Genomic adaptations to an endolithic lifestyle in the coral-associated alga *Ostreobium*

Cintia Iha\(^a\)*, Javier A. Varela\(^b\), Viridiana Avila\(^c\), Christopher J. Jackson\(^a\), Kenny A. Bogaert\(^d\), Yibi Chen\(^e\), Louise M. Judd\(^f\), Ryan Wick\(^f\), Kathryn E. Holt\(^f,g\), Marisa M. Pasella\(^a\), Francesco Ricci\(^a\), Sonja I. Repetti\(^a\), Mónica Medina\(^c\), Vanessa R. Marcelino\(^b\), Cheong X. Chan\(^c\), Heroen Verbruggen\(^a\)*

\(^a\)School of BioSciences, University of Melbourne, Victoria 3010, Australia
\(^b\)School of Microbiology/Centre for Synthetic Biology and Biotechnology/Environmental Research Institute/APC Microbiome Institute, University College Cork, Cork T12 YN60, Ireland
\(^c\)Pennsylvania State University, University Park, PA, 16802, USA
\(^d\)Phycology Research Group, Ghent University, Krijgslaan 281 S8, 9000 Gent, Belgium
\(^e\)School of Chemistry and Molecular Biosciences and Australian Centre for Ecogenomics, The University of Queensland, Brisbane, Queensland 4072, Australia
\(^f\)Department of Infectious Diseases, Central Clinical School, Monash University, Melbourne, 3004, Australia
\(^g\)London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK

*corresponding author: cintiaiha@gmail.com, heroen@unimelb.edu.au
Abstract

The green alga *Ostreobium* is an important coral holobiont member, playing key roles in skeletal decalcification and providing photosynthate to bleached corals that have lost their dinoflagellate endosymbionts. *Ostreobium* lives in the coral’s skeleton, a low-light environment with variable pH and O₂ availability. We present the *Ostreobium* nuclear genome and a metatranscriptomic analysis of healthy and bleached corals to improve our understanding of *Ostreobium*’s adaptations to its extreme environment and its roles as a coral holobiont member.

The *Ostreobium* genome has 11,735 predicted protein-coding genes and shows adaptations for life in low and variable light conditions and other stressors in the endolithic environment. It presents an unusually rich repertoire of light harvesting complex proteins, lacks many genes for photoprotection and photoreceptors, and has a large arsenal of genes for oxidative stress response. An expansion of extracellular peptidases suggests that *Ostreobium* may supplement its energy needs and compensate for its lysine auxotrophy by feeding on the organic skeletal matrix, and a diverse set of fermentation pathways allow it to live in the anoxic skeleton at night. *Ostreobium* depends on other holobiont members for biotin and vitamin B12, and our metatranscriptomes identify potential bacterial sources. Metatranscriptomes showed *Ostreobium* becoming a dominant agent of photosynthesis in bleached corals and provided evidence for variable responses among coral samples and different *Ostreobium* genotypes. Our work provides a comprehensive understanding of the adaptations of *Ostreobium* to its extreme environment and an important genomic resource to improve our comprehension of coral holobiont resilience, bleaching and recovery.

Introduction

Coral health depends on the harmonious association between the coral animal and its microbial associates, together known as the holobiont. A wealth of studies shows that climate change threatens coral health by disrupting the association between the coral and its
photosynthetic dinoflagellate endosymbionts (Symbiodiniaceae), culminating in coral bleaching and death. The role of other microbes in coral resilience is just starting to be understood, and our current knowledge is largely based on correlations between coral health status and the presence of microbial taxa inferred from metabarcoding. Whole-genome sequences of corals and their symbionts are an invaluable resource to obtain a mechanistic understanding of the functioning and resilience of the holobiont. Recent genomic studies have shown that the coral host, dinoflagellate symbionts, and prokaryotes living in the coral tissue have complementary pathways for nutrient exchange, highlighting the interdependence and possible co-evolution between these members of the coral holobiont. To date, most work has focused on the coral animal and microbiota associated with its living tissue, with very little work done on the highly biodiverse and functionally important microbiota inhabiting the skeleton of the coral.

The green alga Ostreobium sp. is an important eukaryotic symbiont living inside the coral skeleton (Fig. 1a). This endolithic alga is the principal agent of coral reef bioerosion, burrowing through the limestone skeleton of corals and other reef structures and dissolving up to a kilogram of reef CaCO$_3$ per m$^2$ per year. These algae bloom in the skeleton when corals bleach (Fig. 1b), boosted by the extra light, and – hypothetically – the extra nitrogen and CO$_2$ that may reach the skeleton in the absence of Symbiodiniaceae. Part of the carbohydrates produced by Ostreobium photosynthesis make their way into the coral tissue, extending the time it can survive without its dinoflagellate partner. While these ecological and physiological phenomena have been described, our knowledge of the molecular mechanisms involved is scarce, severely limiting our understanding of key processes in healthy and bleached holobionts. What is the mechanism of skeletal deterioration and how tightly does Ostreobium metabolism integrate with that of the coral and associated microbiota? Developing this knowledge will be essential to understand and manage the roles that skeletal microbiota play during coral bleaching.

Ostreobium has an extreme lifestyle for a green alga. It lives in a very dimly lit environment, with mostly the low-energy far-red wavelengths not used up by Symbiodiniaceae available. The action spectrum of Ostreobium photosynthesis extends into far-red
wavelengths but the underlying molecular mechanisms are little known\textsuperscript{13,14}. The daily rhythm of oxygenic photosynthesis and respiration within the skeletal matrix leads to strong fluctuations of pH and O\textsubscript{2}, ranging from total anoxia at night to ca. 60\% of air saturation during the day\textsuperscript{11,15} and the skeleton limits diffusion of O\textsubscript{2} and other compounds. These stressful conditions are not normally encountered by free-living algae and the mechanisms allowing \textit{Ostreobium} to thrive in this extreme habitat are virtually unknown.

