Title: Whole-genome duplication in an algal symbiont serendipitously confers thermal tolerance to corals

Authors:

Katherine E. Dougan[1,2]*, Anthony J. Bellantuono[3], Tim Kahlke[4], Raffaela M. Abbriano[4], Yibi Chen[1], Sarah Shah[1], Camila Granados-Cifuentes[2], Madeleine J. H. van Oppen[5,6], Debashish Bhattacharya[7], David J. Suggett[4], Cheong Xin Chan[1]*, Mauricio Rodriguez-Lanetty[2]*

Affiliations:

[1]The University of Queensland, School of Chemistry and Molecular Biosciences, Australian Centre for Ecogenomics, Brisbane, QLD 4072, Australia

[2]Florida International University, Department of Biological Sciences, Miami, FL 33099, U.S.A.

[3]Florida International University, Department of Biological Sciences, Biomolecular Science Institute, Miami, FL 33099, U.S.A.

[4]University of Technology Sydney, Climate Change Cluster, NSW 2007, Australia

[5]School of Biosciences, The University of Melbourne, Parkville, VIC 3010, Australia

[6]Australian Institute of Marine Science, Townsville, QLD 4810, Australia

[7]Rutgers University, Department of Biochemistry and Microbiology, New Brunswick, NJ 08901, U.S.A.

*Correspondence: k.dougan@uq.edu.au (K. E. Dougan), c.chan1@uq.edu.au (C. X. Chan), rodmauri@fiu.edu (M. Rodriguez-Lanetty)

Abstract:

Coral reefs are fundamentally sustained by symbioses involving dinoflagellate algae in the Family Symbiodiniaceae. The coral symbiont *Durusdinium trenchii* is notable for enhancing the resilience of coral holobionts under thermal stress. Believed to have experienced whole-genome duplication (WGD), *D. trenchii* offers a valuable model system to understand how selection acts on the genome of a facultative symbiont after WGD. We present genome assemblies for two isolates of *D. trenchii* and confirm WGD in these taxa, providing the first example of this phenomenon in a single-celled eukaryotic symbiont. We assess how the facultative lifestyle has contributed to the retention and divergence of duplicated genes, and how these results intersect with the observed thermotolerance of corals hosting *D. trenchii* symbionts. Our findings reveal that the free-living lifestyle is the main driver of post-WGD evolution, however, they also implicate symbiosis in this process, with both lifestyles increasing algal fitness. Our results demonstrate that WGD, driven by selection in the free-living phase, has converted *D. trenchii* into a coral symbiont that serendipitously provides increased thermal stress protection to the host coral.

Main Text:

Uncovering the foundations of biotic interactions, particularly symbiosis, remains a central goal for research given that virtually no organism lives in isolation. Coral reefs are marine biodiversity hotspots that are founded upon symbioses involving dinoflagellate algae in the Family Symbiodiniaceae (*1*). These symbionts are the solar power plants of reefs, providing photosynthetically fixed carbon and other metabolites to the coral holobiont (*2, 3*). Breakdown of the coral-dinoflagellate symbiosis (i.e. coral bleaching), often due to ocean warming, puts corals at risk of starvation, disease, and eventual death. Symbiodiniaceae microalgae are diverse with at least 15 clades and 11 named genera (*1, 4-6*), encompassing a broad spectrum of symbiotic associations and host-specificity. Most of these taxa are facultative symbionts (i.e. they can live freely or in symbiosis), although solely free-living species are also known (*1*). Genomes of Symbiodiniaceae are believed to reflect the diversification and specialization of these taxa to inhabit distinct ecological niches. The genomes of symbionts, due to spatial confinement, are predicted to undergo structural rearrangements, streamlining, and enhanced genetic drift (*7*). This hypothesis is supported by the relatively high level of structural rearrangement, pseudogenization, and duplication in genomes of symbiotic Symbiodiniaceae (*8*).

Whole-genome duplication (WGD) is an evolutionary mechanism for generating functional novelty and genomic innovation (*9, 10*), and can occur within species following errors in meiosis, i.e. via autopolyploidy. Following WGD, the evolutionary trajectory of duplicated sequence regions generally proceeds from large-scale purging, temporary retention and/or divergence, to fixation (*11, 12*); WGD-derived genes (i.e. *ohnologs* (*13, 14*)) that are retained can provide a selective advantage and enhance fitness through increased gene dosage, specialization in function, and/or the acquisition of novel functions (*11, 12*).

