Journal of Experimental Marine Biology and Ecology 408 (2011) 120–125

Contents lists available at ScienceDirect

Journal of Experimental Marine Biology and Ecology

journal homepage: www.elsevier.com/locate/jembe

Recent progress in *Symbiodinium* transcriptomics

William Leggat [a,b,*], David Yellowlees [a], Monica Medina [c]

[a] ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland, 4811, Australia
[b] School of Pharmacy and Molecular Sciences, James Cook University, Townsville, Queensland, 4811, Australia
[c] School of Natural Sciences, University of California Merced, 5200 N. Lake Rd., Merced, CA 95343, USA

ARTICLE INFO

Available online 21 September 2011

Keywords:
Genomics
Minicircles
Symbiodinium
Symbiosis

ABSTRACT

Dinoflagellate symbionts of the genus *Symbiodinium* are integral to the success of the coral holobiont (a coral host and the microbial community it harbours), however despite their importance we currently have a very limited knowledge of the genes which they possess and their genomic organisation. Analysis shows that the number of expressed sequence tags (genes that are expressed) available for *Symbiodinium* (7964) is less than 1/10 of those available for the scleractinian coral host (103,434). This lack of DNA sequence information limits the functional genomic studies that can be undertaken from the symbiont perspective. In addition these sequences are from only three *Symbiodinium* types (C3, A1, A3) and do not represent the large diversity of clades and subclades seen. Here we summarise our current understanding of the *Symbiodinium* genomic content with reference to our knowledge of other dinoflagellates. The genetic information of *Symbiodinium* is encompassed in the nuclear, plastid and mitochondrial genomes. As is the case with other dinoflagellates these three genomes are significantly different from the "general" phototrophic eukaryote. Firstly the nuclear genome of dinoflagellates is extremely large, utilises modified DNA bases not normally found in eukaryotes, and tandem repeat regions seem to contain the most highly expressed genes. Meanwhile the plastid genome, which normally contains between 40 and 250 genes in other eukaryotes, has been reduced to 18 genes encoded in "minicircles." Finally the dinoflagellate mitochondrial genome only encodes for 2 or 3 proteins instead of the normal 40–50 in other eukaryotes. While we have some knowledge of *Symbiodinium* genome structure, little is known about its transcriptome. With the advent of inexpensive high throughput sequencing technologies, our understanding of the *Symbiodinium* genome will rapidly increase and we will begin to be able to look into the responses of these important single celled organisms.

1. Introduction

Recently the concept of the tripartite coral holobiont (Rohwer et al., 2001), consisting of the coral host, dinoflagellate symbiont and associated prokaryotes, has come to dominate how we view corals (this idea can be further expanded to include viruses (Vega-Thurber and Correa, 2010)). In any successful symbioses there must be a degree of integration and co-ordination of gene expression and metabolism in the different partners (Yellowlees et al., 2008). In the coral holobiont this integration is complicated by the fact that two of the partners (coral and dinoflagellate) are eukaryotic, in contrast to other partnerships such as the legume-*Rhizobium* or the squid-*Vibrio* symbioses, which involve eukaryote/prokaryote associations. An additional unusual aspect of the coral holobiont is that the genome size of the symbiont is larger than that of the host. As with any other photoautotroph, the *Symbiodinium* genetic makeup consists of three distinct genomes, the nuclear genome, the mitochondrial and the plastid genome, all of these have unique characteristics in *Symbiodinium* and more generally in dinoflagellates. In addition to this dinoflagellates have a variety of characteristics that make them unique amongst eukaryotic photoautotrophs, for example unique pigments and proteins associated with photosynthesis. This review will summarise our current state of knowledge of the *Symbiodinium* genomes, compare and contrast this information with our understanding of other dinoflagellates, and discuss how this knowledge can be used for functional genomics, which allows us to link gene expression changes to cellular, physiological and eventually ecological responses.

Dinoflagellates are the second largest group of eukaryotic photoautotrophs, yet despite this there has only been a slow accumulation of genomic information of this diverse group (Fig. 1). As with all genetic characterisation before the arrival of high-throughput sequencing, our early understanding of the *Symbiodinium* genome arose from studies on single genes or a few gene families. The first molecular work, focused on phylogenetic analysis of the small subunit ribosomal DNA gene revealed a great cryptic diversity in the genus (Rowan and Powers, 1992). Currently *Symbiodinium* are classified into nine separate clades (A–I) based solely

Abbreviations: EST, expressed sequence tag.

* Corresponding author at: Biochemistry and Molecular Biology, James Cook University, Townsville, 4811, Australia. Tel.: +61 7 47816923; fax: + 61 7 47816078.
E-mail addresses: bill.leggat@jcu.edu.au (W. Leggat), david.yellowlees@jcu.edu.au (D. Yellowlees), mmedina@ucmerced.edu (M. Medina).