From an evolutionary perspective, \textit{Ostreobium} is a member of the green plant lineage (Viridiplantae), which includes the land plants that originated in the Streptophyta lineage and a broad diversity of algae in the Chlorophyta lineage\textsuperscript{16} (Fig. 1c). \textit{Ostreobium} is in the Bryopsidales, an order of marine algae that has evolved in the Chlorophyta. While \textit{Ostreobium} forms small microscopic filaments, many representatives of this order are larger seaweeds\textsuperscript{17}.

So far, the genomic resources available for coral holobiont research have been limited to the coral animal (7 genomes), its dinoflagellate photosymbionts (14 genomes) and a small fraction of the prokaryotes associated with its tissue (52 metagenome-assembled genomes\textsuperscript{2}). Here, we present the first nuclear genome of an \textit{Ostreobium} species to extend the available genomic toolkit into the coral skeleton, an element of the holobiont crucial for our comprehension of coral resilience, bleaching, and recovery. We expect that these genomic resources will spur new insights into processes of coral bleaching encompassing the entire holobiont, as this knowledge will be essential to safeguard the future of coral reefs in a changing climate. Based on comparative analyses of the \textit{Ostreobium} genome with those of other green algae, we show \textit{Ostreobium}’s innovations in light harvesting antennae and derive insights and hypotheses about its functions as a coral symbiont, and as an alga living in an extreme environment.

\textbf{Results and Discussion}

We obtained the nuclear genome data of \textit{Ostreobium quekettii} (SAG culture collection strain 6.99, non-axenic) by assembling the sequenced reads generated by Illumina and Nanopore
platforms (Supplementary Table 1). This strain was discovered growing on a culture of a small tropical red alga and isolated into culture. The strain is nested in a lineage of *Ostreobium* species found in scleractinian coral\(^\text{18}\) and it readily colonizes coral skeleton when that is provided as a substrate. This clearly shows that the strain is an appropriate representative for this limestone-burrowing coral-associated genus, despite it being initially isolated from a non-coral source. *Ostreobium* is known to produce flagellated spores and we think that a spore residing on the surface of the red alga is the most likely source of the strain.

We assembled the nuclear genome of *Ostreobium quekettii* in 2,997 scaffolds (total 150.4 Mb of assembled bases; N50 length 71,393 bp), with an average sequence coverage of 116x for short-read data. The GC content of the *Ostreobium* nuclear genome is 52.3%, which is higher than 40.4% of *Caulerpa lentillifera* (the closest relative of *Ostreobium* for which a nuclear genome has been sequenced), but less than most other green algae (Supplementary Table 2).

The annotation process resulted in 11,735 predicted protein-coding genes, of which ~52% were supported by RNA sequencing data. We recovered 74.3% of complete core conserved eukaryote genes (based on BUSCO\(^\text{19}\)) in the assembled genome, comparable to the 75.9% in both *C. lentillifera* and *Ulva mutabilis* (Supplementary Fig. 1).

**Photobiology in a dark place**

*Ostreobium* has developed adaptations to the peculiar light environment in which it lives, but its mechanisms are not well understood\(^\text{13,14}\). Our investigation of the light harvesting complex (LHC) proteins shows that *Ostreobium* has the highest number of LHC proteins observed among green algae and *Arabidopsis thaliana* (Fig. 2a, Supplementary Figs 1 and 2).

This expansion of the light harvesting protein arsenal is found in both photosystems, with duplications of the *Lhca1* and *Lhca6* gene families associated with PSI (Supplementary Fig. 2) and the presence of both *Lhcp* and *Lhcb* families associated with PSII (Supplementary Fig. 3).

Most green algae possess a single *Lhca1* gene, while *Ostreobium* has four (Fig. 2); two adjacent genes were found on two different scaffolds, suggesting that this gene family expanded via tandem duplication and translocation (Fig. 2b). The specific amino acid residue as a ligand
for Chl in LCH proteins is essential in determining the chromophore organization, which affects light spectral absorption\textsuperscript{10}. The Lhca1 protein of green plants typically uses histidine as the chlorophyll-binding residue (A5 site in Fig. 2c), but A. thaliana mutants use asparagine at the A5 site, yielding red-shifted absorption spectra\textsuperscript{20,21}. Although most green algae use histidine at the A5 site, the siphonous green algae (Ostreobium, C. lentillifera and Bryopsis corticulans) all have asparagine (Fig. 2c), hinting at their capability to use far-red wavelengths. Lhca6 is located in the outer LHC belt, where it forms heterodimers (with Lhca5), and their long C-terminal loops facilitate interactions between the inner and outer LHC belts\textsuperscript{22,23}. The Lhca6 gene occurs as a single copy in most microalgae and is absent in the seaweeds C. lentillifera and U. mutabilis. Remarkably, we recovered six copies of Lhca6 in Ostreobium (Supplementary Fig. 2). This result is supported by transcriptomic evidence for all Lhca6 genes and different chromosomal contexts (i.e. flanking genes) on the contigs.

In the PSII-associated LHC, Ostreobium possesses an unusual combination of both the Lhcp and major light-harvesting complex (Lhcb) protein families, and has more copies of both genes than do other green algae (Fig. 2a, Supplementary Fig.3). Lhcb is found in most species of the green lineage except prasinophytes that have the Lhcp family instead\textsuperscript{24}. Only the streptophyte Mesostigma viridis is known to encode proteins of both families (Supplementary Fig. 3) suggesting that Lhcp and Lhcb were both part of the LHCII antenna system of the green plant ancestor and the families were differentially lost in different green algal and land plant lineages\textsuperscript{25,26}.