WGD has been described in free-living unicellular eukaryotes such as yeast (*15-18*) and the ciliate *Paramecium* (*19, 20*), but not in symbiotic species. Evidence of WGD is absent among Symbiodiniaceae lineages, with the exception of the genus *Durusdinium*, as observed in microsatellite sequence data (*21*). This genus includes the thermotolerant species *Durusdinium trenchii*, a facultative symbiont known to confer heat-tolerance on corals and thus enhance their resilience under thermal stress (*22*). Given its facultative lifestyle (i.e. free-living *versus* symbiotic), *D. trenchii* offers a valuable model system to understand how selection acts on the genome of a symbiont after a WGD event. To this end, we present *de novo* genome assemblies from two isolates of *D. trenchii* and demonstrate WGD in this lineage. Based on gene expression profiles, we assess how the duality of facultative lifestyle has contributed to the fate of ohnologs in these microalgae, and how these results intersect with the observed thermotolerance of corals hosting *D. trenchii* symbionts.

We generated *de novo* genome assemblies from *D. trenchii* CCMP2556 (total length = 1.71 Gb; N50 = 774.26 kb; 29,137 scaffolds) and *D. trenchii* SCF082 (total length = 1.64 Gb; N50 = 398.48 kb; 44,682 scaffolds) using 10X Genomics linked reads (tables S1 and S2). The two genomes are highly similar in terms of marker genes (fig. S1), whole-genome sequence (~98% shared identity; fig. S2 and table S3), and repeat landscapes (fig. S3), yielding ~54,000 protein-coding genes (table S4) with a high extent of data completeness (table S5; see Methods). To assess WGD in *D. trenchii*, we followed González-Pech *et al.* (*8*) to identify collinear gene blocks within each genome (see Methods); these blocks likely arose via segmental duplication and/or WGD. We identified 864 blocks implicating 27,597 (49.46% of the total 55,799) genes in CCMP2556, and 776 blocks implicating 18,209 (34.02% of the total 53,519) genes in SCF082 (table S6). The proportion of genes present in collinear blocks in *D. trenchii* is ~49-fold greater than that in other Symbiodiniaceae and the outgroup polar dinoflagellate *Polarella*, which have not experienced WGD (Fig. 1A). We also observed a high extent of conserved synteny (22,041 CCMP2556 genes syntenic with 21,094 SCF082 genes), with ohnologs predominant in these syntenic blocks (CCMP2556: 15,395 [69.85%]; SCF082: 12,617 [59.31%]) (Fig. 1B and table S6). Using homologous protein sets derived from available whole-genome data, our inference of lineage-specific duplicated genes (see Methods) revealed 7,945 gene duplication events specific to *D. trenchii*, which is an order of magnitude greater than in other Symbiodiniaceae (fig. S4). Whereas the distribution of synonymous substitution sites (*Ks*) lacks the distinct peak (fig. S5) expected in ohnologs, this is not surprising for the relatively recent WGD expected in *D. trenchii* (*23*). The timing of WGD in *D. trenchii*, as observed in other taxa, likely coincides with its split from the sister taxon, *D. glynnii* ~1 million years ago (*1*). These results based on independently assembled genomes from two isolates, combined with the extent and size of the gene blocks (table S7 and fig. S6), provide unambiguous evidence for WGD in *D. trenchii*.

To assess the fate of ohnologs in *D. trenchii*, we focused on CCMP2556 from which transcriptome data exist (*24*) for cells from two lifestyles: free-living in culture or engaged in symbiosis with the anemone *Exaiptasia pallida*, with both under ambient (28°C) and thermal stress (34°C) conditions. We assessed conservation of expression in ohnologs using the gene-expression modules (fig. S7 and table S8) inferred from weighted gene co-expression network analysis (WGCNA). We adopted an integrated approach (fig. S8; see Methods) to classify each ohnolog-pair into distinct evolutionary scenarios based on their expression profiles and correlation to lifestyle and/or to temperature (Fig. 1, C and D, and table S9). Most ohnolog-pairs that correlated to lifestyle exhibit “Conserved” (4,830 of 9,349 [51.67%]) expression profiles: 2,284 (24.43%) and 2,546 (27.23%) correlated to the free-living and symbiotic lifestyles, respectively (Fig. 1C); this indicates strong functional constraints by each lifestyle, likely due to the benefit from increased gene dosage. Other ohnolog-pairs with “Divergent” expression profiles indicate selection (i.e. specialization) based on lifestyle (2,539 of 9,349 [27.16%]). Of the ohnolog-pairs that correlate with temperature (Fig. 1C), very few are “Conserved” (28°C: 79 of 2,759 [2.86%]; 34°C: 70 of 2,759 [2.54%]) and most exhibit “Gain/Loss” of correlation in one ohnolog (28°C: 1,294 [46.90%]; 34°C: 1,107 [40.12%]); this clearly indicates that lifestyle is the main driver of post-WGD evolution.