0022-0981/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jembe.2011.07.032

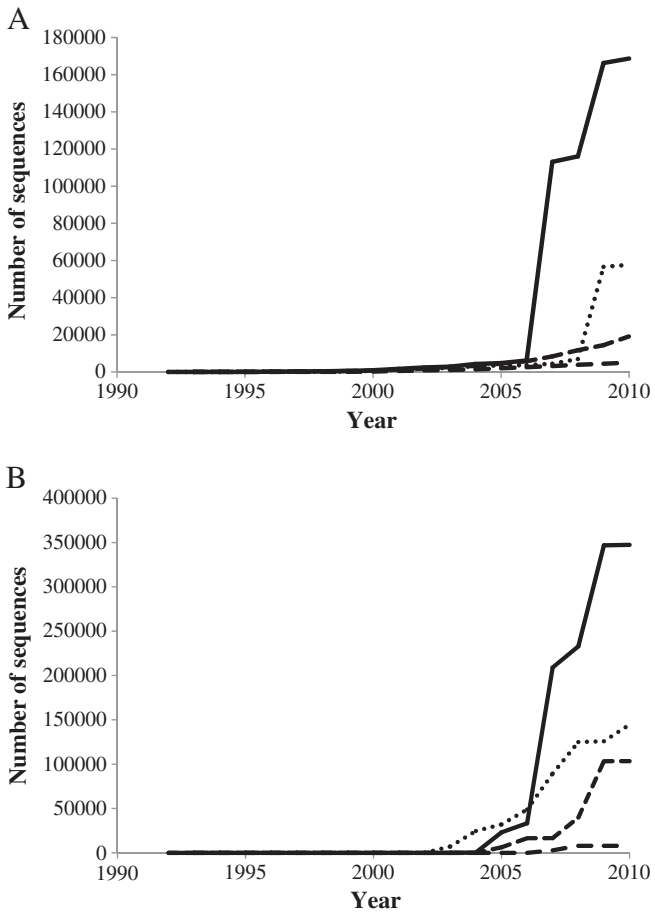

Fig. 1. Accumulation of DNA sequences deposited into the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov) for the (A.) nucleotide database and (B.) the Expressed Sequence Tag database. Data represents the number of sequences deposited for the Classes Anthozoa (——) and Dinophyceae (····) and the Orders Scleractinia (····) and Suessiales (– –) which respectively include corals and *Symbiodinium*.

upon genotypic markers (Pochon and Gates, 2010). The genotype diversity within the genus *Symbiodinium* is comparable to that seen amongst different orders of non-symbiotic dinoflagellates (Rowan and Powers, 1992). Each *Symbiodinium* clade contains numerous sub-clades (phylotypes) which are often differentiated by only a single base pair. Additional early characterisation of other genes, such as peridinin–chlorophyll *a/c* binding protein (PCP) (Norris and Miller, 1994), Rubisco (Morse et al., 1995; Rowan et al., 1996; Whitney et al., 1995; Whitney and Yellowlees, 1995) and a unique dinoflagellate light harvesting protein, integral chlorophyll *a*-chlorophyll c_2–peridinin protein complex (acpPC, Hiller et al. (1995)), demonstrated that there was a large amount of novelty in the dinoflagellate genome when compared to other photoautotrophs. Recently, a number of expressed sequence tag projects (Bachvaroff et al., 2004; Hackett et al., 2004; Hackett et al., 2005; Leggat et al., 2007; Patron et al., 2005; Patron et al., 2006; Tanikawa et al., 2004; Yoon et al., 2005) have dramatically increased our knowledge of the dinoflagellate transcriptome; with over 50,000 EST sequences now available, these studies have demonstrated that dinoflagellates have a unique genetic makeup.

1.1. Genomic organisation of dinoflagellates

Despite the ecological importance of dinoflagellates, until recently, we knew little about their genetic compliment, and what was known indicated that they were extremely different to other alga. Estimates of the genome size of dinoflagellates indicate that it ranges between 2 and 200 gigabases (Gb) (LaJeunesse et al., 2005), up to two orders of magnitude larger than the human genome (approximately 2.9 Gb). Interestingly *Symbiodinium* contains some of the smallest genome estimates for dinoflagellates (LaJeunesse et al., 2005); whether this reflects their smaller cell size, or a genome reduction due to their symbiotic life history, as is seen in other microbial endosymbionts (Cavalier-Smith, 2005; Keeling and Slamovits, 2004; Moran and Wernegreen, 2000; Sakharkar et al., 2004), is unknown (Lin, 2006).

The structure and organisation of the dinoflagellate nuclear genome are also unlike other eukaryotes (for review see Diaz de la Espina et al. (2005)) as they lack histones (the proteins involved in DNA packaging and ordering), and nucleosomes (the basic method of DNA packaging in eukaryotes). Especially unusual is the fact that the genomic DNA of dinoflagellates is permanently condensed, as if ready for division. There is also debate over the number of chromosomes present in *Symbiodinium*. In their published description of *S. microadriaticum*, *S. goreauii*, *S. kawagutii* and *S. pilosum* Trench and Blank (1987) reported the chromosome numbers to be 97 ± 2, 72 ± 2, 26 and 78 ± 2 respectively. However, this can be contrasted to the study of Udy et al. (1993) who found highly variable numbers of condensed DNA regions, ranging from less than 10 to over 100. This large diversity in chromosome number can also be seen in other dinoflagellates, where counts vary from 4 to 270 (Moreno Diaz de la Espina et al., 2005). Another unique feature of the dinoflagellate genome is the DNA bases used. In addition to the normal four DNA bases (adenine, cytosine, guanine and thymine), dinoflagellates are the only known eukaryote to utilise a large amount of the modified base hydroxymethyluracil, which can replace between 12 and 68% of thymidine residues (Rae, 1976).

The large genome size in dinoflagellates is thought to be due to high gene copy numbers. To this end, *Symbiodinium* Rubisco (Rowan et al., 1996), PCP (36 ± 12 copies/cell; Reichman et al. (2003)) and a variety of acPCP genes (Boldt et al. (2010)) has been shown to be multi-copy. Similarly other genes in dinoflagellates have also shown to have numerous copies within the genome (luciferin binding protein ~1000 copies/cell of *Gonyaulax polyedra*, Lee et al. (1993), PCP ~5000 copies/cell of *G. polyedra* Le et al. (1997), see Hou and Lin (2009) for an exhaustive list)). Using regression models based upon sequenced eukaryotic and prokaryotic genomes Hou and Lin (2009) estimate that *Symbiodinium* contains 38,188 protein-coding genes per genome (in comparison the human genome is estimated to contain between 20,000 and 25,000 (Human Genome Sequencing Consortium, 2004)). This is supported by estimates from high throughput sequencing of the dinoflagellate *Alexandrium tamarense* transcriptome which produces similar results, in the order of 40,000 genes, for dinoflagellate gene numbers (Moustafa et al., 2010).