Although Ostreobium shows a high diversity of LHC genes, it lacks many genes for photoprotection and photoreceptors. The non-photochemical quenching (NPQ) genes for LHCSR and PsbS are both absent from Ostreobium and C. lentillifera genomes (Supplementary Fig. 3 and Supplementary Table 3). Energy-dependent quenching (qE) was not observed in other siphonous green algae\textsuperscript{27} and our genomic data provide clear evidence that these algae are not capable of this mechanism. Ostreobium also lacks the light-harvesting complex-like (LIL) genes coding for OHP1 and OHP2 (Supplementary Table 3). While the function of LIL proteins has not been comprehensively determined, their involvement in response to light stress is
known\textsuperscript{28}. The loss of genes involved in high-light sensitivity also extends to the chloroplast
genome of \textit{Ostreobium}, which lacks the chloroplast envelope membrane protein gene (\textit{cemA})
that is needed in \textit{Chlamydomonas} to persist in high-light conditions\textsuperscript{29}.

\textit{Ostreobium} has fewer known photoreceptors than do most other green algae. We found
three blue light photoreceptor genes: a phototropin (not shown), a plant cryptochrome and a
photolyase/blue-light receptor 2 (PHR2; Supplementary Fig. 4). Phototropins are widespread in
green plants\textsuperscript{30}. In \textit{C. reinhardtii}, phototropins are involved in the sexual cycle\textsuperscript{31}, phototactic
response, light-dependent expression of LHC proteins, chlorophyll and carotenoid biosynthesis,
and LHCSR3-based non-photochemical quenching\textsuperscript{32}. However, the function of phototropin in
\textit{Ostreobium} remains unknown, and we did not recover LHCSR in the genome. Surprisingly,
given the predominance of far-red light in coral skeletons, no genes coding for the red/far-red
phytochromes were found in the \textit{Ostreobium} genome. Although phytochromes are widely found
in the green plant lineage, they have been lost in most of the green algal lineage Chlorophyta\textsuperscript{33}.

although not in \textit{Micromonas}, a prasinophyte microalga\textsuperscript{34}. Most green algae appear to have no
specific photoreceptors for red light\textsuperscript{35}. However, \textit{Chlamydomonas} has an animal-like
cryptochrome that can be activated by either red or blue light\textsuperscript{36}. The function of this protein is
related to transcription of genes involved in photosynthesis, pigment biosynthesis, cell cycle
control and circadian clock\textsuperscript{37}. This gene is not found in \textit{Ostreobium}, which also lacks the Cry-
DASH-type cryptochromes observed in many other green algae (Supplementary Fig. 4).

Functional information for Cry-DASH is still scarce\textsuperscript{35}, but this protein is known to be involved
in DNA repair in \textit{Arabidopsis}\textsuperscript{38}. \textit{Ostreobium} does have two genes similar to the \textit{Arabidopsis}
putative blue-light receptor protein called PAS/LOV (Uniprot O64511)\textsuperscript{39}. The UV resistance
locus8 (UVR8), an important photoreceptor to induce UV-B acclimation, is also absent in
\textit{Ostreobium} as well as \textit{C. lentillifera}. We did not find different types of rhodopsin-like
photoreceptors in \textit{Ostreobium}.

Besides the photoprotective LHC genes and photoreceptors, \textit{Ostreobium} and \textit{C. lentillifera} appear to have lost the light-dependent protochlorophyllide oxidoreductase (LPOR),
another important light signaling gene involved in chlorophyll biosynthesis. The reduction of
protochlorophyllide to chlorophyllide can be catalysed by either of two non-homologous enzymes: the nuclear encoded LPOR and the plastid-encoded light-independent (DPOR) protochlorophyllide oxidoreductase\textsuperscript{40}. Most green algae have both systems, and DPOR has been lost in many eukaryotes\textsuperscript{41}. The Ostreobium genome provides the first evidence for the loss of LPOR in any eukaryote, and a screening of C. lentillifera also came back negative, suggesting that the loss may have occurred in the common ancestor of Bryopsidales. Both genera encode DPOR in their chloroplast genomes\textsuperscript{29,42}. Although both enzymes catalyse the same reaction, they have different features: DPOR is encoded, synthesized and active in the plastid, and is highly sensitive to oxygen\textsuperscript{43}, while the nucleus-encoded LPOR is synthesized in the cytosol and active in the plastid, and requires light to be activated\textsuperscript{44}. DPOR might be an advantage over LPOR for endolithic photosynthetic organisms because of the low-light, low-oxygen environments they inhabit\textsuperscript{4}.

The Ostreobium genome clearly reflects its evolutionary trajectory into a peculiar light habitat, with an unparalleled arsenal of LHC proteins but few known mechanisms to sense the light or defend itself against excessive light. Some of these genome features are shared with other Bryopsidales, including the loss of LPOR, qE-type non-photochemical quenching and the UV resistance locus 8. This suggests that the common ancestor of Bryopsidales may have been a low-light-adapted organism, possibly an endolithic alga like Ostreobium is now, a hypothesis supported by Ostreobium being the sister lineage of all other Bryopsidales\textsuperscript{45} and other bryopsidalean lineages also containing old, largely endolithic families\textsuperscript{46}. Bryopsidales originated in the Neoproterozoic with Ostreobium diverging in the early Paleozoic\textsuperscript{17,47}. One could speculate that the bryopsidalean ancestor inhabited a low-light environment, possibly on the dimly lit seafloor beneath Cryogenian ice sheets. The different lineages emerging from this ancestor could then have followed different evolutionary trajectories during the onset of Paleozoic grazing, with the Ostreobium lineage fully committing to an endolithic lifestyle while other bryopsidalean lineages engaged in an evolutionary arms race with grazers to form larger and chemically defended macroalgae.
The unusually large arsenal of LHC proteins appears to be confined to the *Ostreobium* lineage. The genus is present in a diverse range of light environments, from old oyster shells in the intertidal, where it experiences full spectrum sunlight, to coral skeleton (far-red enriched) and in mesophotic habitats (blue light), hinting at its capability to harvest energy from photons of many different wavelengths \(^48\). While from the genome data alone, we cannot derive that the different LHC proteins convey a capability for photosynthesis at different wavelengths, it is known that different molecular arrangement of LHC proteins and the specific pigments binding to them have an impact on their spectral properties \(^49\), which could help *Ostreobium* acclimate to different light environments.