WGD enables the retention of complete metabolic pathways, which we assessed in both *D. trenchii* isolates following Aury *et al.* (*19*). Of the 98 metabolic pathways retained in duplicate (table S10), specialization driven by lifestyle was detected in central metabolic pathways (figs. S9-S16), such as glycolysis/gluconeogenesis (Fig. 1E). Ohnolog specialization in glycolysis/gluconeogenesis reflects the contrasting functions of this pathway during symbiotic *versus* free-living phases. That is, a high rate of gluconeogenesis, inferred using ohnolog expression data, supplies glucose for translocation to the coral host during symbiosis, whereas a high rate of glycolysis fuels dynamic energetic needs inherent to free-living cells tolerating more-variable environments (*7*). Development of minor or partitioned functionality following WGD has been described in duplicate glycolysis pathways (*25*). In yeast, these pathways diverged and became semi-independent, with each specialized for low and high glucose levels

(*25*); in *D. trenchii*, this might represent a capacity for fine-tuning carbon metabolism to the contrasting energetics of the two lifestyles.

Whereas most ohnolog pairs were expressed at similar levels (fig. S17) or with one ohnolog more highly or dominantly expressed (table S11; see Methods), some have diverged sufficiently to each be dominantly expressed under different scenarios of lifestyle and/or temperature; these pairs represent putative instances of sub-functionalization or neo-functionalization. We identified 90 such divergent ohnolog-pairs in *D. trenchii* (table S12 and fig. S18). Most of these (73 [81.11%]) diverged between the two lifestyles, with many (44 from clusters I-IV and X; fig. S18) exhibiting peak expression when free-living, implicating this lifestyle as the major driver of gene-expression divergence (Fig. 2A and fig. S18). These ohnolog-pairs highlight strong specialization at key nodes in metabolic pathways with broader ohnolog retention and divergence related to nitrogen cycling (including metabolisms of alanine, aspartate, and glutamate; fig. S9) and glutathione metabolism (fig. S10). Notably, this includes a glutamine synthetase (GS; Cluster X; Fig. 2A and fig. S18) that has been connected to rapid symbiotic establishment with hosts by *D. trenchii* (*26*) and an ammonium transporter (Cluster III; fig. S18); both exhibit peak expression in the free-living phase. Along with other transporters among the 90 ohnolog pairs such as a sugar phosphate/phosphate translocator and Na+/dicarboxylate transporter (fig. S18), this enhanced metabolite exchange likely reflects a concerted response of nutrient cycling due to limited sources that are otherwise available during symbiosis. We also observed a similar pattern in mRNA editing based on these data, suggesting the highest functional diversity during the free-living phase (Fig. 2B, fig. S19, and tables S13-14; see Methods).

Our results suggest that the divergence in ohnolog-pair expression is driven by changes in gene regulation and transcript processing (figs. S20 and S21). We observed a greater extent of alternative splicing (AS) in ohnologs compared to non-ohnologs (figs. S22 and S23, and table S15), along with considerable divergence within ohnolog-pairs in both the number and conservation of splice sites (table S16). The increased AS among the ohnologs (table S17 and Figure S24) yielded distinct patterns of differential exon usage (Fig. 2C and fig. S25); we observed asymmetric distributions of exon expression among those in “Gain/Loss” (panel iii) and the 90 divergent ohnolog-pairs (panel iv), compared to singletons (panel i) and those in “Conserved” and “Retained” (panel ii). This asymmetry is observed among ohnologs that have gained and/or retained its specificity to lifestyle within the “Gain/Loss” pairs, and more so among the 90 divergent ohnolog-pairs. This result suggests an accumulation of beneficial exons or purging of superfluous exons, reflecting the lifestyle that drove ohnolog-pair divergence and fixation. Exon restructuring (*27*) and the increase in alternative splicing (*28*) appear to drive ohnolog gene expression divergence in *D. trenchii* vis-à-vis algal lifestyle.

In summary, we demonstrate WGD in a microalgal endosymbiont, and provide strong evidence that lifestyle is the key driver of post-WGD evolution in *D. trenchii*. Given that these algae transition frequently between the free-living and symbiotic lifestyle, we present a hypothetical framework of how this duality drives post-WGD genome evolution (Fig. 3). Under the null hypothesis (i.e. free-living), we expect adaptations to be driven by nutrient availability and fluctuating environmental conditions, whereas under the alternative hypothesis (i.e. symbiosis), we expect adaptations to reflect maintenance of a stable host-symbiont relationship and tightly integrated nutrient/metabolite cycling within the coral holobiont. Whereas our results provide stronger support for the null hypothesis as the driving force behind post-WGD evolution, they implicate both lifestyles in impacting the maintenance and expression divergence of ohnologs. These combined selective forces increase overall fitness in *D. trenchii* and likely explain the high thermotolerance of this species within corals (*29*). Benefits conferred by WGD to a free-living lifestyle in more-variable environments serendipitously primed *D. trenchii* to better assist or exploit the coral holobiont when faced with thermal stress. Whether symbiosis may also have negative effects on fitness post-WGD is unknown (*30*). It should be noted that the dual lifestyle is widespread in Symbiodiniaceae (*1*), yet WGD is not. Therefore, the key feature of *D. trenchii* that we are addressing is not the dual lifestyle, but rather how this trait impacts post-WGD genome evolution and adaptation to the symbiotic *versus* free-living phase. Since the algae propagate to very high densities in coral tissues (0.5–5.0 × 10^6 cells/cm^{-2}) (*31, 32*), the symbiotic lifestyle may also indirectly provide a mechanism for propagation of successful algal genotypes while resident in host tissues. Consequently, these genotypes could re-seed free-living populations upon dissociation from the coral due to colony death, bleaching, or other mechanisms of symbiont population control. The maintenance of multi-gene copies combined with fixed, adaptive changes likely makes *D. trenchii* more capable of metabolic maintenance under dynamic, often stressful environments, and hence a more-resilient symbiont. This may explain the vast geographic and expanded host range for *D. trenchii* (*22*) and its well-known capacity for increasing coral survival under heat waves. Therefore, in an interesting and unexpected twist, WGD, driven by selection under the free-living lifestyle has converted *D. trenchii* into the ideal coral symbiont, able to protect the host coral from thermal stress while increasing its population size during symbiosis.