Complicating analysis of gene numbers is that a proportion of the genome contains pseudo-genes, which are either not transcribed or translated, or do not encode functional proteins. For example Rowan et al. (1996) characterised a Rubisco gene (*rbcG*) from a clade A *Symbiodinium* strain which was missing almost half of the active site residues thought to be required for activity. Similarly, a genomic survey of *Amphidinium carterae* found that 10 of 46 actin genes were pseudogenes (Bachvaroff and Place, 2008), these high numbers of pseudogenes may also account for some of the large genome size of dinoflagellates.

Currently there is little genomic sequence information available for *Symbiodinium*, or any other dinoflagellate, and our understanding of genome organisation comes from small scale studies. PCR analysis of the genomic organisation of *A. carterae* indicated that there are at least two general categories of gene organisation in dinoflagellates. The first are highly expressed genes which are found in tandem arrays, with a short intergenic region, averaging 283 bp between the two genes, with low intron densities. In contrast those genes with a lower expression level are found as single isolated copies and are relatively intron rich (Bachvaroff and Place, 2008). Whether this arrangement holds true for *Symbiodinium* is unknown.

Another feature of the dinoflagellate transcriptome is a high proportion of trans-splicing, where a short conserved non-coding spliced leader sequence is added to the 5′ end of other pre-mRNA transcripts (Zhang et al., 2007). Identical spliced leader sequences have been identified for *Symbiodinium* (Zhang et al. (2007), Leggat unpub. data) and may be useful in obtaining full length transcripts as has been done in other dinoflagellates (Lapointe et al., 2008). Spliced leader sequences can serve a variety of roles, including regulating gene translation (the process by which mRNA is used as a template for protein synthesis). The mechanisms of gene regulation in dinoflagellates are still unclear, given that their permanently condensed chromosomes would not seem to allow for the usual methods of transcription seen in other eukaryotes. In fact, analysis of the few microarray studies that have been conducted on dinoflagellates would seem to indicate that there may not be large transcriptional changes seen even with large changes in conditions (Okamoto and Hastings, 2003a; Okamoto and Hastings, 2003b; Van Dolah et al., 2007), instead it has been suggested that much of the regulation may occur post-translationally (Bachvaroff and Place, 2008; Fagan et al., 1999; Okamoto et al., 2001). This suggestion that expression of much of the transcriptome does not significantly change under contrasting conditions is supported by a recent massive parallel signature sequencing study conducted on the dinoflagellate *A. tamarense* (Moustafa et al., 2010). When this dinoflagellate was exposed to nutrient replete, nitrogen limited, phosphorus limited and xenic culture conditions approximately 73% of the transcriptome showed no significant changes in expression. Interestingly, the treatment that produced the largest change in expression profile occurred in response to the presence of bacteria, which caused the expression of 487 genes which were not found in any other treatment (Moustafa et al., 2010). Given that the coral holobiont model (Rohwer et al., 2001) consists of a diverse prokaryote community, future research will need to determine if a similar transcriptomic response to bacteria is seen in *Symbiodinium*.

Epigenetic mechanisms of gene regulation, i.e. DNA methylation, that may be important in dinoflagellates, have so far received little attention. Methylation occurs when a methyl group is added to a cytosine residue in a cytosine-guanine di-nucleotide sequence (CpG) or a cytosine-any nucleotide-guanine tri-peptide (CpNpG) and functions to prevent transcription. Restriction digest analysis of *Symbiodinium* genomic DNA indicates that a high proportion of CpG and CpNpG motifs are methylated (ten Lohuis and Miller, 1998a). Decreased methylation patterns near or within the coding regions of two genes (PCP and acpPCP), were associated with an up-regulation of the transcripts (ten Lohuis and Miller, 1998b).

1.2. Plastid genome

The chloroplast in all eukaryotic photoautotrophs is thought to have arisen from an initial endosymbiotic event in which a eukaryote engulfed a photosynthetic cyanobacterium. Subsequent divergence, approximately 500 mya, leads to three divergent chloroplast lineages, the green (Chlorophyta — higher plants and green algae), the red lineage Rhodophytes — red algae) and the glaucophytes (Delwiche, 1999; Gould et al., 2008). Secondary endosymbiotic events (engulfing of both green and red algae by other eukaryotes) gave rise to other photoautotrophs lineages, such as euglenophytes, kelps, diatoms, coccolithophorids and dinoflagellates (there are also examples of tertiary engulfment in some dinoflagellates). The *Symbiodinium* chloroplast is derived from secondary endosymbiosis of a red alga.

Chloroplast genomes are normally composed of a single circular genome of approximately 70–200 kb located in the chloroplast stroma, which encodes for between 42 and 251 chloroplast proteins (Green, 2004; Wolfe et al., 1992), the remaining chloroplast proteins (estimated to be in the range of 2100–4800 in higher plants (Richly and Leister, 2004)) are found in the nuclear genome. In contrast dinoflagellate plastid genomes have been fragmented and encoded genes are found on small plasmid like minicircles (Barbrook et al., 2001; Barbrook and Howe, 2000; Moore et al., 2003; Zhang et al., 1999; Zhang et al., 2001) of between 2 and 4 kb. These minicircles have been shown to encode eight photosystem proteins (*psaA*, *psaB*, *psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbI*), subunits of ATP synthase (*atpA*, *atpB*) and the cytochrome b_6f complex (*petB*, *petD*), two ribosomal proteins (*rpl28*, *rpl33*), two unknown open reading frames (*ycf16*, *ycf24*) and the large (23S) and small (16s) ribosomal RNA in a number of different dinoflagellate species, including *Symbiodinium* (Moore et al., 2003). Most of those gene, which are normally found in the plastid genome, have instead been transferred to the nuclear genome. Each of the minicircles encodes normally 1, or occasionally 2 proteins, although where two are contained within the one minicircle they are not encoded as a polycistronic transcript (Barbrook et al., 2001). Recently an exception to this rule was reported where it was found that in the dinoflagellate *H. triquetra* minicircles encoding for *petD*, *psbE* and *psbA* also encoded for either tRNA-Met, tRNA-Trp, tRNA-Pro which were co-transcribed and then cleaved from the protein encoding sequence (Nelson et al., 2007). There is also evidence that non-traditional start codons (GTA) may be found in some of the minicircles (Barbrook and Howe, 2000) while some minicircles do not appear to contain any recognisable coding region. An exception to this pattern of chloroplast organisation was recently found in the dinoflagellate *G. polyedra*, where the *psbA* gene seems to be associated with DNA of roughly 50–150 kb, the majority of which appears to be non-coding (Wang and Morse, 2006). Whether this organisation represents a derived or ancestral state is unknown. There is also the possibility that in some species minicircles are found not in the chloroplast but instead in the nucleus (Laatsch et al., 2004), how these proteins are then transported to the chloroplast without the presence of obvious plastid targeting sequences remains unknown.