**Life in an extreme environment**

The endolithic environment, and the coral skeleton, is an extreme environment in many ways. Oxygen levels vary strongly, from high concentrations caused by photosynthesis during the day to complete anoxia due to respiration during the night, and this trend is mirrored in strong diurnal pH fluctuations \(^4,15\). Reactive oxygen species (ROS) can be produced in high quantities in these conditions, particularly in the morning when photosynthesis starts \(^50\).

The *Ostreobium* genome exhibits a strong genetic capacity for oxidative stress response, with ROS scavenging and neutralizing genes present in large numbers compared to other green algae (Fig. 3). We found six genes coding for superoxide dismutases (SOD; Fig. 3 and Supplementary Fig. 5), metal-containing enzymes that form the first line of defence to scavenging ROS by catalysing the very reactive superoxide radical (\(O_2^-\)) into hydrogen peroxide (\(H_2O_2\)) \(^51\). All three categories of SOD are present in *Ostreobium*.

Hydrogen peroxide is commonly processed by catalase (CAT), of which *Ostreobium* has eight copies while most other green algae have one or two (Fig. 3), or none in the case of prasinophytes. Seven of the *Ostreobium* catalases formed a unique lineage in our phylogenetic analysis (Supplementary Fig. 5) and are found in tandem on three different scaffolds, indicating diversification of this gene in the *Ostreobium* lineage. Hydrogen peroxide can also be processed through the glutathione-ascorbate cycle, a metabolic pathway neutralizing hydrogen peroxide.
through successive oxidations and reductions of ascorbate, glutathione, and NADPH\textsuperscript{51} (Fig. 3). \textit{Ostreobium} featured high copy numbers of the enzymes to quickly purify ascorbate, with eight copies of ascorbate peroxidase (APX) and seven monodehydroascorbate reductases (MDHAR) (Supplementary Fig. 5). The glutathione-ascorbate cycle is important to keep ascorbate free for \(\text{H}_2\text{O}_2\) scavenging (Fig. 3).

Most algae live in environments with higher oxygen concentrations and can produce energy via the respiratory electron transport chain. In the coral skeleton, however, any oxygen produced through photosynthesis is completely consumed by respiration within an hour of the onset of darkness\textsuperscript{15}, and the environment is anoxic for long periods of time, hence we expect \textit{Ostreobium} to use other electron acceptors to produce ATP and re-oxidize NAD(P)H and FADH\textsubscript{2}. \textit{Ostreobium} possesses the enzymes required to produce succinate, lactate, formate, acetate, ethanol, alanine, and glycerol (Supplementary Fig. 6), but lacks \(\text{H}_2\) and acetate production from acetyl-CoA (PAT1/PAT2 and ACK1/ACK2). Several fermentation-related genes are present in multiple copies (Supplementary Table 4), including two lactate dehydrogenases, four tandem copies of ALDH (aldehyde dehydrogenase), and eight copies of malate dehydrogenase, implying that acetate, succinate and lactate are the preferred fermentation products in \textit{Ostreobium}. We did not find the gene for pyruvate decarboxylase, which catalyses the production of acetaldehyde, the substrate of ALDH. This gene is also absent in \textit{C. lentillifera}, \textit{U. mutabilis} and prasinophytes, indicating it has been lost multiple times, and it is reasonable to assume that another yet uncharacterised enzyme may catalyse this reaction in these algae.

A comparison of the number of genes having particular InterPro annotations between \textit{Ostreobium} and \textit{Caulerpa} showed a large number of depleted IPR terms in \textit{Ostreobium}, which can be attributed to an enrichment in the \textit{Caulerpa} genome, as counts in \textit{Ostreobium} are comparable to those of other green algae. However, genes coding peptidases were strongly enriched in \textit{Ostreobium} (e.g. Peptidase S1, PA clan – 179 genes; Serine-proteases trypsin domain – 142 genes; Supplementary Fig. 7). About 40% of the genes are predicted to be secreted (signal peptide), and 25% are membrane bound, compared to 31% and 12% (of total 57
genes) respectively in *C. lentillifera*, suggesting an expanded potential for external protein degradation in *Ostreobium*. It is known that *Ostreobium* can penetrate the organic-inorganic composite material of the black pearl oyster nacreous layer, and the combination of proteolytic enzyme proliferation and its growth at excessively low light intensities lends support to the hypothesis that this alga could complement its energy needs by feeding on the organic matrix of the coral skeleton and shells.

**The coral holobiont**

*Ostreobium* plays a number of important roles in the coral holobiont, particularly during periods of coral bleaching, but current knowledge is far from complete and the genome can help define hypotheses of how the species may interact with other holobiont members.

**Molecular mechanism of CaCO₃ dissolution**

*Ostreobium* is the main agent of microbial bioerosion of the coral skeleton. While the molecular mechanism behind this phenomenon is not yet known, the *Ostreobium* genome reveals the molecular toolkit available for this process and allows us to make conjectures about how bioerosion by *Ostreobium* might occur (Fig. 4a). A working model for microbial carbonate excavation was first described for cyanobacteria, involving passive uptake of Ca²⁺ at the boring front, decreasing the ion concentration in the extracellular area below calcite saturation levels and leading to dissolution of adjacent calcium carbonate. Imported Ca²⁺ is transported along the cyanobacterial filament and excreted away from the growing tip, likely by P-type Ca²⁺-ATPases that pair transport of Ca²⁺ with counter transport of protons. Whereas the Ca²⁺-ATPase mechanism appears to be somewhat widespread among endolithic cyanobacteria, it is unclear whether this mechanism extends to microboring algae like *Ostreobium*.