References:

1. T. C. LaJeunesse *et al.*, Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology* **28**, 2570-2580. e2576 (2018).

2. N. Rädecker *et al.*, Heat stress destabilizes symbiotic nutrient cycling in corals. *Proceedings of the National Academy of Sciences* **118**, (2021).

3. T. Xiang *et al.*, Symbiont population control by host-symbiont metabolic interaction in Symbiodiniaceae-cnidarian associations. *Nature Communications* **11**, 1-9 (2020).

4. T. C. LaJeunesse *et al.*, Revival of *Philozoon* Geddes for host-specialized dinoflagellates, ‘zooxanthellae’, in animals from coastal temperate zones of northern and southern hemispheres. *European Journal of Phycology*, 1-15 (2021).

5. X. Pochon, T. C. LaJeunesse, *Miliolidium* n. gen, a new symbiodiniacean genus whose members associate with soritid foraminifera or are free-living. *Journal of Eukaryotic Microbiology* **68**, e12856 (2021).

6. M. R. Nitschke *et al.*, Description of *Freudenthalidium* gen. nov. and *Halluxium* gen. nov. to formally recognize clades Fr3 and H as genera in the family Symbiodiniaceae (Dinophyceae). *Journal of Phycology* **56**, 923-940 (2020).

7. R. A. González-Pech, D. Bhattacharya, M. A. Ragan, C. X. Chan, Genome evolution of coral reef symbionts as intracellular residents. *Trends in Ecology & Evolution* **34**, 799-806 (2019).

8. R. A. González-Pech *et al.*, Comparison of 15 dinoflagellate genomes reveals extensive sequence and structural divergence in family Symbiodiniaceae and genus *Symbiodinium*. *BMC Biology* **19**, 1-22 (2021).

9. Y. Van de Peer, S. Maere, A. Meyer, The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics* **10**, 725-732 (2009).

10. M. E. Schranz, S. Mohammadin, P. P. Edger, Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current Opinion in Plant Biology* **15**, 147-153 (2012).

11. G. C. Conant, K. H. Wolfe, Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Molecular Systems Biology* **3**, 129 (2007).

12. G. C. Conant, K. H. Wolfe, Turning a hobby into a job: how duplicated genes find new functions. *Nature Reviews Genetics* **9**, 938-950 (2008).

13. S. Ohno, U. Wolf, N. B. Atkin, Evolution from fish to mammals by gene duplication. *Hereditas* **59**, 169-187 (1968).

14. P. P. Singh, J. Arora, H. Isambert, Identification of ohnolog genes originating from whole genome duplication in early vertebrates, based on synteny comparison across multiple genomes. *PLoS Computational Biology* **11**, e1004394 (2015).

15. M. Kellis, B. W. Birren, E. S. Lander, Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617-624 (2004).

16. B. Gallone *et al.*, Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts. *Cell* **166**, 1397-1410. e1316 (2016).

17. J. Peter *et al.*, Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* **556**, 339-344 (2018).

18. F. Bigey *et al.*, Evidence for Two Main Domestication Trajectories in *Saccharomyces cerevisiae* Linked to Distinct Bread-Making Processes. *Current Biology* **31**, 722-732.e725 (2021).

19. J.-M. Aury *et al.*, Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature* **444**, 171-178 (2006).

20. C. L. McGrath, J.-F. Gout, T. G. Doak, A. Yanagi, M. Lynch, Insights into three whole-genome duplications gleaned from the *Paramecium caudatum* genome sequence. *Genetics* **197**, 1417-1428 (2014).

21. D. T. Pettay, T. C. Lajeunesse, Microsatellite loci for assessing genetic diversity, dispersal and clonality of coral symbionts in ‘stress-tolerant’clade D *Symbiodinium*. *Molecular Ecology Resources* **9**, 1022-1025 (2009).