Despite the key role that protein products of these genes play in photosynthesis, as yet we have little understanding of how the expression of the minicircle genes is regulated. Getting a good handle on this question is particularly relevant today, given that the D1 protein, which is coded for by *psbA*, is one of the possible points of thermal damage causing coral bleaching (Warner et al., 1999). What is known about minicircle regulation is that copy numbers are between 1 and 5 minicircle copies/cell, during exponential growth phase, and increase dramatically to between 50 and 420 copies during slower growth periods (Koumandou and Howe, 2007). Other unique features of the minicircles include a common “core” region which is conserved amongst minicircles of the one species and may also be conserved within genera. Given this conservation it has been suggested that this core region and linked variable region in the *psbA* gene could be used as a phylogenetic marker for closely related *Symbiodinium* strains (Moore et al., 2003).

The fact that dinoflagellates have highly reduced plastid genomes has meant that nearly all plastid proteins are nuclear encoded, when compared to other eukaryote photoautotrophs. A number of these nuclear encoded plastid genes encode polyproteins, i.e. multiple copies of the one protein are transcribed as one mRNA, translated as one protein, imported into the chloroplast and then cleaved into their final mature states. Polyproteins, or transcripts encoding multiple copies of protein domains, have been identified for Rubisco (Rowan et al., 1996), PCP (Norris and Miller, 1994; Reichman et al., 2003), chlorophyll *a*-chlorophyll c_2-peridinin (acpPCP) binding proteins (Hiller et al. (1995); Boldt, Yellowlees and Leggat unpub.), carbonic anhydrase (Bobeszko, Yellowlees and Leggat unpub.) and a protein of unknown function (Bacl) (Leggat et al., 2007).

1.3. Mitochondrial genome

As with the chloroplast, mitochondria arose through an ancient endosymbiotic engulfment of an α-proteobacteria. In all organisms

there has been a transfer of genes from the mitochondria to the nuclear genome, with mitochondrial genomes generally containing between 40 and 50 genes on a single circular genome (Burger et al., 2003). However, as with the plastid genome, this is not the case with dinoflagellates. They instead encode only 2 or 3 proteins, which have undergone rRNA fragmentation, lack canonical stop codons and have extensive RNA editing, and as yet, we have little information on the structure of the mitochondrial genome (e.g. circular or linear, number of genomes) (for review see Waller and Jackson (2009)). Given that these characteristics have been found in all dinoflagellates examined to date it is expected that the *Symbiodinium* mitochondrial genome would display similar characteristics.

1.4. Symbiodinium ESTs

Only two studies have examined the transcriptome of *Symbiodinium* (Leggat et al., 2007; Voolstra et al., 2009). Leggat et al. (2007) generated a pooled library of *Symbiodinium* clade C3 genes from a variety of stress and non-stress treatments using both freshly isolated and short term cultures. Of the 1456 unique ESTs generated, the function of only 44% could be bioinformatically assigned, 19% which were of unknown function, 12% were assigned to protein turnover and chaperones, and 12% to energy production and conversion. Interestingly only 10% of these contigs matched any of the 50,000 previously identified dinoflagellate ESTs available at that time. Surprisingly the most abundant gene, making up approximately 2% of the transcriptome, encoded a previously uncharacterised bacterial like protein targeted to the chloroplast, again emphasising the unique nature of the dinoflagellate genome. Similarly studies of this type have identified a number of dinoflagellate specific genes (Hackett et al., 2005; Leggat et al., 2007).

In an effort to understand evolution rates in *Symbiodinium* Voolstra et al. (2009) compared ESTs from *Symbiodinium* clade C3 (3336 unique sequences) and CassKB8 (clade A1; 1484 unique sequences). A comparison of 115 orthologs found that the highest d*N*/d*S* ratio (the ratio of non-synonymous to synonymous nucleotide changes, which is a measure of evolutionary rate) was highest in those genes that were *Symbiodinium* specific. While the study was limited by the small sample size, the data suggested that *Symbiodinium* specific genes are under greater positive selection than those genes that are more widely distributed, hinting at the possibility that their function may be linked to symbiosis (Voolstra et al., 2009). Given that clade A *Symbiodinium* are ancestral, and possibly more adapted to a free living lifestyle, while clade C is more derived and seems more adapted to a symbiotic lifestyle, a more comprehensive analysis with multiple clades and subclades may provide further support for this hypothesis.

Given the current relatively small size of the *Symbiodinium* transcriptomic data (Table 1, Fig. 1), and the evolution of high throughput sequencing technology, the characterisation of a variety of transcriptomes is the obvious next step in understanding the *Symbiodinium* genome.

Table 1
Symbiodinium EST available through NCBI.