*Ostreobium* has an expanded repertoire (33 genes) of calcium transporters (Supplementary Table 5), including 18 voltage-dependent calcium channels and several transient receptor potential transporters, two-pore channels and calcium-transporting ATPases. Calcium uptake, possibly combined with acidification via a Ca²⁺/H⁺ transporter would promote...
decalcification (Fig. 4a), allowing Ostreobium to burrow into the coral skeleton. Calcium
toxicity can be avoided either by accumulation in a vacuole and/or posterior transport out of the
cell. In addition to calcium transport, bicarbonate uptake could also play a role in the burrowing
mechanism. Ostreobium carries two orthologs of the CIA8 transporter responsible for
bicarbonate transport in C. reinhardtii 56, and the imported bicarbonate may be further
transported into the chloroplast to be fixed. Ostreobium has a carbonic anhydrase predicted to
be targeted outside the cell that may further assist the decalcification process. In natural
communities dominated by Ostreobium, CaCO3 dissolution was higher at night, suggesting that
Ostreobium takes advantage of the lower external pH at night 53.

Whilst the Ostreobium genome does not provide confirmation of a specific mechanism
for CaCO3 dissolution, it lends support to a growing hypothesis of how bioerosion might work
in Ostreobium, likely bearing similarities to the process described in cyanobacteria. A detailed
characterisation of this dissolution process should be a priority, as Ostreobium is responsible for
ca. 30-90% of microbial dissolution of skeletal CaCO3. Higher temperatures and lower pH boost
this activity, suggesting that this yet unknown process will lead to major reef deterioration in
future ocean conditions 53.

Interactions with other holobiont members

As far as holobiont functioning goes, interactions between the coral animal and its algal
and bacterial symbionts are best understood, and interactions between the Symbiodiniaceae and
bacteria are just starting to come into focus 57, but little is known about interactions involving
Ostreobium. It is well known that several algae lack genes related to biosynthetic pathways of,
and are auxotrophic for, certain nutrients. The Ostreobium genome shows that the metabolic
pathways involved in the production of vitamins B1, B2, B6 and B9 are complete, but it lacks
the BioA and LysA genes involved in biotin and lysine synthesis. These genes were found in all
other green algae used for comparison, suggesting that this gene loss is specific to Ostreobium.
It is likely that the alga obtains these nutrients from bacteria, and the dependency on external
lysine may also explain the enrichment of extracellular proteases encoded in the Ostreobium
Many algae are auxotrophic for vitamin B12 (cobalamin), a cofactor involved in the synthesis of methionine, instead obtaining this compound from associated bacteria. The *Ostreobium* genome encodes two gene copies of a B12-independent methionine synthase (METE) in addition to the B12-dependent version (METH), so while the alga most likely does not strictly require B12 for growth, the presence of METH suggests that it uses B12 provided by other holobiont members. Corals are also auxotrophic for several vitamins and amino acids that are produced by holobiont members.

The nature of metabolic exchanges between *Ostreobium* and the coral animal is an open question in holobiont research, and a potentially critical one, as coral bleaching and subsequent blooming of the endolithic *Ostreobium* algae become increasingly common due to ocean warming. Typical algal-animal metabolic exchanges include nitrogen and CO$_2$ provision by the animal to the alga and carbohydrate provision to the animal by the alga. Coral polyps are known to secrete nitrogen in the form of ammonia. An expanded repertoire of ammonia transporters was identified in *Ostreobium* (Supplementary Table 5), potentially reflecting an adaptation to increase and diversify ammonia uptake in the alga. This observation is also in line with the presence of diazotrophic bacteria facilitating the conversion of N$_2$ into ammonia in marine limestones (including coral skeletons) and ammonia being the most abundant form of inorganic nitrogen in skeletal pore waters, and adds to the evidence for the roles that endolithic organisms play in the holobiont N cycle. During carbon cycling in the holobiont, glucose has been postulated as one of the main carbohydrates exchanged between Symbiodiniaceae and corals. While carbon compounds fixed by endoliths are known to be transferred to the coral animal and subsequently assimilated, neither the exact transferred molecules nor the molecular mechanisms involved in their translocation have been characterised, but the *Ostreobium* genome encodes two genes coding for H$^+$-glucose transporters that might be involved in this process.
Probing changes in holobiont processes

While these investigations of the *Ostreobium* genome allowed us to evaluate hypotheses about interactions within the holobiont, gaining deeper insight will require approaches that study multiple partners simultaneously. As a first step towards understanding the molecular mechanisms in the coral holobiont during coral bleaching, we screened transcriptomes of healthy and bleached coral holobionts. We exposed fragments of the coral *Orcibella faveolata* to elevated temperatures leading to bleaching, followed by re-acclimation of the bleached samples to ambient temperatures (post-bleaching condition), during which the corals remained bleached and an *Ostreobium* bloom occurred. Total metatranscriptomes (including coral tissue and skeleton) were generated for healthy control samples that were kept under ambient temperature (no bleaching), and the bleached samples.

From the KEGG annotation of the assembled transcriptome, it is clear that the number of transcripts coding for the *Ostreobium* photosynthetic apparatus exceeds that of Symbiodiniaceae (Supplementary Fig. 8), in line with the expanded set of photosynthetic genes (particularly LHC) observed in the *Ostreobium* genome. While the metatranscriptome sequencing depth was not designed to track up- and down-regulation of individual genes, it was clear that expression of the photosynthesis genes *psbA* and *rbcL* from Symbiodiniaceae was drastically lower in bleached samples while expression of these genes in *Ostreobium* increased (Supplementary Fig. 9a), supporting the notion that *Ostreobium* becomes a dominant agent of photosynthesis in the bleached holobiont. The bleached state increases the light available to *Ostreobium*, which likely leads to the need for higher repair rates of PSII protein D1 (encoded by *psbA*) and higher rates of carbon fixation (facilitated by RuBisCO, encoded by *rbcL*). The expression of these genes by Symbiodiniaceae in the bleached samples, albeit at low levels, suggests that some of these endosymbionts have remained or that some re-colonisation has occurred during the post-bleaching period.

