queensland, Australia.
2. Institute of Pharmaceutical Science, King’s College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, United Kingdom.
3. Department of Chemistry King’s College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, United Kingdom.
4. Section for Bioinformatics, Department of Biochemical Engineering, Faculty of Food Technology and Biotechnology, University of Zagreb, Pierottijeva 6, 1000 Zagreb, Croatia.
5. Department of Genetics, University of Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany.

Received: 6 March 2013 Accepted: 15 July 2013
Published: 26 July 2013

### References

1. Freudenthal HD: *Symbiodinium* gen. nov. and *Symbiodinium microadriaticum* sp. nov., a zooxanthella: taxonomy, life cycle, and morphology. *J. Eukaryot Microbiol* 1962, 9:45–52.
2. Muscatine L: The role of symbiotic algae in carbon and energy flux in reef corals. In Coral Reefs: Ecosystems of the World. Edited by Dubinsky Z. Amsterdam: Elsevier; 1990:75–84.
3. Younglees D, Rees TAV, Leggat W: Metabolic interactions between algal symbionts and invertebrate hosts. *Plant Cell Environ* 2008, 31:679–694.
4. Johannes RE, Wiebe WJ, Crossland CJ, Rimmer DW, Smith SV: Latitudinal limits of coral reef growth. *Mar Ecol Prog Ser* 1983, 11:105–111.
5. Spalding MD, Gereffet AM: New estimates of global and regional coral reef areas. *Coral Reefs* 1997, 16:225–230.
6. Hatcher BG: Coral reef primary productivity. A hierarchy of pattern and process. *Trends Ecol Evol* 1990, 5:149–155.
7. Reaka-Kudla ML: The global biodiversity of coral reefs: A comparison with rainforests. In *Biodiversity II: Understanding and Protecting Our Natural Resources*. Edited by Reaka-Kudla ML, Wilson DE, Wilson EO. Washington DC: Joseph Henry/National Academy Press; 1997:83–108.
8. Davy SK, Allemand D, Weis VM: Cell biology of cnidarians-dinoflagellate symbiosis. *Microbiol Mol Biol Rev* 2012, 76:229–261.
9. Bellwood DR, Hughes TP, Folke C, Nyström M: Confronting the coral reef crisis. *Nature* 2004, 429:827–833.
10. Hoegh-Guldberg O: Climate change, coral bleaching and the future of the world’s coral reefs. *Mar Freshw Res* 1999, 50:859–866.
11. Hoegh-Guldberg O, Mumby PJ, Hooten AJ, Steneck RS, Greenfield P, Gomez E, Harvell CD, Sale PF, Edwards AJ, Caldeira K, Knowlton N, Eakin CM, Jorgez-Pietro R, Muthiga N, Bradbury RH, Dubi A, Hatzios ES: Coral reefs under rapid climate change and ocean acidification. *Science* 2007, 318:1737–1742.
12. Porter JV, Dustan P, Jaap WC, Patterson KL, Kosmynin V, Meier OW, Patterson ME, Parsons M: Patterns of spread of coral disease in the Florida Keys. *Hydrobiologia* 2001, 460:1–24.
17. Orr JC, Selig ER, Casey KS, Page CA, Willis BL, Harvell CD, Sweatman H, Melendy AM: Thermal stress and coral cover as drivers of coral disease outbreaks. *Plos Biol* 2007, 5:e124.

18. Brandt ME, Manunco JW: Disease incidence is related to bleaching extent in reef-building corals. *Ecology* 2009, 90:2859–2867.

19. Devlin MJ, Brodie J: Terrestrial discharge into the Great Barrier Reef Lagoon: nutrient behaviour in coastal waters. *Mar Pollut Bull* 2005, 51:9–22.

20. Fabricius KE: Effects of terrestrial runoff on the ecology of corals and coral reefs: review and synthesis. *Mar Pollut Bull* 2005, 50:125–146.

21. Orr JC, Fabry VJ, Aumont O, Bopp L, Doney SC, Feely RA, Gnanadesikan A, Anthony KRN, Kline DI, Diaz-Pulido G, Dove S, Hoegh-Guldberg O: Effects of terrestrial runoff on the ecology of corals and coral reefs: review and synthesis.*Mar Pollut Bull* 2005, 50:125–146.

22. Devlin MJ, Brodie J: Terrestrial discharge into the Great Barrier Reef Lagoon: nutrient behaviour in coastal waters. *Mar Pollut Bull* 2005, 51:9–22.

23. Carpenter KE, Abram M, Aeby G, Aronson RB, Banks S, Bruckner A, Chiriboga A, Cortés J, Dellbeek JC, Devantier L, Edgar GJ, Edwards AJ, Fenster D, Guzmán HM, Hobækera BW, Hodgson G, Johan O, Lecuanan WY, Livingstone SR, Lovell ER, Moore JA, Obura DO, Ochavillo D, Pandolfi BA, Pechet WF, Quilliman MC, Rebcon B, Richards ZT, Rogers AD, Sanciangco J, et al: One third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science* 2008, 321:560–563.

24. Dunlop et al. BMC Genomics http://www.biomedcentral.com/1471-2164/14/509

25. The susceptibility and resilience of corals to thermal stress: a global map of human impact on marine ecosystems. *Science* 2008, 319:948–952.

26. Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JB, Kleya J, Lough JG, Marshall P, Nystrom M, Palumbi SR, Pandolfi JM, Rosen B, Roughgarden J: Climate change, human impacts, and the resilience of coral reefs. *Science* 2003, 301:892–933.

27. Evans TG, Hofmann GE: Defining the limits of physiological plasticity: how gene expression can assess and predict the consequences of ocean change. *Philos Trans R Soc B Biol Sci* 2012, 367:1733–1745.

28. Weis VM: The susceptibility and resilience of corals to thermal stress: adaptation, acclimatization or both? *Trends Ecol Evol* 2010, 25:1518–1517.

29. Edge SE, Morgan MB, Gleason DF, Snell TW: Gene expression in the early development of the scleractinian coral *Acropora millepora*. *Mol Ecol* 2009, 18:1517–1533.

30. DeSalle RM, Sun J, Chen Q, Lun JC, Xu J, Qiu JW: *PcamBase*: Development of a transcriptomic database for the brain coral *Pocillopora damicornis*. *BMC Genomics* 2011, 12:385.

31. Dunlop et al. BMC Genomics http://www.biomedcentral.com/1471-2164/14/509

32. Cardinale BJ, Duk-Rodriguez E, Gitay H, Kapos V, Kowarik I, Jager T, et al: Ecosystem biological diversity: toward a common framework for measurement and monitoring.*Biol Conserv* 2003, 112:141–156.