Clade	Number of EST	Conditions
C3	5184	Freshly isolated from *Acropora aspera* which had been subjected to either elevated temperature, elevated ammonium, or freshly isolated alga subjected to elevated or reduced CO_2
A1 (KB8)	2653	Cultured at 28 °C under 12:12 light:dark cycle
A3 (PL-TS-1)	127	Cultured at 24 °C under 12:12 light:dark cycle

1.5. Horizontal gene transfer

An interesting feature of dinoflagellates is that they have acquired a significant number of genes through the process of horizontal gene transfer (HGT), where organisms acquire genetic material from a different species. HGT is thought to be one of the major driving forces in prokaryote evolution and until recently had not been widely reported in eukaryotes. A variety of studies has now documented cases of horizontal gene transfer from cyanobacteria (Waller et al., 2006) and other eukaryotes (Fagan et al., 1998; Fagan and Hastings, 2002; Takishita et al., 2003; Wisecaver and Hackett, 2010) into dinoflagellate nuclear genomes. Single gene studies have identified two genes which are present in *Symbiodinium* as a result of HGT, these are a glyceraldehyde-3-phosphate-dehydrogenase (Takishita et al., 2003), which is thought to have been transferred from euglenophytes, and the form II Rubisco (Rowan et al., 1996) which is only elsewhere found in α-proteobacteria. Larger scale analysis of ESTs identified another 48 (out of a total of 1456) genes which were possible cases of HGT in *Symbiodinium* as they had previously only been identified in bacteria or other dinoflagellates (Leggat et al., 2007). These genes encoded proteins that were involved in a variety of functions including carbohydrate transport, inorganic ion transport, stress response and proteases, amongst others. Studies of other dinoflagellate ESTs have also hinted at numerous HGT events. Interestingly, Nosenko and Bhattacharya (2007) proposed that ecological niches where parasitism (and presumably symbiosis), and phagotrophy are common, beneficial genes may be more commonly transferred from prokaryotes to eukaryotes. Given the symbiotic nature of *Symbiodinium* it will be interesting to compare rates of HGT into *Symbiodinium* and other free-living dinoflagellates, as well as amongst *Symbiodinium* clades. Certainly it appears that chromalveolates, and in particular dinoflagellates, are adept at acquiring and integrating novel genes into their genome.

1.6. Functional genomics

Functional genomics links the large scale genomic resources that are becoming available to gene expression, cell biology, physiology and eventually ecological processes. Transcriptome analyses allow us to rapidly examine the expression response of a wide variety of genes. The main methods used for trascriptome studies are quantitative PCR (qPCR), gene expression microarrays, and recently RNA-seq. As yet, there have been no dedicated *Symbiodinium* microarray studies, while studies using qPCR have only recently become available. There are two factors limiting the application of qPCR to studies of *Symbiodinium*, the first is the lack of available sequence information for different clades, while the second is the availability of suitable house-keeping genes (HKG) for use as references. We are starting to overcome the sequence data challenge as new sequence data are emerging. Voolstra et al. (2009) have completed the only large scale comparison of genes from different *Symbiodinium* strains, perhaps the most surprising result was the low sequence similarity between clade A1 and C3. Several new large-scale *Symbiodinium* transcriptomes based on next-generation sequencing technologies will soon become available (Leggat unpublished, Medina unpublished). The large sequence divergence between clades, and also large divergence of some genes within a clade (Boldt et al. unpub.), make it difficult to design gene primers for qPCR that are suitable for multiple clades/subclades. Another necessary requirement for qPCR is the availability of effective HKGs. These internal controls are used as a comparator when examining gene expression changes, the assumption being that the expression of HKGs does not change between the control and treatment conditions. To this end, recently two papers (Boldt et al., 2010; Rosic et al., 2010) have presented the first HKG available for *Symbiodinium* under particular treatments filling a critical gap in the application of this methodology (Table 2). This ability to examine gene expression patterns, and integrate this

Table 2
Symbiodinium house-keeping genes used for quantitative PCR.

Gene	*Symbiodinium* clade (where known)	Stress	Citation
Proliferating cell nuclear antigen (PCNA)	C3	Light	Boldt et al. (2010)
Actin	C3	Light	Boldt et al. (2010)
18S RNA	C3	Light	Boldt et al. (2010)
Cytochrome oxidase subunit 1	C1, C3	Light/temperate	Rosic et al. (2010)
Calmodulin	C1, C3	Light/temperate	Rosic et al. (2010)
Ribosomal protein S4	C1, C3	Light/temperate	Rosic et al. (2010)
Cyclophin	C1, C3	Light/temperate	Rosic et al. (2010)
S-adenosyl-L methionine synthetase	C1, C3	Light/temperate	Rosic et al. (2010)
Beta tubulin	C1, C3	Light/temperate	Rosic et al. (2010)

data with other "omic" approaches, such as proteomics and metabolomics (Gordon and Leggat, 2010), will provide a window on subcellular responses as never before imagined. In turn this information will enable researchers to answer many of the fundamental questions which are still unresolved in *Symbiodinium* symbioses.

1.7. Future directions for Symbiodinium genomics

With the advent of high throughput cost effective sequencing there is likely to be an explosion in the EST sequence data available for different *Symbiodinium* strains and there is now the real possibility of a *Symbiodinium* genome (Voolstra, Ravasi and Bajic; http://coral.aoml.noaa.gov/pipermail/coral-list/2010-May/008559.html) using the Illumina sequencing platform. The ability to examine and compare the transcriptome response to different conditions in multiple species may provide an important tool in understanding what drives physiological differences at both the intra- and inter-cladal level. To date over 150 different phylotypes have been identified yet the ecological relevance of these small sequence variations is unknown (cf. Savage et al., 2003, Correa and Baker, 2009 and Sampayo et al, 2008) and the question of differential *Symbiodinium* performance within and between clades has not been thoroughly examined. The ability to utilise molecular tools to dissect differential transcriptome responses will provide us with a new window into how these important organisms function and link symbiont responses to those of the host under different conditions (Ainsworth et al., 2008; Baird et al., 2009). The ability to utilise both *Symbiodinium* genomic resources and those of the coral host (Miller et al., 2011) will be a powerful tool. This effort will be particularly relevant in the near future given that the sensitivity of the coral symbiosis to stress can be significantly dependent upon the phylotype of *Symbiodinium* that reside within the coral.