We detected multiple haplotypes of *psbA* (Supplementary Fig. 9b), indicating that several strains of *Ostreobium* were present in the *O. faveolata* skeletons. While expression levels for
these haplotypes tended to increase, some did not change significantly while others differed by
an order of magnitude or more. These differences suggest that Ostreobium strains may differ
physiologically or change in relative abundance during the experiment. Such differences in the
microbiome – whether Ostreobium strains or other holobiont members – may result in high
variability among coral samples in experimental work. Indeed, we found considerable
differences in expression levels between samples within conditions, suggesting that future
metatranscriptome experiments should be planned to use generous replication.
The metatranscriptomes also provide some hints as to where Ostreobium may source its needs
for vitamin B12 and biotin. We detected 35 bacterial transcripts from the vitamin B12 pathway,
including several Proteobacteria (e.g. Rhodobacteraceae), Bacteroidetes (e.g. Flavobacteriaceae)
and Cyanobacteria that are known to be abundant in the coral skeleton. Seven bacterial
transcripts associated with the biotin pathway were found, with taxonomic affinities to
Actinobacteria, Cyanobacteria and Proteobacteria (e.g. Endozoicomonadaceae,
Rhodospirillaceae). This result narrows down the potential mutualistic relationships between
bacteria and Ostreobium in the holobiont (Fig. 4b).

Conclusion

The complexity of the coral holobiont presents an interesting challenge to reconstruct a
comprehensive model of metabolic exchanges and other interactions among its component
organisms. Recent progress in building such models from genomes of coral and tissue-
associated prokaryotes has not been mirrored in the skeleton. Our results allowed refining
hypotheses from previous physiological work, illuminating the biology of skeletal inhabitants
and their interactions. Despite this progress, many questions remain about the exact mechanisms
at play, their immediate physiological effects on the partners and longer-term ecological
consequences. One challenge in this area of research is that the roles and impacts of Ostreobium
vary in time. In terms of photosynthesis, Ostreobium makes a minor contribution to holobiont
photosynthesis in healthy corals, but our results show that Ostreobium blooms during
bleaching cause a shift of expression levels between microbiome members. Specifically, 

*Ostreobium* takes over the major role of photosynthesis from the Symbiodiniaceae now lost 

from the holobiont. These changes likely have implications flowing through the entire 

microbiome interaction network, but our knowledge of this is in its infancy. From a 

physiological viewpoint, the question remains as to what extent the *Ostreobium* bloom is simply 

an opportunistic response, and whether it is actually leading to a beneficial metabolite exchange 

with the coral host. Similarly, the fitness costs/benefits of *Ostreobium* in the skeleton need to be 

evaluated across the entire coral life cycle and might differ drastically between the period of 

eyear skeleton formation following larval settlement and at later stages. The fact that there are 

>80 different species-level operational taxonomic units in *Ostreobium*, and virtually nothing is 

known their physiological features, complicates the matter further. Genome-scale data from 

*Ostreobium* and other interacting partners critically link these physiological features to 

underlying molecular mechanisms. The results and datasets generated in this study provide a 

foundational reference for future research into the biology of this key holobiont member, and 

the intricate role that *Ostreobium* plays in coral biology. This will be of particular importance in 

the light of global climate change, as the increased frequency of bleaching and lower pH will 

boost *Ostreobium* populations and carbonate dissolution rates.

**Methods**

**Culturing and nucleic acid extraction**

*Ostreobium quekettii* (SAG culture collection strain 6.99, non-axenic) was cultured in F/2 

media on a 14H - 8H light/dark cycle at ~19 degrees celsius. Total DNA was extracted using a 

modified cetyl trimethylammonium bromide (CTAB) method described in Cremen and 

collaborators. Total RNA was extracted using Plant RNA reagent (Thermofisher, Waltham, 

MA, USA).
Illumina sequencing and assembly

Total DNA was sequenced using Illumina sequencing technology (HiSeq2000, 150 bp paired-end reads, ~40 GB data) (Supplementary Table 1), at Novogene, Beijing. Reads were trimmed with Cutadapt v1.12 using the parameters -e 0.1 -q 10 -O 1 -a AGATCGGAAGAGC. To remove pIDNA, mtDNA and bacterial sequences prior to assembly all reads were mapped against the previously published *Ostreobium* chloroplast and mitochondrial sequences and a database of bacterial genomes, using BWA with the bwa-mem algorithm at default settings. The bacterial database was constructed from 5460 prokaryotic representative genomes (https://www.ncbi.nlm.nih.gov/refseq/about/prokaryotes/#representative_genomes) downloaded from NCBI using the tool ncbi-genome-download (https://github.com/kblin/ncbi-genome-download). If one or both pair-end reads mapped to any database sequence, both reads were removed from the full Illumina dataset using the BBtools `filterbyname.sh` program (https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/). The resulting dataset was assembled using Spades v3.12.0 with the --careful option. Contigs produced by Spades were further filtered to identify any remaining bacterial or fungal sequences by BLAST searches of open reading frames against the nr database.

For RNA sequencing, a strand-specific 100 bp paired-end library was constructed and sequenced using Illumina HiSeq 2500 (ENA study accession number PRJEB35267). Quality filtering of reads was performed using Trimmomatic v0.39 with the following settings: LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20. Transcriptome data were assembled using Trinity v2.8.3 in both genome-guided mode and de novo mode.