33. Cardinale BJ, Duk-Rodriguez E, Gitay H, Kapos V, Kowarik I, Jager T, et al: Ecosystem biological diversity: toward a common framework for measurement and monitoring. *Biol Conserv* 2003, 112:141–156.
67. Vassilieva EV, Nusrat A: Vesicular trafficking: molecular tools and targets. 
68. Kristie TM, Liang Y, Vogel JL: Rab and ARF GTPase regulation of exocytosis. 
69. Hong M-C, Huang Y-S, Lin W-W, Fang L-S, Chen M-C: Host-zooxanthellae interactions in four 
70. Kloepper TH, Kienle CN, Fasshauer D: SNARE Database. 
71. Chen MC, Hong MC, Huang YS, Liu MC, Cheng YM, Fang LS: Membrane fusion by synthetic endosomes. 
Methods Mol Biol 2010, 1799:127–138. 
72. Masuda ES, Huang BC, Fisher JM, Luo Y, Scheller RH: Syntaxin-binding protein that may regulate SNARE complex assembly. 
J Biol Chem 2003, 278:10427–10436. 
73. Chen MC, Cheng YM, Sung PJ, Kuo CE, Fang LS: Membrane fusion proteins sheds light on the conservation of the eukaryotic SNARE machinery. 
Mol Biol Cell 2007, 18:3463–3471. 
74. Hay JC, Chao DS, Kuo CS, Scheller RH: Protein interactions regulating vesicle transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells. 
Cell 1997, 89:149–158. 
75. Masuda ES, Huang BC, Fisher JM, Luo Y, Scheller RH: Tomosyn binds t-SNARE proteins via a VAMP-like coiled-coil. 
Neuron 1998, 21:679–680. 
76. Scales SJ, Hesser BA, Masuda ES, Scheller RH: Amisyn, a novel syntaxin-binding protein that may regulate SNARE complex assembly. 
J Biol Chem 2002, 277:28271–28279. 
77. Jahn R, Scheller RH: SNAREs – Engines for membrane fusion. Nat Rev Mol Cell Biol 2006, 7:531–643. 
78. Muscatine L, Pool RR, Crenichia E: Some factors influencing selective release of solubil organic material by zooxanthellae from reef corals. 
Mar Biol 1972, 13:298–308. 
79. Sutton DC, Hoegh-Guldberg O: Host-zooxanthellae interactions in four temperate marine invertebrate symbioses: assessment of the effect of host extracts on symbionts. 
Biol Bull 1990, 178:175–186. 
80. Lee S, Horn EL, Liu Y, Wysocka J, Bowerman B, Hengartner MO, Heer W: Epigenetic regulation of histone H3 serine 10 phosphorylation status by HCF-1 proteins in C. elegans and mammalian cells. 
PLoS One 2007, 2:e1213. 
81. Kristie TM, Liang Y, Vogel JL: Control of α-herpesvirus IE gene expression by HCF-1 coupled chromatin modification activities. 
Biochem Biophys Acta 2010, 1799:257–265. 
82. Christensen RG, Enauameh MS, Noyes MB, Bodsky MH, Wolfe SA, Stormo GD: Recognition models to predict DNA-binding specificities of homeodomain proteins. 
Bioinformatics 2012, 28:884–889. 
83. Mann RS, Lelli KM, Joshi R: Hox specificity: unique roles for cofactors and collaborators. 
Curr Top Dev Biol 2009, 88:56–101. 
84. Finnerty JR, Martindale MQ: Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. 
Dev Evol 1999, 11:26–23. 
85. Nishiwaki T, Satomi Y, Nakajima M, Lee C, Kiyohara R, Kageyama H, Kitayama JR: The cnidarian-bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, Nematostella vectensis. 
Curr Biol 2006, 16:7964. 
86. Saito T, Sawamoto K, Okado H, Anderson DJ, Mikoshiba K: Mammalian Barh homologue is a potential regulator of neural bHLH genes. 
Dev Biol 1998, 216:92–225. 
87. Habas S, Kato Y, He X: Wnt/Frizzled activation of Rho regulates vertebrate gastrulation and requires a novel Formin homology protein Daam1. 
Cell 2001, 107:843–854. 
88. Matussek T, Djiane A, Jankovic F, Brunner D, Mlodzik M, Mihaly J: The Drosophila formin Daam regulates the tracheal cuticle pattern through organizing the actin cytoskeleton. 
Development 2006, 133:567–576. 
89. Lui D, Halkett MA, Zhu W, Rubart M, Liu Y, Yang Z, Chen H, Haneline LS, Chan RJ, Schwartz RJ, Field LJ, Atkinson SJ, Shou W: Dishevelled-associated activator of morphogenesis 1 (Daam1) is required for heart morphogenesis. 
Development 2011, 138:303–315. 
90. Wada Y, Kitamoto K, Kanbe T, Tanaka K, Anraku Y: The SLP7 gene of Saccharomyces cerevisiae is essential for vesicular trafficking and morphogenesis function. 
Mol Cell Biol 1990, 10:2214–2223. 
91. Bragdon B, Moseychuk O, Saidahna S, King D, Julian J, Nohe A: Bone morphogenetic proteins: a critical review. 
Cell Signal 2011, 23:609–620. 
92. Martin VJ: Photoreceptors of cnidarians. Can J Zool 2002, 80:1703–1722. 
93. Mason BM, Cohen JH: Long-wavelength photoreactivity in coral planula larvae. 
Biol Bull 2012, 222:89–92. 
94. Levy OA, Appelbaum L, Leggat W, Gottlieb Y, Hayward DC, Miller DJ, Hoegh-Guldberg O: Light-responsive cryptophytes from a simple multicellular animal, the coral Acropora millepora. 
Science 2007, 318:467–470. 
95. Vize PD: Transcriptome analysis of the circular regulatory network in the coral Acropora millepora. 
Biol Bull 2009, 216:131–137. 
96. Brady AK, Snyder KA, Vize PD: Circadian cycles of gene expression in the coral, Acropora millepora. 
PLoS One 2011, 6:e25072. 
97. Ishiura M, Kutsuna S, Aoki I, Iwasaki H, Andersson CR, Tanabe A, Golden SS, Johnson CH, Kondo T: Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. 
Science 1998, 281:1519–1523. 
98. Nishikawa T, Satomi Y, Nakajima M, Lee C, Kiyohara R, Kageyama H, Kitayama Y, Temamoto M, Yamaguchi A, Hijikata A, Go M, Iwasaki H, Takao T, Kondo T: Role of KaiC phosphorylation in the circular clock system of Synechococcus elongatus PCC 7942. 
Proc Natl Acad Sci USA 2004, 101:3927–3932. 
99. Xu Y, Mori T, Qin X, Yan H, Egli M, Johnson CH: Intracellular regulation of phosphorylation status of the circular clock protein KaiC. 
PLoS One 2009, 4:e7509. 
100. Nakahira Y, Katayama M, Miyashita S, Kutsuna S, Iwasaki H, Oyama T, Kondo T: Global gene repression by KaiC as a master process of prokaryotic circadian system. 
Proc Natl Acad Sci USA 2004, 101:881–885. 
101. Kondo T, Ishiura M: The circadian clock of cyanobacteria. Bioessays 2000, 22:10–15.
107. Hardin PE, Hall JC, Roisbaib M: Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature 1990, 343:536–540.

108. Dunlap JC: Molecular bases for circadian clocks. Cell 1999, 96:271–290.

109. Reiter AM, Behrendt L, Tarrant AM: Light entrained rhythmic gene expression in the sea anemone Nematostella vectensis: the evolution of the animal circadian clock. PLoS One 2010, 5:e12825.