22. D. T. Pettay, D. C. Wham, R. T. Smith, R. Iglesias-Prieto, T. C. LaJeunesse, Microbial invasion of the Caribbean by an Indo-Pacific coral zooxanthella. *Proceedings of the National Academy of Sciences* **112**, 7513-7518 (2015).

23. G. P. Tiley, M. S. Barker, J. G. Burleigh, Assessing the performance of Ks plots for detecting ancient whole genome duplications. *Genome Biology and Evolution* **10**, 2882-2898 (2018).

24. A. J. Bellantuono, K. E. Dougan, C. Granados-Cifuentes, M. Rodriguez-Lanetty, Free-living and symbiotic lifestyles of a thermotolerant coral endosymbiont display profoundly distinct transcriptomes under both stable and heat stress conditions. *Molecular Ecology* **28**, 5265-5281 (2019).

25. G. C. Conant, K. H. Wolfe, Functional partitioning of yeast co-expression networks after genome duplication. *PLoS Biology* **4**, e109 (2006).

26. I. Yuyama, M. Ishikawa, M. Nozawa, M.-a. Yoshida, K. Ikeo, Transcriptomic changes with increasing algal symbiont reveal the detailed process underlying establishment of coral-algal symbiosis. *Scientific Reports* **8**, 1-11 (2018).

27. Z. Chen *et al.*, De novo assembly of the goldfish (*Carassius auratus*) genome and the evolution of genes after whole-genome duplication. *Science Advances* **5**, eaav0547 (2019).

28. Y. Wang, B. Guo, The divergence of alternative splicing between ohnologs in teleost fishes. *BMC Ecology and Evolution* **21**, 1-10 (2021).

29. L. J. Chakravarti, M. J. van Oppen, Experimental evolution in coral photosymbionts as a tool to increase thermal tolerance. *Frontiers in Marine Science* **5**, 227 (2018).

30. L. Carretero-Paulet, Y. Van de Peer, The evolutionary conundrum of whole-genome duplication. *American Journal of Botany* **107**, 1101 (2020).

31. E. A. Drew, The biology and physiology of alga-invertebrates symbioses. II. The density of symbiotic algal cells in a number of hermatypic hard corals and alcyonarians from various depths. *Journal of Experimental Marine Biology and Ecology* **9**, 71-75 (1972).

32. J. W. Porter, L. Muscatine, Z. Dubinsky, P. G. Falkowski, Primary Production and Photoadaptation in Light- and Shade-Adapted Colonies of the Symbiotic Coral, *Stylophora pistillata*. *Proceedings of the Royal Society B: Biological Sciences* **222**, 161-180 (1984).

Acknowledgments:

This project is supported by computational resources of the National Computational Infrastructure (NCI) National Facility systems through the NCI Merit Allocation Scheme (Project d85) awarded to CXC. We thank Brian Kemish and David Green for their technical assistance during genome assembly as well as Joel Burke and Michelle Havlik for technical assistance with the algal cultures. #

Funding:

Australian Research Council grant DP190102474 (CXC, DB)
Australian Research Council grant FL180100036 (MJHvO)
National Science Foundation grant CAREER1453519 (MRL)
City University of New York (CGC)
University of Technology Sydney (TK)

Author contributions:

Conceptualization: KED, AJB, DJS, CXC, MRL

Methodology: KED, AJB, TK, RMA, YC, SS, CXC

Investigation: KED, AJB, TK, RMA, YC, SS

Visualization: KED, DB, CXC

Funding acquisition: TK, CGC, DB, DJS, CXC, MRL

Supervision: DB, DJS, CXC, MRL

Writing – original draft: KED, DB, CXC

Writing – review & editing: KED, AJB, MJHvO, DB, DJS, CXC, MRL

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

The assembled genomes, predicted gene models, and proteins for *D. trenchii* CCMP2556 and SCF082 are available at Cloudstor (https://cloudstor.aarnet.edu.au/plus/s/XNkoZSH5MKEG2WO).

List of Supplementary materials:

Materials and Methods
Supplementary Text
Figs. S1-S20
Tables S1-S17
References 33-62