Note added in proof

Since submission of this manuscript there have been a number of developments in the area of Symbiodinium and coral genomics. Medina et al. have released new expressed sequence tags for two Symbiodinium strains (Mf104b and KB8; http://medinalab.org/zoox/) while genome assemblies have been released for two corals, Acropora digitifera (http://marinegenomics.oist.jp/acropora_digitifera) and Acropora millepora (http://coralbase.org/). **[SS]**

References

Ainsworth, T.D., Hoegh-Guldberg, O., Heron, S.F., Skirving, W.J., Leggat, W., 2008. Early cellular changes are indicators of pre-bleaching thermal stress in the coral host. JEMBE 364, 63–71.

Bachvaroff, T.R., Place, A.R., 2008. From stop to start: tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. PLoS One 3 (8), e2929.

Bachvaroff, T.R., Concepcion, G.T., Rogers, C.R., Herman, E.M., Delwiche, C.F., 2004. Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. Protist 155, 65–78.

Baird, A.H., Bhagooli, R., Ralph, P.J., Takahashi, S., 2009. Coral bleaching: the role of the host. TREE 24 (1), 16–20.

Barbrook, A.C., Howe, C.J., 2000. Minicircular plastid DNA in the dinoflagellate *Amphidinium operculatum*. Mol. Gen. Genet. 263, 152–158.

Barbrook, A.C., Symington, H., Nisbet, R.E.R., Larkum, A.W.D., Howe, C.J., 2001. Organisation and expression of the plastid genome of the dinoflagellate *Amphidinium operculatum*. Mol. Genet. Genomics 266, 632–638.

Boldt, L., Yellowlees, D., Leggat, W., 2010. Measuring *Symbiodinium* sp. Gene Expression Patterns with Quantitative Real-time PCR Proceedings of the 11th ICRS, Ft Lauderdale, Florida 7–11 July 2008, pp. 118–122.

Burger, G., Gray, M.W., Franz Lang, B., 2003. Mitochondrial genomes: anything goes. Trends Genet. 19 (12), 709–716.

Cavalier-Smith, T., 2005. Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Ann. Bot. 95 (1), 147–175.

Correa, A., Baker, A., 2009. Understanding diversity in coral-algal symbiosis: a cluster-based approach to interpreting fine-scale genetic variation in the genus *Symbiodinium*. Coral Reefs 28 (1), 81–93.

de la Espina, S.M.D., Alverca, E., Cuadrado, A., Franca, S., 2005. Organization of the genome and gene expression in a nuclear environment lacking histones and nucleosomes: the amazing dinoflagellates. Eur. J. Cell Biol. 84 (2–3), 137–149.

Delwiche, C.F., 1999. Tracing the thread of plastid diversity through the tapestry of life. Am. Nat. 154, 164–177.

Fagan, T.F., Hastings, J.W., 2002. Phylogenetic analysis indicates multiple origins of chloroplast glyceraldehyde-3-phosphate dehydrogenase genes in dinoflagellates. Mol. Biol. Evol. 19 (7), 1203–1207.

Fagan, T.F., Hastings, J.W., Morse, D., 1998. The Phylogeny of Glyceraldehyde-3-phosphate Dehydrogenase Indicates Lateral Gene Transfer from Cryptomonads to Dinoflagellates, pp. 633–639.

Fagan, T., Morse, D., Hastings, J.W., 1999. Circadian synthesis of a nuclear-encoded chloroplast glyceraldehyde-3-phosphate dehydrogenase in the dinoflagellate *Gonyaulax polyedra* is translationally controlled. Biochemistry 38 (24), 7689–7695.

Gordon, B.R., Leggat, W., 2010. *Symbiodinium*-invertebrate symbioses and the role of metabolomics. Mar. Drugs 8, 2546–2568.

Gould, S.B., Waller, R.F., McFadden, G.I., 2008. Plastid evolution. Annu. Rev. Plant Biol. 59 (1), 491–517.

Green, B.R., 2004. The chloroplast genome of dinoflagellates — a reduced instruction set? Protist 155, 23–32.

Hackett, J.D., Yoon, H.S., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Nosenko, T., Bhattacharya, D., 2004. Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. Curr. Biol. 14 (3), 213–218.

Hackett, J., Scheetz, T., Yoon, H., Soares, M., Bonaldo, M., Casavant, T., Bhattacharya, D., 2005. Insights into a dinoflagellate genome through expressed sequence tag analysis. BMC Genomics 6 (1), 80.

Hiller, R.G., Wrench, P.M., Sharples, F.P., 1995. The light-harvesting chlorophyll a-c-binding protein of dinoflagellates: a putative polyprotein. FEBS Lett. 363 (1–2), 175–178.

Hou, Y., Lin, S., 2009. Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes. PLoS One 4 (9), e6978.

Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. Nature 431 (7011), 931–945.

Keeling, P.J., Slamovits, C.H., 2004. Simplicity and complexity of microsporidian genomes. Eukaryot. Cell 3 (6), 1363–1369.

Koumandou, V.L., Howe, C.J., 2007. The copy number of chloroplast gene minicircles changes dramatically with growth phase in the dinoflagellate *Amphidinium operculatum*. Protist 158 (1), 89–103.

Laatsch, T., Zauner, S., Stoebe-Maier, B., Kowallik, K.V., Maier, U.G., 2004. Plastid-derived single gene minicircles of the dinoflagellate *Ceratium horridum* are localized in the nucleus. Mol. Biol. Evol. 21 (7), 1318–1322.