110. Gibbons IR, Cilia and flagella of eukaryotes. J Cell Biol 1981, 91:1075–124.

111. Linck RW: Tektins and microtubules. Adv Mol Biol Cell 1990, 3:353–63.

112. Fox LA, Sawin KE, Sale WS: Kinesin-related proteins in eukaryotic flagella. J Cell Sci 1994, 107:1455–1500.

113. Amos LA: The tektin family of microtubule-stabilizing proteins. Genome Biol 2008, 9:229.

114. Kozmurzki KG, Beech PL, Rosenbaum JL: Origins of neurogenesis, a cnidarian view. J Exp Zool B Mol Neurosci 2002, 289:536–540.

115. Kozminski KG, Beech PL, Rosenbaum JL: Evolution of two-component signal transduction systems. J Cell Sci 1997, 110:157–163.

116. Bartlett DH, Franz BI, Matsumura P: Enzymes of the shikimate acid pathway encoded in the genome of a marine invertebrate. J Biol Chem 2000, 275:20439–20446.

117. Ferris HU, Furukawa Y, Minamino T, Kroetz MB, Kihara M, Namba K, Macnab RM: The Aer protein and the serine chemoreceptor Tsr have microbial origins. Proc Natl Acad Sci USA 1999, 96:11919–11917.

118. Shichida Y, Matsuyama T: Evolution of opsins and phototransduction. Philos Trans R Soc Lond B Biol Sci 2009, 364:288–2895.

119. Belaev A, Leemrth DA, Saavo-de-la-Sala P: Synthesis and biological evaluation of novel, peripherally selective chromanol imidazoline-thione-based inhibitors of dopamine beta-hydroxylase. J Med Chem 2006, 49:1191–1197.

120. Grimmelikhuijzen CJP, Williamson M, Hansen GN: Neuropeptides in cnidarians. Can J Zool 2002, 80:1690–1702.

121. Attenborough IA, Hayward DC, Kitahara MW, Miller DJ, Ball EE: A “neural” enzyme in nonbilaterian animals and algae: preneural origins for pectinid glycoside ο-amidating monooxygenase. Mol Biol Evol 2012, 29:3092–3109.

122. Allemand D, Tambutté S, Tambutté E, Citta D, Tambutté S: Molecular cloning and localisation of a PMCA P-type calcium ATPase from the cnidarian hydroid Obelia longissima. BMC Genomics 2013, 14:509.
209. Rubio LM, Ludden PW: Biosynthesis of the iron-molybdogenum cofactor of nitrogenase. Annu Rev Microbiol 2008, 62:93–111.

210. Fujita Y, Bauer CE: Reconstitution of the light-independent protocorallophyllide reductase from purified Bchl and BchN-BchB subunits. In vitro confirmation of nitrogenase features of a bacterioclarephyll biosynthesis enzyme. J Biol Chem 2000, 275:23851–23858.

211. Sarmu R, Barney BM, Hamilton TL, Jones A, Seefeld LC, Peters JW: Crystal structure of the L protein of Rhodobacter sphaeroides light-independent protocorallophyllide reductase with MgADP bound: a homologue of the nitrogenase Fe protein. Biochemistry 2006, 45:10034–10051.

212. Munak D, Nomata J, Etaba K, Miazogiuchi T, Shiba T, Taniaki H, Kurisu G, Fujita Y: X-ray crystal structure of the light-independent protocorallophyllide reductase. Nature 2010, 461:110–114.

213. Heyes DJ, Hunter CN: Making light work of enzyme catalysis: protocorallophyllide oxidoreductase. Trends Biochem Sci 2005, 30:642–649.

214. Dean DR, Bolin JT, Zheng L: Nitrogenase metalloclusters: structure, organisation, and synthesis. J Bacteriol 1993, 175:6757–6774.

215. Zheng L, Dean DR: Catalytic formation of a nitrogenase iron-sulfur cluster. J Biol Chem 1994, 269:18723–18726.

216. Shah VK, Stacey G, Brill WJ: Electron transport to nitrogenase. J Biol Chem 1983, 258:12064–12068.

217. Hoover TR, Santero E, Porter S, Kustu S: UV-induced DNA damage in coral plenulae. Proc Natl Acad Sci USA 1987, 84:5590–5593.

218. Muraki N, Nomata J, Ebata K, Mizoguchi T, Shiba T, Tamiaki H, Kurisu G, Kusumawidijaja G: Photoactivation and the coral bleaching phenomenon. J Exp Mar Biol Ecol 2009, 377:1–17.

219. Anderson S, Zepp R, Machada J, Santavy D, Hansen L, Mueller E: Indicators of UV exposure in corals and their relevance to global climate change and coral bleaching. Hum Ecol Risk Assess 2001, 7:1271–1278.

220. Torregiani JH, Lesser MP: The effects of short-term exposures to ultraviolet radiation in the Hawaiian coral Montipora verrucosa. J Exp Mar Biol Ecol 2007, 340:194–202.

221. Benich A, Avishai N, Rabinowitz C: UV incites diverse levels of DNA breaks in different cellular compartments of a branching coral species. J Exp Biol 2005, 208:843–848.

222. Holden HM, Thoden JB, Raushel FM: Making light work of enzyme catalysis: nitrogenase metalloclusters. Annu Rev Plant Biol 1999, 50:277–303.

223. Carpenter LW, Patterson MR, Bromage ES: Coral reefs 2012, 32:207–214.

224. Catmull J, Yellowlees D, Miller DJ: NADP+-dependent glutamate dehydrogenase from Acropora formosa: purification and properties. Mar Biol 1987, 95:559–563.

225. Clode P, Saunders M, Maker G, Ludwig M, Atkins CA: Carbamoyl phosphate synthetase: an amazing biochemical odyssey from substrate to product. Cell Mol Life Sci 1999, 56:507–522.

226. Crandall JB, Teece MA: An overview of the coral bleaching crisis. Proc Natl Acad Sci USA 2008, 105:2451–2458.

227. Dunlap et al. BMC Genomics 2013, 14:509

228. Carpenter LW, Patterson MR, Bromage ES: Water flow influences the spatiotemporal distribution of heat shock protein 70 within colonies of the scleractinian coral Montastrea annularis (Ellis and Solander, 1786) following heat stress: Implications for coral bleaching. J Exp Mar Biol Ecol 2010, 387:52–59.

229. Rosic NN, Pernice M, Dove S, Dunn S, Hoehn-Gulberg O: Gene expression profiles of cytosolic heat shock proteins Hsp70 and Hsp90 from symbiotic dinoflagellates in response to thermal stress: possible implications for coral bleaching. Cell Stress Chaperones 2011, 16:95–106.

230. Aravind L, Anantharaman V, Koonin EV: Monophyly of class I aminocyclitol RNA synthase, USPA, ETP, photolyase and the PP-ATPase nucleotide-binding domains: Implications for protein evolution in the RNA world. Proteins 2002, 48:1–14.
Molecular and evolutionary basis of the cellular stress response. 

272. Levy O, Kaniewska P, Alon S, Eisenberg E, Karako-Lampert S, Bay LK, Reef R, Kültz D: Molecular and evolutionary basis of the cellular stress response. *Plant Physiol.* 2003, 131:1290–1299.

257. Forêt S, Seneva F, de Jong D, Bieller A, Hemmrich G, Hayden DC, Ball EE, Bosch TC, Agata K, Hassel M, Miller DJ: Phylogenomics reveals an anomalous distribution of USP genes in metazoans. *Mol Biol Evol* 2011, 28:153–161.