Nanopore sequencing and hybrid assembly

Total DNA was extracted as above and sequenced using nanopore sequencing technology (MinION, Oxford Nanopore Technologies), producing 1,889,814 reads and ~9.5 GB data. To identify sequences originating from bacteria, all reads were used as blastn queries (megablast
algorithm) against the prokaryotic representative genome dataset. Reads with a hit alignment
>1000 bp long and an e-value < 1e-100 were removed from the nanopore dataset.

A hybrid (combined long-read and short-read) genome assembly was performed with
MaSuRCA v3.2.7 \(^70\), using the filtered Illumina reads as short-read input. For long-read input,
the nanopore reads were complemented with the contigs produced from the Spades assembly of
Illumina data and treated as ‘pseudo-nanopore’ reads.

### Ab initio prediction of protein-coding genes

We adapted the workflow from \(^71\) for ab initio prediction of protein-coding genes in the
Ostreobium genome assembly. Repetitive elements in the genome assembly were first predicted
de novo using RepeatModeler v1.0.11 (http://www.repeatmasker.org/RepeatModeler/). These
repeats were combined with known repeats in the RepeatMasker database (release 20171107) to
generate a customized repeat library. All repetitive elements in the assembled genome scaffolds
were then masked using RepeatMasker 4.0.7 (http://www.repeatmasker.org/) based on the
customized repeat library, before they were subjected to prediction of genes.

The assembled transcripts were used as transcriptome evidence, and vector sequences
were removed using SeqClean \(^72\) based on UniVec (build 10) database. PASA pipeline v2.3.3 \(^73\)
and TransDecoder v5.2.0 \(^74\) were first used to predict protein-coding genes (and the associated
protein sequences) from the vector-trimmed transcriptome assemblies (hereinafter transcript-
based genes). The predicted protein sequences from multi-exon transcript-based genes with
complete 5’ and 3’-ends were searched (BLASTp, E ≤ 10-20) against RefSeq proteins (release
88). Genes with significant BLASTp hits (>80% query coverage and >80% subject coverage)
were retained. Transposable elements were identified using HHblits \(^75\) and Transposon-PSI
(http://transposonpsi.sourceforge.net/), searching against the JAMg transposon database
(https://github.com/genomecuration/JAMg). Proteins putatively identified as transposable
elements were removed. Those remaining were clustered using CD-HITS (ID=75%) \(^76\) to yield a
non-redundant protein set, and the associated transcript-based genes were kept. These genes
were further processed by the Prepare_golden_genes_for_predictors.pl script from the JAMg
package (https://github.com/genomicuration/JAMg). This step yielded a set of high-quality
“golden” genes, which were used as a training set for gene prediction using AUGUSTUS and
SNAP. We also employed GeneMark-ES v4.38 to generate prediction from the genome
scaffolds, and MAKER protein2genome v2.31.10 to make predictions based on homology
evidence to SwissProt proteins (downloaded 27 June 2018).

Subsequently, all genes predicted using GeneMark-ES, MAKER, PASA, SNAP and
AUGUSTUS were integrated into a combined set using EvidenceModeler v1.1.1, following a
weighting scheme of GeneMark-ES 2, MAKER 8, PASA 10, SNAP 2, AUGUSTUS 6. The
resulting genes from the EViidenceModeler prediction were retained if they were constructed
using evidence from PASA, or using two or more other prediction methods.

Functional information of translated predicted genes was retrieved using search BLASTP
against UniProt databases (Swiss-Prot and TrEMBL), KEGG’s annotation tool BlastKOALA.
Gene models of the genome dataset were annotated using InterProScan 5.39 using InterPro
(version 77.0 databases) of Pfam (32.0), SUPERFAMILY (1.75) and TIGRFAMs (15.0). The

**Orthogroups, phylogenetic analysis and BUSCO analysis**

For comparative genomic analyses, we built a dataset containing genomes and
annotations of 20 green algae and two land plants (Supplementary Table 4). We used the
OrthoFinder 2.3.7 pipeline (default parameters) to cluster the potential orthologous protein
families. Enrichment and depletion of domains in *Ostreobium* versus *Caulerpa lentillifera* were
identified using Fisher’s exact tests with a false discovery rate correction (Benjamini-Hochberg
FDR method) of 0.05. All statistical tests were carried out in R. Subcellular localization of
sequences of interest was performed using PSORT, using the WoLF PSORT web server
(https://wolfpsort.hgc.jp/) with the ‘plant’ option, and PredAlgo, using default settings.

We performed phylogenetic analyses for protein families relevant for photobiology and
oxidative stress response, based on the orthogroups obtained from the OrthoFinder analysis
described above. All phylogenetic trees were reconstructed using IQTREE 1.6.12 with the built-
in model selection function, and branch support estimated using ultrafast bootstrap with 1,000 bootstrap replicates\(^{89}\). To identify the light-harvesting protein families associated with PSI (Lhca), we also included Lhca protein sequences from \textit{Bryopsis corticulans}\(^{23}\) and \textit{Chlamydomonas reinhardtii}\(^{22}\).

We ran BUSCO v3.1.0\(^{19}\) on the \textit{Ostreobium} genome using the Eukaryota dataset. This analysis identifies complete, duplicated, fragmented and missing genes that are expected to be present in a set of single-copy genes in the dataset. We ran identical BUSCO analyses on the \textit{C. lentillifera, Ulva mutabilis} and \textit{C. reinhardtii} genomes to allow direct comparison.