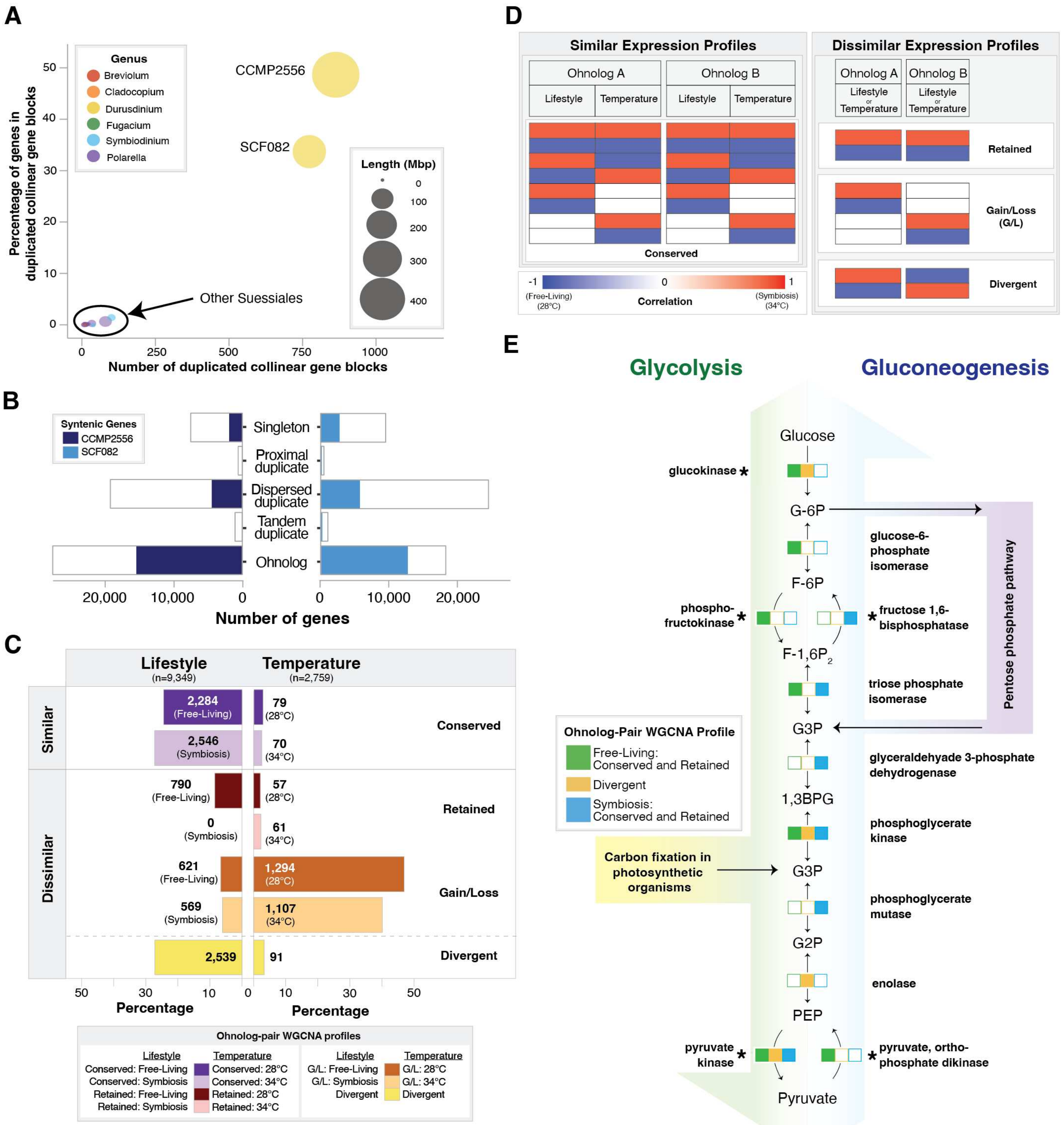

A
Genus
Breviolum
Cladocopium
Durusdinium
Fugacium
Symbiodinium
Polarella
CCMP2556
SCF082
Length (Mbp)
0
100
200
300
400
Other Suessiales
Percentage of genes in duplicated collinear gene blocks
0
10
20
30
40
50
0
250
500
750
1000
Number of duplicated collinear gene blocks
B
Syntenic Genes
CCMP2556
SCF082
Singleton
Proximal duplicate
Dispersed duplicate
Tandem duplicate
Ohnolog
20,000
10,000
0
0
10,000
20,000
Number of genes
C
Lifestyle
(n=9,349)
Temperature
(n=2,759)
Similar
Dissimilar
2,284
(Free-Living)
2,546
(Symbiosis)
79
(28°C)
70
(34°C)
Conserved
790
(Free-Living)
0
(Symbiosis)
57
(28°C)
61
(34°C)
Retained
621
(Free-Living)
569
(Symbiosis)
1,294
(28°C)
1,107
(34°C)
Gain/Loss
2,539
91
Divergent
50
30
10
0
10
30
50
Percentage
Percentage
Ohnolog-pair WGCNA profiles
Lifestyle
Conserved: Free-Living
Conserved: Symbiosis
Retained: Free-Living
Retained: Symbiosis
Temperature
Conserved: 28°C
Conserved: 34°C
Retained: 28°C
Retained: 34°C
Lifestyle
G/L: Free-Living
G/L: Symbiosis
Divergent
Temperature
G/L: 28°C
G/L: 34°C
Divergent
D
Similar Expression Profiles
Ohnolog A
Ohnolog B
Lifestyle
Temperature
Lifestyle
Temperature
Conserved
-1
0
1
(Free-Living)
(28°C)
Correlation
(Symbiosis)
(34°C)
Dissimilar Expression Profiles
Ohnolog A
Ohnolog B
Lifestyle or Temperature
Lifestyle or Temperature
Retained
Gain/Loss (G/L)
Divergent
E
Glycolysis
Gluconeogenesis
Glucose
glucokinase *
G-6P
glucose-6-phosphate isomerase
F-6P
phospho-fructokinase *
* fructose 1,6-bisphosphatase
F-1,6P2
triose phosphate isomerase
G3P
Pentose phosphate pathway
glyceraldehyde 3-phosphate dehydrogenase
1,3BPG
phosphoglycerate kinase
Carbon fixation in photosynthetic organisms
G3P
phosphoglycerate mutase
G2P
enolase
PEP
pyruvate kinase *
* pyruvate, ortho-phosphate dikinase
Pyruvate
Ohnolog-Pair WGCNA Profile
Free-Living: Conserved and Retained
Divergent
Symbiosis: Conserved and Retained