259. Kerk D, Bulgrien J, Smith DW, Gribskov M: Differential gene expression during thermal stress and bleaching in the Caribbean coral *Montastraea faveolata*. *Mol Evol* 2008, 17:3952–3971.

260. Forêt S, Seneca F, de Jong D, Bieller A, Hemmrich G, Hayden DC, Ball EE, Bosch TC, Agata K, Hassel M, Miller DJ: Phylogenomics reveals an anomalous distribution of USP genes in metazoans. *Mol Biol Evol* 2011, 28:153–161.

265. Hernández MP, Sullivan WP, Toft DO: The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction.

267. Ellis RJ, van der Vies SM: The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction.

277. Murata S, Yashiroda H, Tanaka K: Gene expression in coral-algal symbiosis. *Annu Rev Physiol* 2010, 72:331–352.

278. Park BH: Hypoxia-inducible factor-2 (HIF-2) regulates erythropoietin gene expression in vivo.

286. Rao PV, Horvitz J, Zigler JS: δ-Crystallin, a molecular chaperone, forms a stable complex with carnitine arylhydrate upon heat denaturation. *Biochem Biophys Res Commun* 1993, 190:786–793.

287. Carricat-Garvet JP, Caballeras-Terán N, Cruz-Ortega J, Blanchon P: Sensitivity of calcification to thermal stress varies among genera of massive reef-building corals. *PloS One* 2012, 7:e32859.

289. Santo Ede O, Alves N Jr, Dias GM, Mazotto AM, Vermelho A, Vora GJ, Wilson B, Beltran VH, Bourne DG, Le Roux F, Thompson FL: Genomic and proteomic analysis of the pathogen *Vibrio corallivorus* reveals a diverse virulence repertoire. *ISME J* 2011, 5:1471–1483.

290. Kimes NE, Grim CJ, Johnson WR, Hasan NA, Tall BD, Kothary MH, Kiss H, Murak MC, Tapias R, Green L, Detter C, Brice DC, Berritt TS, Covell RR, Morris PJ: Temperature regulation of virulence factors in the pathogen *Vibrio corallivorus*. *ISME J* 2012, 6:835–846.

291. Zhang T, Kraus WL: SIRT1-dependent regulation of chromatin and transcription: linking NAD+ metabolism and signaling to the control of cellular functions. *Biochim Biophys Acta* 2010, 1804:1666–1675.

292. Liu TY, Zoya BR, El Gazzar M, Vachharajani VT, McCall CE: NAD+-dependent SIRT1 deacetylase participates in epigenetic reprogramming during endotolerance treatment. *J Biol Chem* 2011, 286:19856–19864.

293. Kataoka S, Imhof A, Sassone-Corsi P: Connecting threads: epigenetics and metabolism. *Cell* 2012, 148:24–28.

294. Lim JH, Lee YM, Chun YS, Chen J, Kim JE, Park JW: Sirtuin 1 modulates cellular responses to hypoxia by deacetylating hypoxia-inducible Factor α. *Mol Cell* 2010, 38:260–270.

295. Lee JH, Song MY, Song EK, Kim EK, Moon WS, Han MK, Park JW, Kwon KB, Park BR: Overexpression of SIRT1 protects pancreatic beta-cells against cytokine toxicity by suppressing the nuclear factor-kappaB signalling pathway. *Diabetes* 2009, 58:344–351.

296. Wang F, Nguyen M, Qin FX, Tong Q: SIRT2 deacetylates FOXO3a in response to oxidative stress and caloric restriction. *Aging Cell* 2007, 6:505–514.

297. Shi T, Wang F, Stieren E, Tong Q: SIRT3, a mitochondrial sirtuin deacetylase, regulates mitochondrial function and thermogenesis in brown adipocytes. *J Biol Chem* 2005, 280:13560–13567.

298. Nairn N, Wu X, Fortier E, Feng Y, Baré OQ, Chen S, Ren X, Wu Z, Streeper RS, Bordonne L: SIRT4 regulates fatty acid oxidation and mitochondrial gene expression in liver and muscle cells. *J Biol Chem* 2010, 285:31995–32002.

299. Du J, Zhou Y, Su X, Yu J, Khan S, Jiang H, Kim J, Woo J, Kim JH, Choi BH, He B, Chen W, Zhang S, Cerione RA, Auwerx J, Hao Q, Lin H: Sirt6 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* 2011, 334:806–809.

300. Mostoslavsky G, Franco S, Murphy MM, Kold J, Patel P, Hsu JT, Hong AL, Ford E, Cheng HL, Kennedy C, Nunez B, Bronson R, Friedewey D, Auerbach W, Valenzuela D, Karov M, Hottinger MO, Hursting S, Barrett JC, Guerente L, Mulligan R, Dernelle B, Yanopoulos GD, et al: Genomic instability and aging-like phenotype in the absence of mammalian SIRT4. *Cell* 2006, 124:315–329.

301. Luo J, Nikolakoy AV, Imai S, Chen D, Su F, Shiloh A, Guerente L, Gu W: Negative control of p53 by Sir2alpha promotes cell survival under stress. *Cell* 2001, 107:137–146.

302. Vakhrukhoeva O, Smolka C, Gajawada P, Kostin S, Boettger T, Kubin T, Braun T, Bober E: Sirt7 increases stress resistance of cardiomyocytes and prevent apoptosis and inflammatory cardiomyopathy in mice. *Circ Res* 2008, 102:703–710.

303. O’Halloran TV, Cullotta VC: Metallochaperones, an intracellular shuttle service for metal ions. *J Biol Chem* 2000, 275:25057–25060.
304. Lin SJ, Culotta VC. The Atx1 gene of Saccharomyces cerevisiae encodes a small metal homeostasis factor that protects cells against reactive oxygen toxicity. *Proc Natl Acad Sci U S A* 1995, 92:3784–3788.

305. Culotta VC, Klop LW, Strain J, Casareno RL, Krems B, Gitlin JD. The copper chaperone for superoxide dismutase. *J Biol Chem* 1997, 272:3469–3472.

306. Wong PC, Waggoner D, Subramaniam JR, Tresorillo L, Batinkas TR, Culotta VC, Price DL, Rothstein J, Gitlin JD. Copper chaperone for superoxide dismutase is essential to activate mammalian Cu/Zn superoxide dismutase. *Proc Natl Acad Sci USA* 2000, 97:2886–2891.

307. Shiek M, Lesser MP, Dunlap WC, Stochaj WR, Chatterjee B, Wu W, Zhong J. Depth-dependent responses to solar ultraviolet radiation and oxidative stress in the zooxanthellate coral Acropora microphthalma. *Mar Biol* 1995, 122:41–51.

308. Lesser MP, Stochaj WR, Tапley DW, Shiek JM. Bleaching in coral–reef anthozoans: effects of irradiance, ultraviolet radiation, and temperature on the activities of protective enzymes against active oxygen. *Coral Reefs* 1990, 8:225–232.

309. Crowe JH, Carpenter JP, Crowe LM. The role of vitrification in anthodisobiosis. *Ann Rev Physiol* 1980, 42:73–103.

310. Salá-Pauza MA, Meyer-Fernández J. Stabilization against thermal inactivation promoted by sugars on enzyme structure and function: Why is trehalose more effective than other sugars? *Arch Biochem Biophys* 1998, 360:10–14.