LaJeunesse, T.C., Lambert, G., Andersen, R.A., Coffroth, M.A., Galbraith, D.W., 2005. *Symbiodinium* (Pyrrophyta) genome sizes (DNA content) are the smallest among dinoflagellates. J. Phycol. 42 (3), 746–748.

Lapointe, M., MacKenzie, T.D.B., Morse, D., 2008. An external δ-carbonic anhydrase in a free-living marine dinoflagellate may circumvent diffusion-limited carbon acquisition. Plant Physiol. 147 (3), 1427–1436.

Le, Q.H., Markovic, P., Hastings, J.W., Jovine, R.V.M., Morse, D., 1997. Structure and Organization of the Peridinin-chlorophyll a-Binding Protein Gene in *Gonyaulax polyedra*, pp. 595–604.

Lee, D.H., Mittag, M., Sczekan, S., Morse, D., Hastings, J.W., 1993. Molecular cloning and genomic organization of a gene for luciferin-binding protein from the dinoflagellate *Gonyaulax polyedra*. J. Biol. Chem. 268, 8842–8850.

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

Lin, S., 2006. The smallest dinoflagellate genome is yet to be found: a comment on LaJeunesse et al. "*Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates". J. Phycol. 42 (3), 746–748.

Miller, D.J., Ball, E.E., Foret, S., Satoh, N., 2011. Coral genomics and transcriptomics - ushering in a new era in coral biology. JEMBE 408, 114–119 (this issue).

Moore, R.B., Ferguson, K.M., Loh, W.K.W., Hoegh-Guldberg, O., Carter, D.A., 2003. Highly organized structure in the non-coding region of the psbA minicircle from clade C *Symbiodinium*. Int. J. Syst. Evol. Microbiol. 53 (6), 1725–1734.

Moran, N.A., Wernegreen, J.J., 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. TREE 15 (8), 321–326.