Fig. 1. Ohnolog retention and divergence in gene expression.

(**A**) The percentage of genes in duplicated collinear gene blocks is shown relative to the number of duplicated collinear gene blocks identified within the genomes of Suessiales species. (**B**) The number of genes in CCMP2556 and SCF082 across the different MCScanX duplication categories. Colored portions of bars represent the number of syntenic genes between the two isolates from that category. (**C**) Breakdown of ohnolog pairs classified into each evolutionary scenario of conservation/divergence according to their distribution in the WGCNA modules. (**D**) Summary of the types of ohnolog-pair expression profiles represented by each category of evolutionary conservation. (**E**) Diagram showing the divergence in glycolysis/gluconeogenesis pathways predominantly found within putative WGD-duplicated regions in context of "Conserved", "Retained", and "Divergent" ohnolog-pairs.

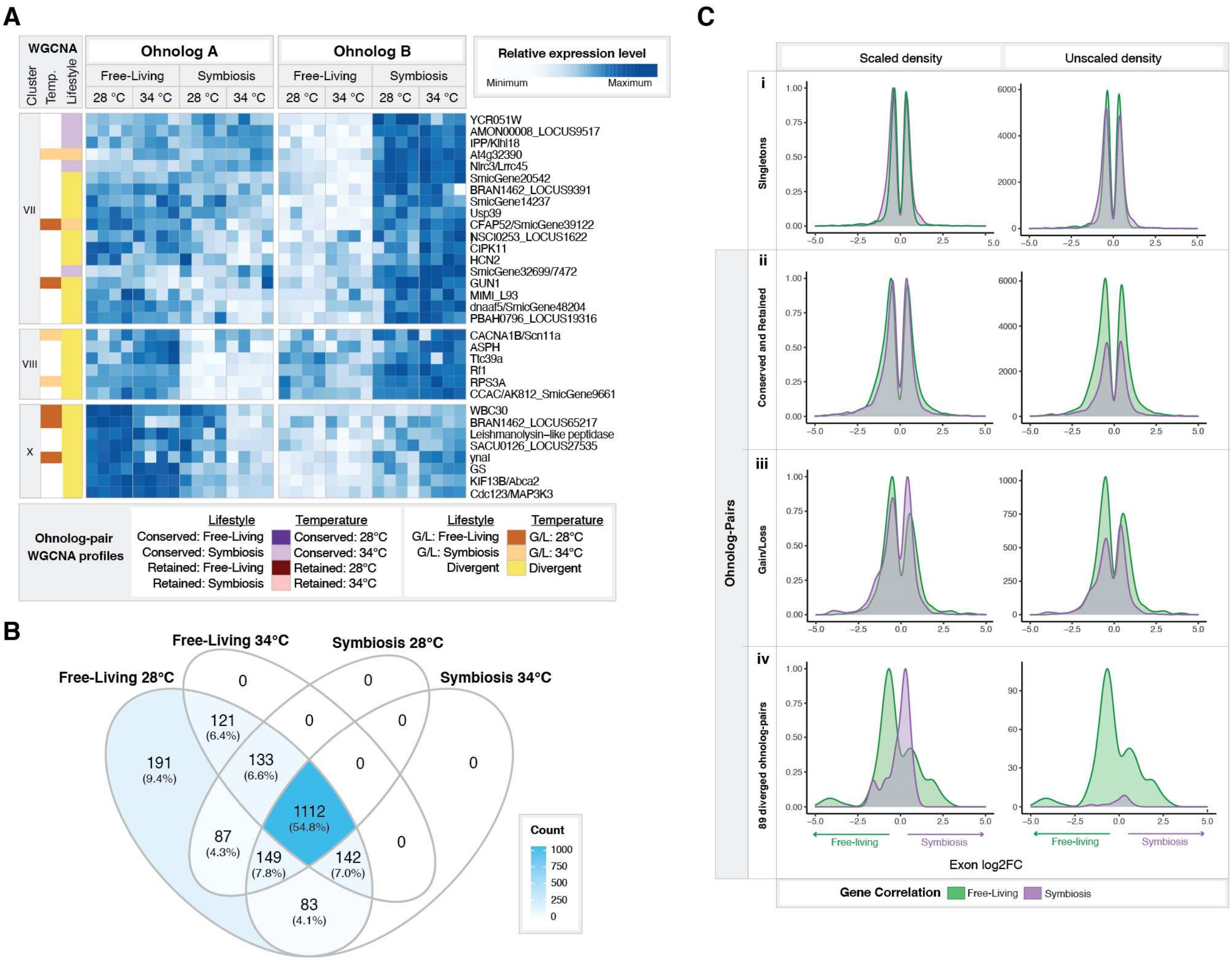
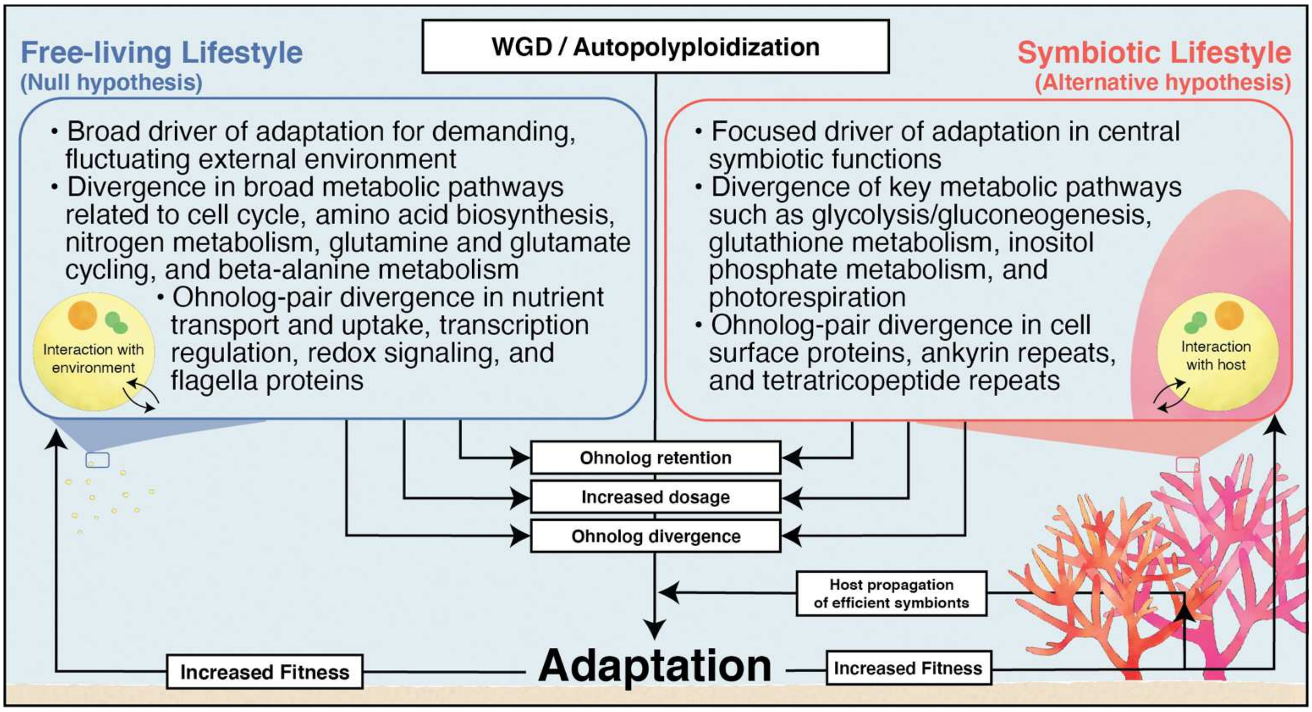


Fig. 2. Lifestyle drives ohnolog specialization via exon restructuring.

(**A**) Gene expression heatmap displaying a subset of the ohnolog-pairs exhibiting divergence and specialization to different growth conditions with ohnolog-pairs clustered according to their expression patterns using Euclidean distances. Expression levels were scaled within each row from the minimum to maximum value to allow comparison of expression between the two genes in an ohnolog-pairs. (**B**) Venn diagram depicting the number of genes displaying mRNA editing across the treatments. (**C**) Scaled and unscaled density plots of the log$_2$FC of differentially used exons ($p < 0.001$) for (i) singletons, (ii) all "Conserved" and "Retained" in context of lifestyle,

(iii) the ohnolog from "Gain/Loss" pairs exhibiting a correlation to lifestyle, and (iv) all ohnologs exhibiting a correlation to a particular lifestyle from the 90 divergent ohnolog-pairs. Directionality of the $\log_2$FC change is indicated along the x-axis, with different colors indicating their gene-level correlation to either the free-living (green) or symbiotic (purple) lifestyles.

330 

331 **Fig. 3. Model of how free-living and symbiotic lifestyles influenced a facultative symbiont's**
332 **evolution after WGD.**