311. Veron JEN. *Corals of Australia and the Indo-Pacific*. Singapore: University of Hawaii Press; 1993:3644.

312. Shiek JM. Solar UV and oxidative stress in algal–animal symbioses. In: *Frontiers of Photobiology*. Edited by Shima A, Inagaki M, Fujikura Y. Takehbe H. Amsterdam: Elsevier Science Publishers; 1993:561–564.

313. Turrens JF. Mitochondrial formation of reactive oxygen species. *J Physiol* 2003, 552:335–344.

314. Lesser MP. Oxidative stress causes coral bleaching during exposure to elevated temperatures. *Coral Reefs* 1997, 16:187–192.

315. Brown BE, Downs CA, Dunn RP, Gibb SW. Oxidative stress in coral–photobiont communities and the role of mitochondria in apoptosis. *Coral Reefs* 2006, 25:610–614.

316. Hallwcll B. Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. *Plant Physiol* 2004, 139:312–322.

317. Mihatá-Cabello MA, Lesser MP, Flores-Ramírez LA, Zenteno-Savin T, Reyes-Borilla H. Oxidative stress in coral–photobiont communities. In: *Oxidative Stress in Aquatic Ecosystems*. Edited by Abele O, Vázquez-Medina JP, Zenteno-Savin T, Chichester: Wiley-Blackwell; 1993:561–564.

318. Jönsson TJ, Lowther WT. The peroxiredoxin repair proteins. *Subcell Biochem* 2007, 44:115–141.

319. Cheng Z, Zhang J, Balow DP, Williams CH Jr. Reactivity of thiodoxin as a protein thiol-disulfide oxidoreductase. *Chem Rev* 2011, 111:5768–5783.

320. Arné ESJ, Holmgren A. Physiological function of thioredoxin and thiorhodanide reductase. *Eur J Biochem* 2000, 267:6102–6108.

321. Ainsworth TD, Hoeh-Guldberg OD. Cellular processes of bleaching in the Mediterranean coral *Oculina patagonica*. *Coral Reefs* 2008, 27:593–597.

322. Kenkel CD, Aglyamova G, Alamrani J, Abagoli G, Capper R, Cunningham R, Richier S, Sabourault C, Courtiade J, Zucchini N, Allemand Furla P. Regulation of apoptotic mediators reveals dynamic responses to thermal stress in the reef building coral Acropora millepora. *Proc Natl Acad Sci U S A* 2011, 108:9095–9099.

323. Tchernov D, Zhang Q, Ye Y, Godzik A. Surprising complexity of the ancestral apoptotic network. *Genome Biol* 2007, 8:B22.

324. Schiep S, Wittig K, Stiening B, Böttger A, Leitz T. Morphogenesis in *Hydractinia echinata* (Cnidaria) is caspase-dependent. *Int J Dev Biol* 2006, 50:625–70.

325. Dunn SR, Weiss VM. Apoptosis as a post-phagocytic mechanism in a coral dinoflagellate mutualism. *Environ Microbiol* 2009, 11:268–276.

326. Adrían C, Brunetti G, Martin SJ. Apoptosomes: protease activation platforms to die for. *Trends Biochem Sci* 2005, 31:243–247.

327. Yang X, Khosravi-Far R, Chang HY, Baltimore D. Daxx, a novel Fas-binding protein that activates JNK and apoptosis. *Cell* 1997, 89:1067–1076.

328. Cohen O, Inbal B, Kissil JL, Raveh T, Berissi H, Spivak-Kozianzam T, Feinest E, Kimchi A. DAPK-kinase participates in TNF-α- and Fas-induced apoptosis and its function requires the death domain. *J Cell Biol* 1999, 146:141–144.

329. Riggs B, Chu T. Programmed cell death 6 (PCDE) protein interacts with death associated protein kinase 1 (DAPK1): additive effect on apoptosis via caspase-3 dependent pathway. *Biotechnol Lett* 2005, 27:1011–1015.

330. Kaiser WK, Upton JW, Long AB, Livingstone-Rosanoff D, Daley-Bauer LP, Haslem R, Caspary T, Mocanu ES. RIP3 mediates the embryonic lethality of caspase-8 deficient mice. *Nature* 2003, 421:368–372.

331. Krumschnabel G, Sohm B, Bock F, Mandl C, Villunger A. The enigma of caspase-2: the layman’s view. *Cell Death Differ* 2009, 16:195–207.

332. Krochid A, Ikura M. How calpain is activated by calcium. *Nat Struct Biol* 2002, 9:239–341.

333. Janssens S, Timel A, Lippens T, Tsuchop J. PIDD mediates NF-κB activation response to DNA damage. *Cell* 2005, 123:1079–1092.

334. Wolenski FS, Garbati MR, Lubinski TL, Taylor-Knowles N, Dresselhaus E, Stefanik DJ, Goucher H, Finnerty JR, Gilmore TD. Characterisation of the core elements of the NF-κB signalling pathway of the sea anemone *Nematostella vectensis*. *Mar Cell Biol* 2011, 31:1076–1087.

335. Dunlap SR, Phillips WS, Sapatara JW, Green DR, Wies VM. Highly conserved caspase-8 homologues from the sea anemone *Aiptasia pallida*: lower metazoans for the study of apoptosis evolution. *J Mol Evol* 2006, 63:95–107.

336. Knowlton N, Rohwer F. Multiples microbial mutualism on coral reefs: the host as a habitat. *Am Nat* 2003, 162:Suppl:S51–S52.

337. Littman RA, Willis BL, Pfeffer C, Bourne DG. Cysteine-rich protein family in reef-building corals. Identification and gene expression analysis of a taxonomically restricted *Coral reefs*: corals expanded caspase family and pro- and anti-apoptotic Bcl-2 proteins. *J Mol Evol* 2006, 63:95–107.

338. Knowlton N, Rohwer F. Diversity of coral-associated bacteria differ with location, but not species, for three acroporid corals on the Great Barrier Reef. *FEMS Microbiol Ecol* 2009, 68:152–153.

339. Rohwer F, Segurian V, Azam F, Knowlton N. Diversity and distribution of coral-associated bacteria. *Mar Ecol Prog Ser* 2002, 243:1–10.

340. Gojn J, DeSalvo MC, Voolstra CR, Reyes-Bermúdez A, Medina M. Identification and gene expression analysis of a taxonomically restricted cysteine-rich protein family in reef-building corals. *Proc Natl Acad Sci U S A* 2006, 103:12115–12120.

341. Sonagawa S, del Salvo MC, Voolstra CR, Reyes-Bermúdez A, Medina M. Identification and gene expression analysis of a taxonomically restricted cysteine-rich protein family in reef-building corals. *Proc Natl Acad Sci U S A* 2006, 103:12115–12120.

342. Rosenberg E, Koren O, Reshef L, Efrony R, Zilber-Rosenberg I. The role of microorganisms in coral health, disease, and evolution. *Nat Rev Microbiol* 2007, 5:355–362.

343. Sousa PR, Adl SP, D’Sa AH, Guerra CA, Troy K, Barasingha N. The role of microorganisms in coral health, disease, and evolution. *Nat Rev Microbiol* 2007, 5:355–362.

344.崙se MG, 崙se MG. Corals shed bacteria as a potential mechanism of resistance to organic matter enrichment. *ISME J* 2012, 6:1159–1165.