Morse, D., Salois, P., Markovic, P., Hastings, J.W., 1995. A nuclear-encoded form II RuBisCO in dinoflagellates. Science 268, 1622–1624.
Moustafa, A., Evans, A.N., Kulis, D.M., Hackett, J.D., Erdner, D.L., Anderson, D.M., Bhattacharya, D., 2010. Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. PLoS One 5 (3), e9688.
Nelson, M.J., Dang, Y., Filek, E., Zhang, Z., Yu, V.W.C., Ishida, K.i., Green, B.R., 2007. Identification and transcription of transfer RNA genes in dinoflagellate plastid minicircles. Gene 392 (1–2), 291–298.
Norris, B.J., Miller, D.J., 1994. Nucleotide sequence of a cDNA clone encoding the precursor of the peridinin-chlorophyll a-binding protein from the dinoflagellate *Symbiodinium* sp. Plant Mol. Biol. 24 (4), 673–677.
Nosenko, T., Bhattacharya, D., 2007. Horizontal gene transfer in chromalveolates. BMC Evol. Biol. 7 (1), 173.
Okamoto, O.K., Hastings, J.W., 2003a. Genome-wide analysis of redox-regulated genes in a dinoflagellate. Gene 321, 73–81.
Okamoto, O.K., Hastings, J.W., 2003b. Novel dinoflagellate clock-related genes identified through microarray analysis. J. Phycol. 39 (3), 519–526.
Okamoto, O.K., Robertson, D.L., Fagan, T.F., Hastings, J.W., Colepicolo, P., 2001. Different regulatory mechanisms modulate the expression of a dinoflagellate iron-superoxide dismutase. J. Biol. Chem. 276 (23), 19989–19993.
Patron, N.J., Waller, R.F., Archibald, J.M., Keeling, P.J., 2005. Complex protein targeting to dinoflagellate plastids. J. Mol. Biol. 348 (4), 1015–1024.
Patron, N.J., Waller, R.F., Keeling, P.J., 2006. A tertiary plastid uses genes from two endosymbionts. J. Mol. Biol. 357 (5), 1373–1382.
Pochon, X., Gates, R.D., 2010. A new *Symbiodinium* clade (Dinophyceae) from soritid foraminifera in Hawaii. Mol. Phylogenet. Evol. 56 (1), 492–497.
Rae, P.M., 1976. Hydroxymethyluracil in eukaryote DNA: a natural feature of the pyrrophyta (dinoflagellates). Science 194 (4269), 1062–1064.
Reichman, J.R., Wilcox, T.P., Vize, P.D., 2003. PCP gene family in *Symbiodinium* from *Hippopus hippopus*: low levels of concerted evolution, isoform diversity, and spectral tuning of chromophores. Mol. Biol. Evol. 20 (12), 2143–2154.
Richly, E., Leister, D., 2004. An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. Gene 329, 11–16.
Rohwer, F., Breitbart, M., Jara, J., Azam, F., Knowlton, N., 2001. Diversity of bacteria associated with the Caribbean coral *Montastraea franksi*. Coral Reefs 20, 85–91.
Rosic, N., Pernice, M., Rodriguez-Lanetty, M., Hoegh-Guldberg, O., 2010. Validation of housekeeping genes for gene expression studies in *Symbiodinium* exposed to thermal and light stress. Mar. Biotechnol. 1–11.
Rowan, R., Powers, D.A., 1992. Ribosomal RNA sequences and the diversity of symbiotic dinoflagellates (zooxanthellae). PNAS 89, 3639–3643.
Rowan, R., Whitney, S.M., Fowler, A., Yellowlees, D., 1996. Rubisco in marine symbiotic dinoflagellates: Form II enzyme in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family. Plant Cell 8, 539–553.
Sakharkar, K.R., Dhar, P.K., Chow, V.T.K., 2004. Genome reduction in prokaryotic obligatory intracellular parasites of humans: a comparative analysis. Int. J. Syst. Evol. Microbiol. 54 (6), 1937–1941.
Sampayo, E.M., Dove, S., LaJeunesse, T.C., 2008. Cohesive molecular genetic data delineate species diversity in the dinoflagellate genus *Symbiodinium*. Mol. Ecol. 18 (3), 500–519.
Savage, A.M., Trapido-Rosenthal, H.G., Douglas, A.E., 2003. On the functional significance of molecular variation in *Symbiodinium*, the symbiotic algae of Cnidaria: photosynthetic response to irradiance. MEPS 244, 27–37.
Takishita, K., Ishida, K.i., Maruyama, T., 2003. An enigmatic GAPDH gene in the symbiotic dinoflagellate genus *Symbiodinium* and its related species (the Order Suessiales): possible lateral gene transfer between two eukaryotic algae, dinoflagellate and euglenophyte. Protist 154 (3–4), 443–454.
Tanikawa, N., Akimoto, H., Ogoh, K., Chun, W., Ohmiya, Y., 2004. Expressed sequence tag analysis of the dinoflagellate *Lingulodinium polyedrum* during dark phase. Photochem. Photbiol. 80 (1), 31–35.
ten Lohuis, M.R., Miller, D.J., 1998a. Hypermethylation at CpG-motifs in the dinoflagellates *Amphidinium carterae* (Dinophyceae) and *Symbiodinium microadriaticum* (Dinophyceae): evidence from restriction analyses, 5-azacytidine and ethionine treatment. J. Phycol. 34 (1), 152–159.
ten Lohuis, M.R., Miller, D.J., 1998b. Light-regulated transcription of genes encoding peridinin chlorophyll a proteins and the major intrinsic light-harvesting complex proteins in the dinoflagellate *Amphidinium carterae* Hulburt (Dinophycae) — changes in cytosine methylation accompany photoadaptation. Plant Physiol. 117 (1), 189–196.
Trench, R.K., Blank, R.J., 1987. *Symbiodinium microadriaticum* Freudenthal, *Symbiodinium goreauii* Sp-Nov, *Symbiodinium kawagutii* Sp-Nov and *Symbiodinium pilosum* Sp-Nov — Gymnodinioid dinoflagellate symbionts of marine-invertebrates. J. Phycol. 23 (3), 469–481.
Udy, J.W., Hinde, R., Vesk, M., 1993. Chromosomes and DNA in *Symbiodinium* from Australian hosts. J. Phycol. 29 (3), 314–320.
Van Dolah, F.M., Lidie, K.B., Morey, J.S., Brunelle, S.A., Ryan, J.C., Monroe, E.A., Haynes, B.L., 2007. Microarray analysis of diurnal- and circadian-regulated genes in the Florida red-tide dinoflagellate *Karenia brevis* (Dinophyceae). J. Phycol. 43 (4), 741–752.
Vega-Thurber, R.L., Correa, A.M.S., 2010. Viruses of reef-building scleractinian corals. JEMBE.
Voolstra, C.R., Sunagawa, S., Schwarz, J.A., Coffroth, M.A., Yellowlees, D., Leggat, W., Medina, M., 2009. Evolutionary analysis of orthologous cDNA sequences from cultured and symbiotic dinoflagellate symbionts of reef-building corals (Dinophyceae: *Symbiodinium*). Comp. Biochem. Physiol. D:Gen. Prot. 4 (2), 67–74.
Waller, R.F., Jackson, C.J., 2009. Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. BioEssays 31, 237–245.
Waller, R.F., Slamovits, C.H., Keeling, P.J., 2006. Lateral gene transfer of a multigene region from cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. Mol. Biol. Evol. 23 (7), 1437–1443.
Wang, Y., Morse, D., 2006. The plastid-encoded psbA gene in the dinoflagellate *Gonyaulax* is not encoded on a minicircle. Gene 371 (2), 206–210.
Warner, M.E., Fitt, W.K., Schmidt, G.W., 1999. Damage to photosystem II in symbiotic dinoflagellates: a determinant of coral bleaching. PNAS 96, 8007–8012.
Whitney, S.M., Yellowlees, D., 1995. Preliminary investigations into the structure and activity of ribulose bisphosphate carboxylase from two photosynthetic dinoflagellates. J. Phycol. 31, 138–146.
Whitney, S.M., Shaw, D.C., Yellowlees, D., 1995. Evidence that some dinoflagellates contain a ribulose-1,5-bisphosphate carboxylase/oxygenase related to that of the α-proteobacteria. Proc. R. Soc. Lond. B 259, 271–275.
Wisecaver, J., Hackett, J., 2010. Transcriptome analysis reveals nuclear-encoded proteins for the maintenance of temporary plastids in the dinoflagellate *Dinophysis acuminata*. BMC Genomics 11 (1), 366.
Wolfe, K.H., Morden, C.W., Palmer, J.D., 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. PNAS 89 (22), 10648–10652.
Yellowlees, D., Rees, T.A.V., Leggat, W., 2008. Metabolic interactions between algal symbionts and invertebrate hosts. Plant Cell Environ. 31, 679–694.
Yoon, H.S., Hackett, J.D., Van Dolah, F.M., Nosenko, T., Lidie, K.L., Bhattacharya, D., 2005. Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. Mol. Biol. Evol. 22 (5), 1299–1308.
Zhang, Z., Green, B.R., Cavalier-Smith, T., 1999. Single gene circles in dinoflagellate chloroplast genomes. Nature 400, 155–159.
Zhang, Z., Cavalier-Smith, T., Green, B.R., 2001. A family of selfish minicicular chromosomes with jumbled chloroplast gene fragements from a dinoflagellate. Mol. Biol. Evol. 18 (8), 1558–1565.
Zhang, H., Hou, Y., Miranda, L., Campbell, D.A., Sturm, N.R., Gaasterland, T., Lin, S., 2007. Spliced leader RNA trans-splicing in dinoflagellates. PNAS 104 (11), 4618–4623.