**Metatranscriptome analysis of healthy and bleached corals**

**Experiment**

\textit{Obericella faveolata} fragments (4cm\(^2\)) from three different colonies were collected in Petempiche Puerto Morelos, Quintana Roo, Mexico (N 20° 54’17.0”, W 86° 50’11.9”) at a depth between eight and nine meters in June 2013 (Permit registration MX-HR-010-MEX, Folio 036). The coral skeleton and live tissue were collected using a hammer and chisel and were transported in seawater to perform the experiment at the Instituto de Ciencias del Mar y Limnología, UNAM. Three fragments were placed in each control and “experimental” tank at ~28°C. Following 18 days of acclimation, heaters were turned on in the treatment tank to reach ~32°C. After 5 days of severe heat exposure, bleached corals were moved back into the control tank to recover at 28°C, where they remained for 38 days (post-bleaching period). Post-bleached and controlled coral fragments were flash frozen and preserved in liquid nitrogen.

**RNA library preparation and sequencing and data availability**

Coral fragments were ground to a fine powder in liquid nitrogen. Total RNA was extracted using the mirVana miRNA Isolation Kit (Life Technologies) since this kit resulted in higher coral holobiont RNA yield and quality. RNA was purified and concentrated using the RNA Clean and Concentrator kit (Zymo Research, Irvine, USA). RNA quantification was...
assessed on a Nanodrop and Qubit using 2.0 RNA Broad Range Assay Kit (Invitrogen). Quality was verified using the Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA). Total RNA samples were sent for metatranscriptome sequencing to the US Department of Energy’s Joint Genome Institute (JGI), California. Samples were depleted in ribosomal RNA and enriched in mRNA from whole holobionts employing the RiboZero kits (Epicenter). Following the RiboZero protocol, mRNA was converted to cDNA and amplified. The libraries were sequenced on the Illumina HiSeq 2000 platform using 2 X 151 bp overlapping paired-end reads. Raw and filtered metatranscriptome sequence data, statistics and quality sequencing reports for the experiment are available at the US Department of Energy Joint Genome Institute (JGI)'s genome portal (https://genome.jgi.doe.gov/portal/) with accession codes 1086604, 1086606, 1086608, 1086610, 1086612, 1086614, Community Sequencing Project No. 1622).

Metatranscriptome analysis

The raw reads were quality-trimmed to Q10, adapter-trimmed and filtered for process artifacts using BBDeuk^{90}. Ribosomal RNA reads were removed by mapping against a trimmed version of the Silva database using BBmap (http://sourceforge.net/projects/bbmap). To generate a de novo reference metatranscriptome, cleaned reads from all samples per species (replicates from control and treatment fragments) were pooled and assembled using Trinity version 2.1.1^{74}. We performed a BLASTn search to identify and separate the different members of the holobiont using a local database. This database was built with the Ostreobium genome, the sea anemones Nematostella vectensis^{91} and Exaiptasia diaphana (syn. Aiptasia pallida)^{92}, the corals Acropora digitifera^{93} and Orbicella faveolata^{94}, and the Symbiodiniaceae genomes Breviolum minutum^{95}, Fugacium kawagutii^{9} and Symbiodinium microadriaticum^{96}, as well as several Breviolum spp. transcriptomes^{97}.

Kallisto^{98} was used to perform a pseudoalignment and quantify transcript abundances, using the Ostreobium contigs derived from the metatranscriptome assembly as a reference. A comparison of counts per million, a correlation matrix and principal component analysis among samples was performed for a quality check of the replicates per species. Expression analysis
was performed using the DESeq2 software\textsuperscript{99}. Differential Expressed Genes (DEGs) were defined by using a cutoff threshold of False Discovery Rate FDR <0.001 and log fold change of 2. Enzyme Commission numbers (EC) were retrieved from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using MEGAN5\textsuperscript{100} for genes from each holobiont member (i.e. host, Symbiodiniaceae and \textit{Ostreobium}).

**Acknowledgements**

Funding was provided by the Australian Research Council (FT110100585 to HV, DP150100705 to HV and CXC, DP200101613 to HV and MM), the University of Melbourne (CBRI to HV & KAH), the DoE Joint Genome Institute (CSP grant 1622 to MM and VAM) and the National Science Foundation (OCE 1442206 and IOS 0644438 to MM), Pennsylvania State University (to MM), the Canon Foundation (to MM) and CONACyT (216837 to VAM). Computational resources were provided through the Nectar Research Cloud. We thank Roberto Iglesias-Prieto, Claudia Tatiana Galindo and Michele Weber for assisting with field experiments and John Beardall, J. Clark Lagarias and Andrew H. Knoll for discussions. We thank Alexander Fordyce for the bleached coral photograph.

**References**


**References**


Figures

**Figure 1.** (a) Cross-section of *Paragoniastrea australensis* coral showing *Ostreobium* that inhabits the skeleton. The inset shows *Ostreobium* filaments after skeletal decalcification. (b) Bleached coral with evident *Ostreobium* bloom (indicated by white arrows). Photograph by Alexander Fordyce. (c) Phylogenetic tree of Viridiplantae (Streptophyta + Chlorophyta) showing the position of *Ostreobium quekettii*. 
Figure 2. (a) Number of different light harvesting complex (LHC) proteins in *Ostreobium*, other green algae and *Arabidopsis thaliana*. (b) Synteny of *Lhca1* copies in the *Ostreobium* genome. (c) Amino acid sequence comparison between *Lhca1* proteins, showing asparagine (N) at the chlorophyll-binding residue A5 in *Ostreobium*. 
Figure 3. Simplified oxidative response pathway comparing the number of genes for enzymes found in the genomes from Ostreobium and some other green algae. Compared with C. lentillifera, U. mutabilis and C. reinhardtii, Ostreobium does have more copies of genes related to quick response to neutralize ROS, such as superoxide dismutase (SOD), catalase (CAT), ascorbate peroxidase (APX) and monodehydroascorbate reductase (MDHAR). DHAR, dehydroascorbate reductase. GR, glutathione reductase.
Figure 4. Roles of Ostreobium in the coral holobiont. (a) Potential mechanisms available to Ostreobium for excavation of the CaCO₃ skeleton of corals near the growing tip of Ostreobium filament. (b) Possible interactions between members of the holobiont derived from genome sequence data.