345. Boume DG, 崙se MG, Work TM, 崙se MG, Smith GW, Harvell CD. Microbial disease and the coral holobiont. *Trends Microbiol* 2009, 17:554–562.
354. Reshef L, Koren Q, Loy A, Zilber-Rosenberg J, Rosenberg E: The coral probiotic hypothesis. Environ Microb 2006, 8:2068–2073.

355. Vollmer SV, Kline DI: Natural disease resistance in threatened staghorn corals. PLoS One 2008, 3:e3778.

356. Carroll SP, Hendry AP, Reznick DN, Fox CW: Evolution on ecological time scales. Funct Ecol 2007, 21:287–292.

357. van Oppen MHJ, Souper P, Howells EJ, Hayward A, Berkelmans R: Novel genetic diversity through somatic mutations: fuel for adaptation of reef corals? Diversity 2011, 3:405–423.

358. Mydlarz LD, McGinty ES, Harvell CD: What are the physiological and immunological responses of coral to climate warming and disease? J Exp Biol 2010, 213:934–945.

359. Hamada M, Shoguchi E, Shinzato C, Kawashima T, Miller DJ, Satoh N: The complex NOD-like receptor repertoire of the coral Acropora digitifera includes novel domain combinations. Mol Biol Evol 2013, 30:167–176.

360. Boyes DC, Nam J, Dangl JL: The Arabidopsis thaliana RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. Proc Natl Acad Sci U S A 1998, 95:15849–15854.

361. Axtell MJ, Stakawicz BJ: Pathogenesis-related proteins: research progress in the last 15 years. Trends Genet 2011, 27:13–24.

362. Edreva A: The versatile bacterial type IV secretion systems. Cell Microbiol 2008, 10:2377–2386.

363. Hamada M, Shoguchi E, Shinzato C, Kawashima T, Miller DJ, Satoh N: The complex NOD-like receptor repertoire of the coral Acropora digitifera includes novel domain combinations. Mol Biol Evol 2013, 30:167–176.

364. Mydlarz LD, Couch SJ, Well E, Smith G, Harvell CD: Immune defenses of healthy, bleached and diseased Montastrea faveolata during a natural bleaching event. Dis Aquat Organ 2009, 87:567–577.

365. Palmer C, Bythell JC, Willis BL: A comparative study of phenoloxidase activity in diseased and bleached colonies of the coral Acropora millepora, Dis. Aquat Org. 2011, 11:3549–3581.

366. Cascales E, Christie PJ: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

367. Perrone F, Sestini P: Intestinal immune responses in humans: The path less explored: innate immune reactions in cnidarians. In: The path less explored: innate immune reactions in cnidarians. In: Trends Genet 2005, 21:633–639.

368. Khalturin K, Hemmich G, Fraune S, Augustin R, Bosch TC: Neglected biological features in cnidarians self-nonself recognition. Adv Exp Med Biol 2009, 641:109–116.

369. Kuznetsov SG, Bosch TC: Self/nonself recognition in Cnidaria: contact of allergic tissue does not result in elimination of nonself cells in Hydra vulgaris. Zool Jena 2003, 106:109–116.

370. Rinkevich B: Allorecognition and xenorecognition in reef corals: a decade of interactions. Hydrobiologia 2004, 530–531:443–450.

371. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

372. Patten NL, Seymour JR, Mitchell JG: Flow cytometric analysis of virus-like particles and heterotrophic bacteria within coral-associated reef water. J Mar Biol Assoc UK 2006, 86:563–566.

373. Lenuste A, Bourvier T, Bettarel Y: Enumerating viruses in coral mucus. Appl Environ Microbiol 2012, 78:6377–6379.

374. Patten NL, Harrison PL, Mitchell JG: Prevalence of virus-like particles within the staghorn scleractinian coral (Acropora muriocata) from the Great Barrier Reef. Coral Reefs 2008, 27:569–580.

375. Wilson WH, Francis J, Ryan K, Davy SK: Temperature induction of viruses in symbiotic dinoflagellates. Aquat Microb Ecol 2001, 25:59–102.

376. Pakhomov EA, Kinsey MC, Char Al, Ainsworth RJ: Metagenomics analysis indicates that stressors induce production of herper-like viruses in the coral Porites compressa. Proc Natl Acad Sci U S A 2008, 105:18413–18418.

377. Patten NL, Seymour JR, Mitchell JG: Flow cytometric analysis of virus-like particles in reef corals. Coral Reefs 2005, 24:141–145.

378. Seymour JR, Patten NL, Bourne DG, Mitchell JG: Novel nucleopolyanomalous dsDNA and +ssRNA viruses are associated with dinoflagellate endosymbionts of corals. EMSE J 2013, 7:13–27.

379. Rinkevich B: Allorecognition and xenorecognition in reef corals: a decade of interactions. Hydrobiologia 2004, 530–531:443–450.

380. Rinkevich B: Neglected biological features in cnidarians self-nonself recognition. Adv Exp Med Biol 2012, 738:86–93.

381. Rinkevich B: Allorecognition and xenorecognition in reef corals: a decade of interactions. Hydrobiologia 2004, 530–531:443–450.

382. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

383. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

384. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

385. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

386. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

387. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

388. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

389. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

390. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

391. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

392. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

393. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

394. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

395. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

396. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

397. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

398. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

399. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.

400. Rinkevich B: The evolution of immunity: a low-life perspective. Trends Immunol 2007, 28:449–454.
403. Widder WR, Wicker RB: Evidence that the SIV viral system of Sarcocccomyses cervinea uses actin by blocking expression of viral mRNA. Mol Cell Biol 1993, 13:4331–4341.

404. Brown JT, Bai X, Johnson AW: The yeast antiviral proteins Skp2p, Skp3p and Skl8p exist as a complex in vivo. RNA 2000, 6:499–457.

405. Houseley J, Tollervey D: The many pathways of RNA degradation. Cell 2009, 136:763–776.

406. Arraiano CM, Matos RG, Barbosa A: RNAse II: The finer details of the modus operandi of a molecular killer. RNA Biol 2010, 7:276–281.

407. Samuel CE: Antiviral actions of interferon. Interferon-regulated proteins and their surprisingly selective antiviral activities. Virology 1991, 183:1–11.

408. Haller O, Kochs G, Webe F: The interferon response circuit: Induction and suppression by pathogenic viruses. Virology 2006, 344:119–130.

409. Fensterl V, Sen GC: Interferons and viral infections. BioFactors 2009, 35:14–20.

410. Taniguchi T, Ogasawara K, Takaoka A, Tanaka N: IFN-alpha/IFN-beta upregulates ISG15 to restrict hepatitis C virus. J Virol 2013, 87:5016–5025.

411. Spear PG, Longnecker R: Interferon consensus sequence binding protein is a negative regulator of NF-kappaB and IRF3. J Virol 2004, 78:10769–10776.

412. Salkowski CA, Barber SA, Detore GR, Vogel SN: RNA, defines an alternative antiviral pathway against RNA genomic RNA. ISG20, a new interferon-induced RNase specific for single-stranded RNA, interacts with RNA polymerase II. RNA Biol 2013, 10:147–159.

413. Brown JT, Bai X, Johnson AW: Identification of a novel GTPase, the inducibly expressed Ski8p exist as a complex

414. Dolman D, Ebert M, Wallinga M, Shimizu H, Ramsay E, Hupp TR, Ball KL: Interferon regulatory factor 2 binding to p300 stimulates DNA-dependent acetylation of p53. Mol Cell Biol 2004, 24:10083–10098.

415. Salkowski CA, Barber SA, Detore GR, Vogel SN: Identification of a novel GTPase, the inducibly expressed Ski8p exist as a complex

416. Beerens J, Tollervey D: The many pathways of RNA degradation. Cell 2009, 136:763–776.

417. Pichlmair A, Schulz O, Tan CP, Rehwinkel J, Kato H, Takeuchi O, Akira S, Tschopp J, Dunlap WC, Morandini AC, Marques AC, Moura-da-Silva AM, Ward M, Padilla G, da Silva LF, Andreakis N, Long PF: Proteomic characterisation of toxins isolated from nematocysts of the South Atlantic jellyfish Olindias sambucina. Toxicon 2013, 62:1400–1407.

418. Leplaga SH: Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. Proc Natl Acad Sci USA 1982, 79:3162–3166.

419. Tang WJ, Guo G: The andrenyl cyclase activity of anthrax edema factor. Mol Asp Med 2009, 30:423–436.

420. Van Deldem C, Iglesias BH: Cell-to-cell signalling and Pseudomonas aeruginosa infections. Emerg Infect Dis 1998, 4:651–656.

421. Yates SP, Merril AR: Elucidation of eukaryotic elongation factor-2 contact sites within the catalytic domain of Pseudomonas aeruginosa exotoxin A. Biochem J 2004, 379:563–572.

422. Hinnebusch BJ, Rudolph AE, Cherepanov P, Dixon JE, Schwan TG, Fonberg J: Role of Yersinia murine toxin in survival of Yersinia pestis in the midgut of the flea vector. Science 2002, 296:733–735.

423. Fagerlund A, Lindbäck T, Storset AK, Granum PE, Hardy SP: Bacillus cereus Nhe is a pore-forming toxin with structural and functional properties similar to the ClyA (HlyE, SheA) family of haemolysins, able to induce osmotic lysis in epithelial. Microbiology 2008, 154:693–704.

424. Marrack P, Kappler J: The staphylococcal enterotoxins and their relatives. Science 1990, 248:705–711.

425. Nemerow GR, Wolfert R, McNaughton ME, Cooper NR: The multifunctional NS1 protein of influenza A viruses. J Gen Virol 2008, 89:2359–2376.

426. Koppert J, Marnay N, Mathew H, Oduko A: The role of interaction factor in gene expression in Escherichia coli. Mol Microbiol 1992, 6:525–5263.

427. Anders B, Moxe P: Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actiniaria). Toxicon 2002, 40:111–114

428. Taniguchi T, Ogasawara K, Takaoka A, Tanaka N: IFN-alpha/IFN-beta upregulates ISG15 to restrict hepatitis C virus. J Virol 2013, 87:5016–5025.

429. Miller VL, Taylor RK, Mekalinos J: Cholera toxin transcriptional activator toxT is a transmembrane DNA binding protein. Cell 1987, 48:271–279.

430. Childers BH, Rose KE: Regulation of virulence in Vibriochloroce: The ToxR regulon. Future Microbiol 2007, 2:335–344.

431. Lanka E, Wilkins BM: DNA processing reactions in bacterial conjugation. Annu Rev Biochem 1995, 64:141–169.

432. Iredell JR, Mannan PA: The toxin-coregulated pilus of Vibrio cholerae 01: a model for type A pilus biosynthesis. Trends Microbiol 1994, 2:187–192.

433. Miller VL, Taylor RK, Mekalinos J: Cholera toxin transcriptional activator toxT is a transmembrane DNA binding protein. Cell 1987, 48:271–279.

434. French CT, Panina EM, Yeh SH, Griffith N, Arambula DG, Miller JF: The Bordetella type III secretion system effectoR BltE contains a conserved
N-terminal motif that guides bacterial virulence factors to lipid raft. Cell Microbiol 2009, 11:735–1749.

455. Espinosa A, Alfaro JR: Disabling surveillance: bacterial type III secretion effectors that suppress innate immunity. Cell Microbiol 2004, 6:1027–1040.

456. Deane JE, Roversi P, Cordes FS, Johnson S, Kenjali R, Daniell S, Booy F, Picking WD, Picking WL, Blocker AJ, Lea SM: Molecular model of a type III secretion system needle: implication for host cell sensing. Proc Natl Acad Sci U S A 2006, 103:12529–12533.

457. Jungo F, Bouguerel L, Xenarios I, Poux S. The UniProtKB/Swiss-Prot Tox-Prot program: a central hub of integrated venom protein data. Toxicon 2012, 60:551–557.

458. Chang D, Duda TF Jr: Extensive and continuous duplication facilitates rapid evolution and diversification of genome family. Mol Biol Evol 2012, 29:2019–2027.

459. Wong ESW, Belov K: Venom evolution through gene duplication. Genes 2012, 4:4961–7.

460. Moura-da-Silva AM, Furlan MS, Caparroni MC, Grego KR, Portes-Junior JA, Clissa PB, Valente RH, Magalhães GS: Diversity of metalloproteinases in Bothrops neuwiedi snake venom transcripts: evidences for recombination between different classes of SMMPs. BMC Genet 2011, 12:94.

461. Dutertre S, Jin AH, Kaas Q, Jones A, Alewood PF, Lewis RJ: Deep venomomics reveals the mechanism for expanded peptide diversity in cone snail venom. Mol Cell Proteomics 2012, 12:312–329.

462. Starcevic A, Long PF: Diversification of animal venom peptide – were jellyfish amongst the first combinatorial chemists? ChemBioChem 2013, 14:491–503.

463. Fabricius KE: Factors determining the resistance of coral reefs to eutrophication: A review and conceptual model. In Coral Reefs: An Ecosystem in Transition. Edited by Dubinsky Z, Stambler N. Heidelberg: Springer (Dordrecht); 2011:493–508.

464. Ferrier-Pages C, Gatusso J-P, Daulet S, Jaultert J: Effect of nutrient enrichment on growth and photosynthesis of the zooxanthellate coral Stylophora pistillata. Coral Reefs 2000, 19:103–113.

465. Kramarsky-Winter E, Downs CA, Downs A, Loya Y: Factors determining the resistance of coral reefs to environmental and engineering systems. Environ Sci Technol 2008, 42:3530–3530.

466. Inouye BW, Sato Y, Senechal M, Tompkins AL, Hidaka H, Hidaka S: Sublethal effects of mercury and its distribution in the fertilization success of several scleractinian coral species. Coral Reefs 2010, 29:524–556.

467. Victor S, Richmond RH: Effect of copper on fertilisation success in the reef coral Acropora seliculosa. Mar Pollut Bull 2005, 51:1448–1451.

468. Nefzi AP, Howard AJ: Inhibition of coral fertilisation and larval metamorphosis by tributyltin and copper. Mar Environ Res 2001, 51:17–27.

469. Reichelt-Brushett AJ, Harrison PL: The effect of selected trace metals on the fertilization success of several scleractinian coral species. Coral Reefs 2005, 24:524–524.

470. Rossini L, Christmas EM: The UniProtKB/Swiss-Prot: evidences for recombination between different classes of SMMPs. BMC Genet 2011, 12:502.

471. Bothrops neuwiedi snake venom transcripts: evidences for recombination between different classes of SMMPs. BMC Genet 2011, 12:94.

472. Purvis R, Rambaut A, Penny D: Extensive and continuous duplication facilitates rapid evolution and diversification of genome family. Mol Biol Evol 2004, 21:1259–1253.

473. Kramarsky-Winter E, Downs CA, Downs A, Loya Y: Factors determining the resistance of coral reefs to environmental and engineering systems. Environ Sci Technol 2008, 42:3530–3530.

474. Inouye BW, Sato Y, Senechal M, Tompkins AL, Hidaka H, Hidaka S: Sublethal effects of mercury and its distribution in the fertilization success of several scleractinian coral species. Coral Reefs 2010, 29:524–556.

475. Victor S, Richmond RH: Effect of copper on fertilisation success in the reef coral Acropora seliculosa. Mar Pollut Bull 2005, 51:1448–1451.

476. Nefzi AP, Howard AJ: Inhibition of coral fertilisation and larval metamorphosis by tributyltin and copper. Mar Environ Res 2001, 51:17–27.

477. Reichelt-Brushett AJ, Harrison PL: The effect of selected trace metals on the fertilization success of several scleractinian coral species. Coral Reefs 2005, 24:524–556.

478. Karntanut W, Pascoe D: Effects of the multiple stressors copper and reduced salinity on the metabolism of the hermatypic coral Porites lutea. Mar Environ Res 2001, 52:289–299.

479. Nyström M, Nordemar I, Tedengren M: Simultaneous and sequential stress from increased temperature and copper on the metabolism of the hermatypic coral Porites cylindrica. Mar Biol 2001, 138:1225–1311.

480. Mitchellmore D, Weis VM: Upshift and partitioning of copper and cadmium in the coral Ficusococcum damicornis. Aquat Toxicol 2007, 85:48–56.

481. Belmyer GA, Grosell M, Bhagooli R, Baker AC, Langdon C, Gillette P, Capo TR: Differential effects of copper on three species of scleractinian corals and their algal symbionts (Symbiodinium spp.). Aquat Toxicol 2010, 97:125–133.

482. Bastidas C, Garcia EM: Sublethal effects of mercury and its distribution in the coral Porites astreoides. Mar Ecol Prog Ser 2004, 267:133–143.

483. Faria N, Ramos R, Bastidas C, Garcia E: Biochemical response of cnidarian larvae to mercury and benzo[a]pyrene exposure. Bull Environ Contam Toxicol 2008, 81:553–557.

484. Negri AP, Howard AJ: Inhibition of coral fertilisation and larval metamorphosis by tributyltin and copper. Mar Environ Res 2001, 51:17–27.

485. Reichelt-Brushett AJ, Harrison PL: The effect of selected trace metals on the fertilization success of several scleractinian coral species. Coral Reefs 2005, 24:524–556.

486. Victor S, Richmond RH: Effect of copper on fertilisation success in the reef coral Acropora seliculosa. Mar Pollut Bull 2005, 51:1448–1451.

487. Jones RJ: Zooxanthellae loss as a bioassay for assessing stress in corals. Mar Ecol Prog Ser 1997, 149:163–171.

488. Retzel AM, Sullivan JC, Taylor-Knowles N, Finnerty JR: Genomic survey of candidate stress-response genes in the estuarine anemone Nematostella vectensis. Biol Bull 2008, 214:233–254.

489. Steele RA, Opella SJ: Structures of the reduced and mercury-bound forms of MerP, the periplasmic protein from the bacterial mercury detoxification system. Biochemistry 1997, 36:688–6895.

490. Mulkopadhyay R, Rosen BP: Arsenate reductases in prokaryotes and eukaryotes. Environ Health Perspect 2002, 110:745–748.

491. Hayakawa T, Kobayashi Y, Toyok X, Hirano S: A new metabolic pathway of arsenate/arsenite-glutathione complexes are substrates for human arsenic methytransferase Cy19. Arch Toxicol 2005, 79:183–191.

492. Price RE, London J, Wallaschäger D, Ruiz-Chancho M, Pichler T: Enhanced bioaccumulation and biotransformation of As in coral reef organisms surrounding a marine shallow-water hydrothermal vent system. Chem Geol 2012, doi:10.1016/j.chemgeo.2012.02.023.

493. David CB: Heavy metal concentrations in growth bands of corals: a record of mine tailings input through time (Marindique Island, Philippines). Mar Poll Bull 2003, 46:187–196.

494. Pessi G, Hoss D: Transcriptional control of the hydrogen cyanide biosynthetic genes hcnABC by the anaerobic regulator ANR and the quorum-sensing regulators LasR and RhlR in Pseudomonas aeruginosa. J Bacteriol 2000, 182:5960–5964.

495. Bird A: Perceptions of epigenetics. Nature 2007, 447:396–398.

496. Jablokova E, Lamb MJ: Epigenetic inheritance in evolution. J Evol Biol 1998, 11:59–183.

497. Bird A: Epigenetic inheritance in evolution. J Evol Biol 1998, 11:59–183.

498. Mirożewski M, Paskowski J: Epigenetic contribution to stress adaptation in plants. Curr Opin Plant Biol 2011, 14:267–274.

499. Srith B, Allis CD: The language of covalent histone modifications. Nature 2009, 460:41–45.

500. Wang M, Mock MW, Harper H, Lee WH, Min J, Knapp S, Oppermann U, Marsden B, Schapira M: Structural genomics of histone tail recognition. Bioinformatics 2010, 26:2629–2630.

501. Goll MG, Becker TH: Eukaryotic histone methyltransferases. Annu Rev Biochem 2005, 74:481–514.
503. Slokin RK, Martienssen R: Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 2007, 8:272–285.
504. Wolffe AP, Jones PL, Wade PA: DNA demethylation. Proc Natl Acad Sci U S A 1999, 96:5894–5896.
505. Ooi SK, Bestor TH: The colorful history of active DNA demethylation. Cell 2008, 133:145–1148.
506. Zhu JK: Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet 2009, 43:143–166.
507. Branco MR, Ficz G, Reik W: Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nat Rev Genet 2011, 13:27–13.
508. Guo AJ, Su Y, Zhong C, Meng GL, Song H: Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell 2011, 145:623–634.
509. Roberts SB, Gavery MR: Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? Front Physiol 2012, 2:116.
510. Brown BE, Cossins AR: The potential for temperature acclimatisation of reef corals in the face of climate change. In Coral Reefs: And Ecosystem in Transition. Edited by Dubinski Z, Stambler N. Berlin: Springer-Verlag; 2011:421–433.
511. Gilbert SF, McDonald E, Boyle N, Buttino N, Gyi L, Mai M, Prakash N, Robinson J: Symbiosis as a source of selectable epigenetic variation: taking the heat for the big guy. Phil Trans R Soc B 2010, 365:671–678.
512. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG: Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011, 7:539.
513. Apache-Tomcat.[http://tomcat.apache.org/download-70.cgi]

doi:10.1186/1471-2164-14-509
Cite this article as: Dunlap et al.: KEGG orthology-based annotation of the predicted proteome of Acropora digitifera: ZoophyteBase – an open access and searchable database of a coral genome. BMC Genomics 2013 14:509.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit