

# Current Biology

Volume 35  
Number 5  
March 10, 2025



Report

# Phylogenomics of coral-infecting corallicolids reveal multiple independent losses of chlorophyll biosynthesis in apicomplexan parasites

Victoria K.L. Jacko-Reynolds,<sup>1,8,\*</sup> Waldan K. Kwong,<sup>1,2</sup> Samuel J. Livingston,<sup>1</sup> Morelia Trznadel,<sup>1</sup> Anthony M. Bonacolta,<sup>1</sup> Gordon Lax,<sup>1</sup> Jade Shivak,<sup>1</sup> Nicholas A.T. Irwin,<sup>1,3</sup> Mark J.A. Vermeij,<sup>4,5</sup> Javier del Campo,<sup>6</sup> and Patrick J. Keeling<sup>1,7</sup>

<sup>1</sup>Department of Botany, University of British Columbia, 3156-6270 University Blvd., Vancouver, BC V6T 1Z4, Canada

<sup>2</sup>Gulbenkian Institute of Molecular Medicine, Rua da Quinta Grande, 6, 2780-156 Oeiras, Portugal

<sup>3</sup>Gregor Mendel Institute of Molecular Plant Biology, Austrian Academy of Sciences, Dr. Bohr-Gasse 3, 1030 Vienna, Austria

<sup>4</sup>Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 700, Amsterdam, 1098 XH, the Netherlands

<sup>5</sup>CARMABI Foundation, P.O. Box 2090, Piscaderabaai z/n, Willemstad, Curaçao

<sup>6</sup>Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Catalonia, Spain

<sup>7</sup>Bluesky: @pjkeelinglab

<sup>8</sup>Lead contact

\*Correspondence: [victoriak.jacko-reynolds@botany.ubc.ca](mailto:victoriak.jacko-reynolds@botany.ubc.ca)

<https://doi.org/10.1016/j.cub.2025.01.028>

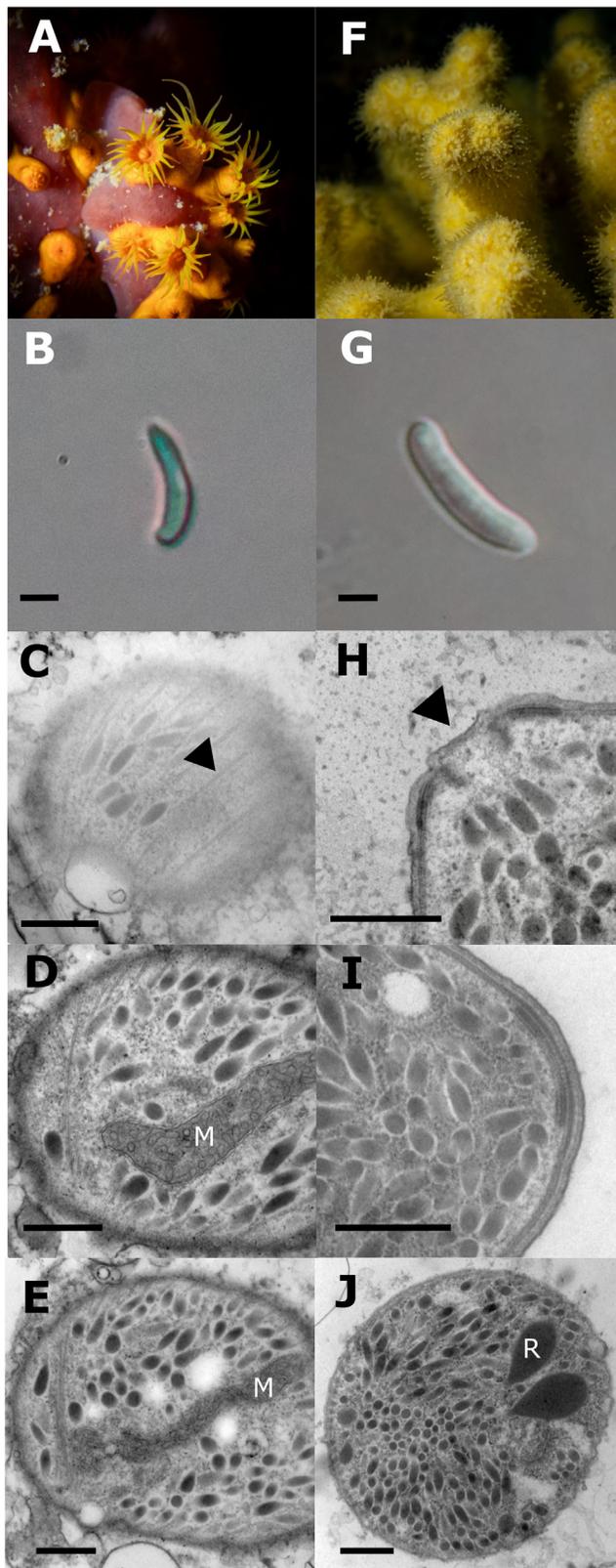
## SUMMARY

The transition from free-living to parasitic lifestyles induces major shifts in evolution, and nowhere is this more acute than in apicomplexans—obligate intracellular parasites of animals that evolved from photosynthetic algae.<sup>1</sup> In other cases where photosynthesis has been lost, including most apicomplexans, chlorophyll is also absent, but in coral-infecting apicomplexans (corallicolids), chlorophyll biosynthesis genes are retained in the plastid genome despite their lack of photosystems.<sup>2</sup> This suggests that the loss of photosynthesis and chlorophyll were decoupled in this lineage, but because these observations are only based on plastid genomes, two fundamental questions remain unclear. First, how this impacted apicomplexan evolution as a whole is unclear because there are conflicting phylogenetic positions for corallicolids: plastid gene phylogenies place them at the base of the apicomplexans, whereas nuclear rRNA places them with late-branching coccidians (suborder Eimeriorina).<sup>2,3</sup> Second, it is unclear if chlorophyll or a metabolic intermediate is synthesized, as most chlorophyll biosynthesis enzymes are encoded in the nucleus. To address these questions, we have sequenced transcriptomes from two corallicolids, infecting *Parazoanthus swiftii* and *Madracis mirabilis* hosts. Phylogenomic data strongly support a late-branching relationship closer with coccidians, specifically with the protococcidians and the newly discovered ichthyocolids. We also find evidence for the expression of nucleus-encoded enzymes involved in chlorophyll biosynthesis in corallicolids and protococcidians. Overall, we conclude that chlorophyll synthesis was likely retained through the early evolution of the group and then lost approximately 10 times independently, emphasizing the impact of parallel evolutionary changes in parasitic transitions.

## RESULTS AND DISCUSSION

Since the discovery that corals and their close relatives around the world are frequently infected with corallicolid apicomplexan parasites,<sup>2,4–6</sup> a number of questions surrounding the evolution and biology of the parasites have remained out of reach due to an almost complete lack of data of nuclear genomes from this group. However, acquiring genomic data from uncultured intracellular parasites of corals, which are also hosts to complex symbiotic communities consisting of many other microbial species, has been challenging. Extensive and deep sequencing of numerous corals together with their symbiotic communities yielded data only from consistently high-copy-number regions of the genome, including mitochondrial or plastid genomes, as

well as nuclear rRNA.<sup>2,3,6</sup> At the same time, transcriptome sequencing from individual corallicolid cells yielded poor results, likely because of the small cell size.<sup>7</sup> To attempt to circumvent these problems, we took two different approaches. First, we sequenced transcriptomes from multiple pools of 5–32 cells isolated from the same infected host, and second, we sequenced meta-transcriptomes from infected tissue samples that were enriched for parasites by Percoll gradient centrifugation. Both approaches were applied in the field over five seasons in Curaçao, from 2018 to 2023, focusing on two hosts previously found to have a 100% infection rate: the yellow pencil coral *Madracis mirabilis* and the golden zoanthid *Parazoanthus swiftii* (collections summarized in [Table S1](#)). *Madracis mirabilis* is a largely heterotrophic stony coral ([Figure 1A](#)) with low levels of symbiotic



**Figure 1. Coralicolids from two hosts have similar, parasitic apicomplexan features**

Left: *Anthozoaphila gnarlus* infecting *Parazoanthus swiftii*.  
(A) The host, *P. swiftii* living symbiotically on a sponge.

zooxanthellae, while *Parazoanthus swiftii* is a zoanthid that lives on sponges and lacks symbiotic zooxanthellae altogether (Figure 1F). For both hosts, density-gradient centrifugation was optimized to enrich for corallicolids, which were identified by light microscopy (Figures 1B and 1G), and more closely inspected using transmission electron microscopy (Figures 1C–1E and 1H–1J).

Overall, these approaches yielded 24 corallicolid transcriptomes: six enrichment libraries and nine cell-pool libraries from *P. swiftii* and six enrichment libraries and three cell-pool libraries from *M. mirabilis*. Although the cell-pool libraries were found to be superior to the single-cell libraries reported previously, they still suffered from poor coverage (containing 5%–7% of our 201 gene phylogenomic dataset). The enrichment libraries, on the other hand, yielded very good coverage (45%–87% of the 201 genes). Unenriched libraries of whole *P. swiftii* polyps and sponge host were also sequenced to aid in the discrimination between host and parasite genes since the genomes of these species have not been characterized. To further test the quality of the libraries and, in particular, whether libraries of parasites from the same host can be combined for analysis, we confirmed the host identity and that hosts were consistently infected with the same parasite by extracting host and parasite SSU and LSU rRNA from each library. In each case, presence of host rRNA was consistent with field identifications. For parasite rRNA, phylogenies were inferred using existing corallicolid sequences, including previously reported sequences derived from both hosts.<sup>3</sup> This phylogeny (Figure S2A) showed that all new samples from each host form a strongly supported clade to the exclusion of one another, once again confirming the lack of cross-contamination and showing each host was consistently infected by a single, distinct parasite. In addition, all new *P. swiftii* samples branch with previously sampled parasites from the same host<sup>3</sup>; however, the previously sampled parasites from *M. mirabilis* did not cluster with the new *M. mirabilis* samples but instead also clustered with parasites from *P. swiftii* (Figure S2A). There is no evidence for host misidentification in the earlier samples (i.e., there is no *P. swiftii* rRNA), suggesting that *M. mirabilis* was infected with different parasites in previous sampling years and that parasites can infect distantly related anthozoan hosts, as biogeographical data have also begun to suggest.<sup>5</sup> The parasite in *P. swiftii* was described as *Anthozoaphila gnarlus*, but with a type host of *M. mirabilis*, so we continue

(B) Light micrograph of a single *A. gnarlus* trophozoite. Scale bar: 1  $\mu$ m.

(C–E) Three transmission electron microscopy (TEM) cross sections through a single *A. gnarlus* cell.

(C) Arrow highlights the microtubule cytoskeleton of the corallicolid extending toward the apex of the cell.

(D) Section highlighting the large mitochondria.

(E) Section with broad overview of the cell and showing micronemes. Scale bars: 500 nm in (C)–(E).

Right: Corallicolid ex. *Madracis mirabilis*.

(F) The host *M. mirabilis*.

(G) Light micrograph of corallicolid ex. *M. mirabilis*. Scale bar: 1  $\mu$ m.

(H) The conoid structure of the apical complex of a corallicolid ex. *M. mirabilis* indicated by a black arrow.

(I) Several micronemes.

(J) Rhoptry-like organelles.

Scale bars: 500 nm in (H)–(J). R, rhoptry-like organelles; M, mitochondria.

See also Figure S1.

to use that name here and propose a host range expansion to include *P. swiftii*, where it was also originally detected.<sup>3</sup> Altogether, these data confirm that the transcriptomes from each host consistently yield closely related parasites and that the parasites from the two hosts sampled from 2021–2023 are distinct and distantly related within the corallicolid tree.

The transcriptome data were next used to resolve the conflicting positions of the corallicolids in the apicomplexan phylogeny. Previous analyses of plastid genome sequences placed corallicolids as the sister to all other apicomplexans with strong support,<sup>2,4</sup> whereas nuclear rRNA phylogenies contrastingly showed corallicolids diverging from other apicomplexans much more recently, as sister to the coccidians (subord. Eimeriorina).<sup>2,8</sup> To resolve this discrepancy, we inferred a phylogenomic tree based on 201 genes with 61,714 sites across 48 taxa, including high coverage from corallicolids from both *P. swiftii* (covering 81% of total sites) and *M. mirabilis* (89% of sites), using maximum-likelihood (ML) and Bayesian methods (Figures 2 and S2B). All analyses consistently show corallicolids branching late, close with the coccidians, with full support. However, ML and Bayesian trees differ in one detail: in ML trees, corallicolids branch specifically with the protococcidian *Eleutheroschizon*, whereas the Bayesian phylogeny supports corallicolids branching as sister to coccidians (Figure S2B). The two alternatives were tested using approximately unbiased (AU) tests on 11 topologies where corallicolids were constrained to be sister to each subgroup of apicomplexans and the outgroups. All topologies were rejected at a 5% significance, except for those where corallicolids branch as sister to both coccidians and protococcidians, sister to protococcidians (the ML topology, Figure 2), or as sister to coccidians (the Bayesian topology, Figure S2B).

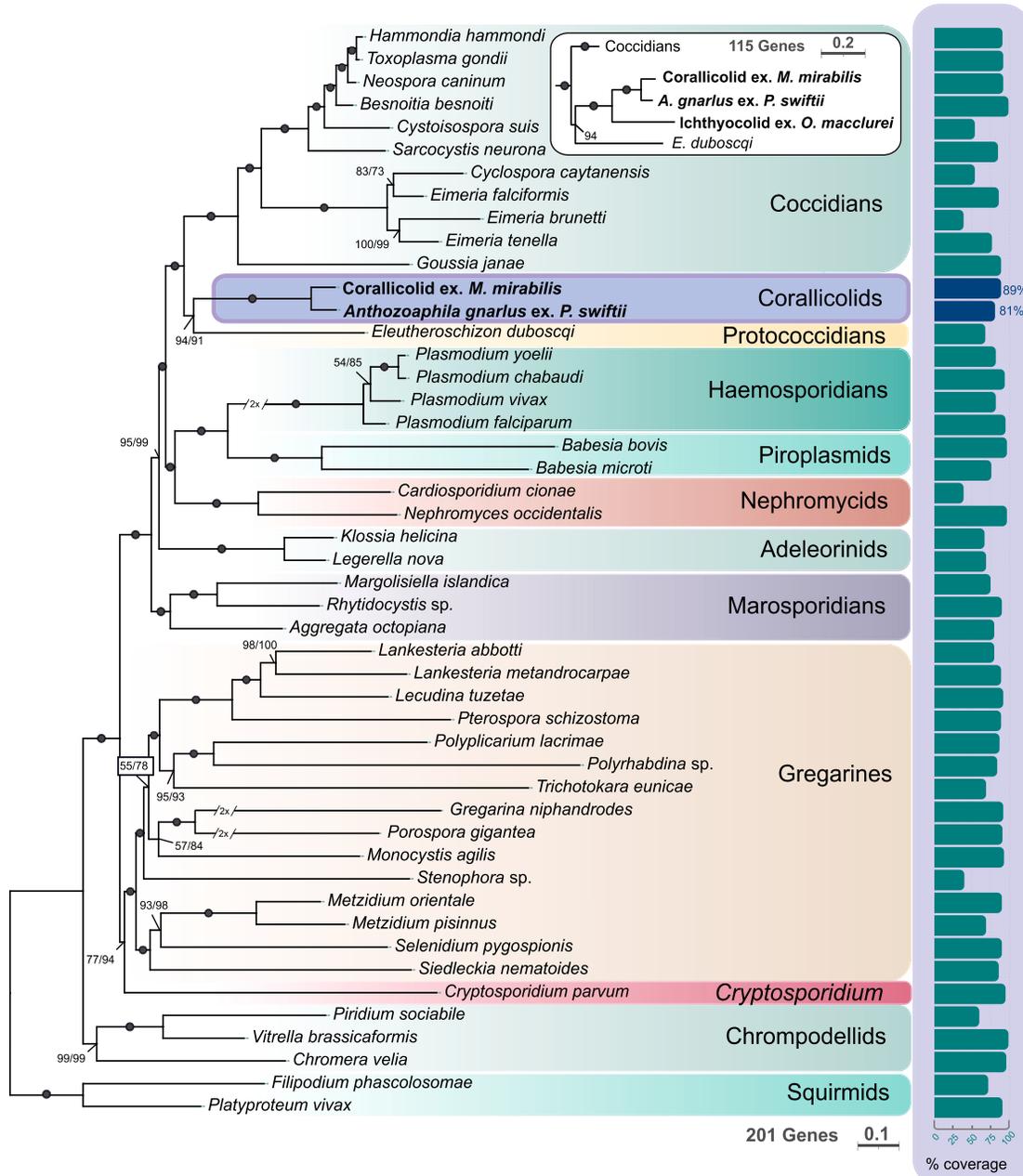
To test whether the inclusion of the ichthyocolids, a recently described group of fish-infecting apicomplexans, affects this result, we searched for ichthyocolid nuclear genes in infected fish blood metagenomic data<sup>8</sup> and found 125 genes of apicomplexan origin. There was limited overlap in the metagenomic gene sampling and the transcriptomic gene sampling, but we were able to construct a phylogenomic tree consisting of 115 genes with high coverage for both groups (Figure S2C). Phylogenomic analysis of this dataset showed complete support for a sister relationship between corallicolids and ichthyocolids, and in this case, both Bayesian and ML analyses were congruent with one another and with the 201-gene ML analysis in supporting their branching with the protococcidia (Figures 2 inset, S2B inset, and S2C).

The phylogenetic relationship of corallicolids to protococcidians or coccidians makes any explanation for the distribution of chlorophyll biosynthesis genes in apicomplexans complicated. Chlorophyll is obviously fundamental to photosynthesis, and in other cases where photosynthesis has been lost, the loss of chlorophyll biosynthesis is so closely linked as to appear concurrent. This is also true in the non-photosynthetic relatives of apicomplexans (squirmids and some chrompodellids), but corallicolids are the exception where the two processes have evidently been decoupled. Had the phylogeny shown corallicolids branching at the base of the apicomplexan tree, like plastid phylogenies originally suggested,<sup>4</sup> they could be interpreted as descendants of an intermediate stage that retained chlorophyll

for some reason specific to their biology. But because they diverged from other apicomplexans much more recently, we infer that chlorophyll biosynthesis must have been retained throughout the early evolution of apicomplexans, ultimately being lost in parallel in ancestors of almost every major group but retained in the corallicolids. This raises numerous questions about the photosystem-independent function of this pathway.

Corallicolid plastid genomes encode four enzymes for chlorophyll biosynthesis<sup>2,6</sup> (*chlL*, *chlN*, *chlB*, and *acsF*), which together mediate light-independent conversion of magnesium-protoporphyrin IX monomethyl ester to protochlorophyllide, the precursor of chlorophyll a. These enzymes are absent in all other apicomplexan plastid genomes sequenced to date, including the closely related ichthyocolids,<sup>8</sup> and coccidians.<sup>9</sup> No plastid genome of any protococcidian has yet been sequenced; however, we found plastid DNA in the only available transcriptome from this group, *Eleutheroschizon duboscqui*,<sup>10</sup> and assembling its draft plastid genome revealed no evidence for these four genes (Figure S3). Moreover, the complete pathway also requires many other enzymes that are typically encoded in the nucleus (Figure 3). The distribution of these enzymes in apicomplexans is unknown: their presence in corallicolids has been impossible to test without nuclear genomic data, and they have not been directly or thoroughly examined in other apicomplexans because there was no reason to suspect presence of chlorophyll biosynthesis. We therefore searched for chlorophyll biosynthesis enzymes in our corallicolid transcriptomes and all other apicomplexan genomes and transcriptomes. The phylogenetic position of all putative hits from profile hidden Markov models (HMMs) and BLAST searches was checked to exclude contamination and members of related gene families that are not directly involved in chlorophyll biosynthesis (phylogenies are available on UBC Borealis: <https://doi.org/10.5683/SP3/2LXG4D>). Altogether we found eight positive hits: all four plastid-encoded genes (*chlL*, *chlN*, *chlB*, and *acsF*) plus nucleus-encoded genes, *chlI*, *chlM*, and *chlG* in the corallicolid *Anthozoaphila*, as well as nucleus-encoded *chlG* in *Eleutheroschizon* (Figures 3 and S2D–S2F). This does not account for the entire pathway, as *chlH*, *chlD*, and *dvr* were undetected (Figure 3). The transcripts we recovered had low read coverage, suggesting that the genes in this pathway are expressed at low levels, which indicates that transcripts for the missing enzymes might be present but undetected. The presence of *chlG* in *Anthozoaphila* and *Eleutheroschizon* is particularly significant because, as the final enzyme in the pathway, its retention suggests the product of the pathway is likely chlorophyll and not some intermediate like protochlorophyllide. The presence of *chlG* in *Eleutheroschizon* is also unusual, however, since its plastid appears to lack *chlL*, *chlN*, *chlB*, and *acsF*. The possibility that chlorophyll biosynthetic genes may be found in nuclear genomes of both protococcidians and ichthyocolids will require further investigation, as will the end product and function of the pathway in corallicolids. We found no evidence of any enzymes for chlorophyll biosynthesis from any other apicomplexans (Figure 3).

We also identified genes involved in other metabolic pathways known to be retained across apicomplexan plastids and found that most genes in the MEP/DOXP and FASII pathways were represented in at least one corallicolid library, whereas a smaller



**Figure 2. Multi-protein phylogenomic tree showing corallicolids are related to coccidians**

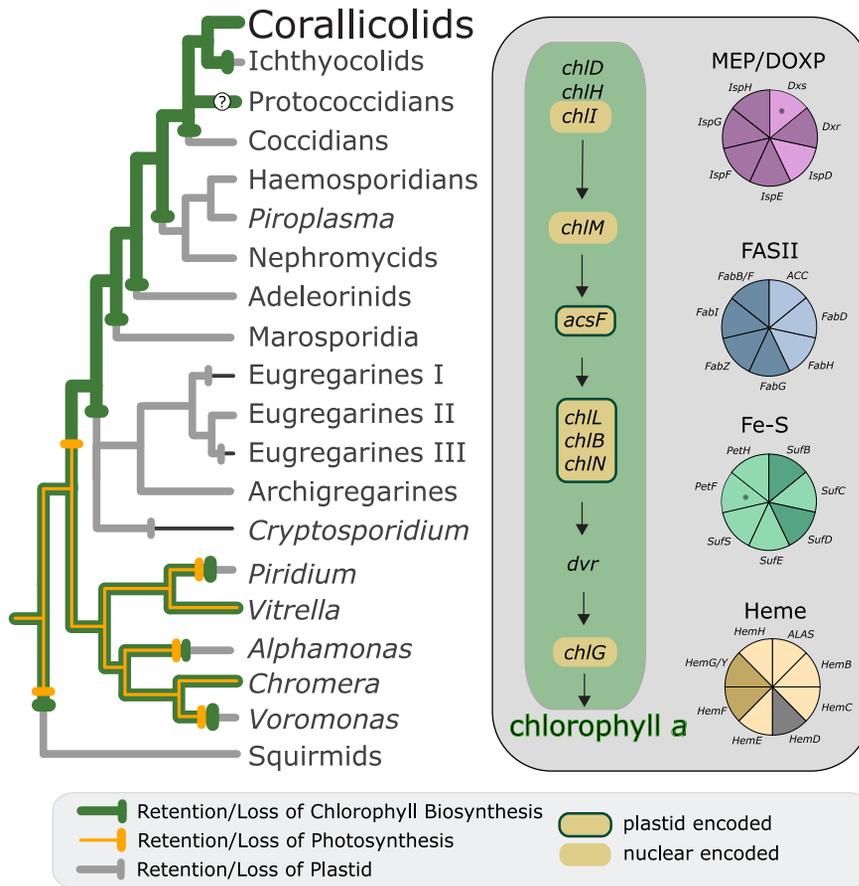
The maximum-likelihood tree was generated from an alignment of 201 genes with 61,714 sites across 48 apicomplexan taxa under the LG+C60+F+G substitution model with 1,000 ultrafast bootstraps (UFBs) and 100 non-parametric bootstraps (NPBs) with posterior mean site frequencies (PMSFs). The black circles indicate 100% UFB/NPB support, and values lower than 100% are indicated at their respective branch. The percent coverage of genes across the 201 gene dataset is indicated on the right for each taxon. Inset: portion of the maximum-likelihood tree generated from an alignment of 115 genes across the same 48 apicomplexan taxa with the inclusion of ichthyocolid data under the LG+C60+F+G substitution model.

See also [Figure S2](#).

proportion of Fe-S and heme synthesis pathway genes were identified ([Figure 3](#)). This suggests that corallicolid plastids can also perform the same biochemical functions as those of other apicomplexans.

It has widely been assumed that photosynthesis was lost once at the origin of apicomplexans, generally conceptually coupled

to a single origin of parasitism. However, parsimonious evolutionary interpretations of this lineage can be misleading, especially relating to plastid and mitochondrial evolution.<sup>11–14</sup> Since both the loss of photosynthesis and the origin of parasitism have taken place multiple times in the closely related chrompodellids and squirmids,<sup>12–14</sup> we should not categorically rule out



**Figure 3. Multiple parallel losses of photosynthesis, chlorophyll biosynthesis, and plastid organelles in apicomplexans and relatives**

Left: schematic of the phylogenetic tree highlighting that corallicolids are related to the coccidians. Gray branches trace the retention of the actual plastid organelle in apicomplexans and their close relatives, with losses indicated by a perpendicular bar, whereas yellow branches trace the retention of photosynthesis, and green branches trace the retention of chlorophyll biosynthesis. The question mark highlights the uncertainty of chlorophyll biosynthesis retention/loss, as only a single gene, *chlG*, from the pathway has been detected in protococcidians. Here, we propose that ancestral apicomplexans underwent a single loss of photosynthesis and decoupling from chlorophyll biosynthesis following the transition to parasitism. As chromodellids have free-living members, the multiple origins of parasitism are also paired with multiple losses of photosynthesis. Considering apicomplexans do not have any known free-living members, we choose the most parsimonious explanation for a single loss of photosynthesis; however, we do not rule out the possibility that photosynthesis too may have been lost several times in parallel in apicomplexans.

Center: chlorophyll biosynthesis genes identified in corallicolids and protococcidians. Dark green-outlined yellow boxes indicate the positive identification of a plastid-encoded gene in the pathway, whereas yellow boxes without an outline indicate the positive identification of a nucleus-encoded plastid-targeted gene from the

pathway. The chlorophyll biosynthetic pathway: magnesium chelatase ATPase subunits D, H, and I (*chID*, *chlH*, and *chlI*); magnesium-protoporphyrin IX methyltransferase (*chlM*); magnesium-protoporphyrin IX monomethyl ester (oxidative) cyclase (*acsF*); light-independent protochlorophyllide oxidoreductase, iron-sulfur ATP-binding protein (*chlL*); light-independent protochlorophyllide oxidoreductase (DPOR) subunits B and N (*chlB/N*); 3,8 divinyl reductase (*dvr*); chlorophyll synthase (*chlG*).

Right: the retention of other metabolic plastid pathway genes found across the 24 corallicolid transcriptomes and *Eleutheroschizon*. Plastid metabolic pathways include the non-mevalonate isoprenoid biosynthesis pathway (MEP/DOXP), fatty acid biosynthesis pathway (FASII), iron-sulfur cluster synthesis pathway (Fe-S), and heme biosynthetic pathway. Darker-colored sections indicate the gene was identified in both corallicolid and *Eleutheroschizon* transcriptomic data, whereas lighter-colored sections indicate the gene was identified in only *Eleutheroschizon* data. Lighter-colored sections with a dot indicate that the gene was only found in corallicolids. Gray section indicates that the gene was not found in either *Eleutheroschizon* or corallicolids. See also Figure S2.

the possibility that either or both might also have taken place multiple times within the apicomplexans, which could help explain the long ancestral persistence of chlorophyll biosynthesis. We do not favor this explanation, however, for the simple reason that it still lacks direct evidence: in all other cases where such parallelisms are inferred, the character state in question is phylogenetically interspersed with other states—for example, both photosynthetic chrompodellids have non-photosynthetic sisters, making the case for parallel loss of photosynthesis very strong. If further exploration reveals truly free-living or photosynthetic lineages *within* the apicomplexans, then a model where multiple losses of photosynthesis contribute to the distribution of chlorophyll becomes more likely.

Even if we maintain the assumption of a single loss of photosynthesis near to the origin of apicomplexans, the incredibly unparsimonious distribution of chlorophyll biosynthesis has several implications for its still-mysterious function. For example, whatever function it does serve is unlikely to be

related to the ecology of corallicolids: there is no indication that chlorophyll biosynthesis is related in any way to infecting coral, living on coral reefs, or infecting hosts that possess photosynthetic symbionts, or even living in a sunlit habitat since it is highly improbable that these characteristics trace back through the entire backbone of the apicomplexan tree to their photosynthetic common ancestor. For example, it has been proposed that chlorophyll might be used for photo-protection or controlling the effects of light-induced ROS,<sup>15</sup> but this would require that the long line of apicomplexan ancestors consistently lived in high light environments, which is doubtful, and indeed some corallicolids infect anthozoans in the deep, aphotic zone where there would be no selection for such activity.<sup>6</sup> The function of chlorophyll is more likely related to basic cellular function. For example, corallicolid chlorophyll retention could be the result of an ancient feedback inhibition system to control heme biosynthesis. Chlorophyll and heme share tetrapyrrole precursors, and in photosynthetic plants and algae

excess protochlorophyllide can downregulate synthesis of an early tetrapyrrole precursor, ALA.<sup>16–18</sup> Another chlorophyll intermediate, Mg-Protoporphyrin-IX, has been shown to function as a retrograde signal to the nucleus to upregulate expression of nucleus-encoded plastid-targeted proteins.<sup>19</sup> If the photosynthetic ancestor of apicomplexans depended on such a control system (most obviously heme since it shares a biosynthetic root with chlorophyll and its production is one of the few functions retained in these plastids), then the pathway would remain essential after the loss of photosynthesis. This could be lost quickly if a new control system evolved, but it could also persist with different descendants evolving potentially different solutions in parallel. This situation may be more common than is currently appreciated as well, as there is emerging evidence that chlorophyll may be present in other lineages that lack photosystems, including cryptomonads<sup>20</sup> and orchids.<sup>21</sup> In these cases, photosynthesis was lost much more recently than in corallicolids, so these may simply be an intermediate phase of no functional consequence. It is also possible that this situation is relatively common in parasitic plants but difficult to recognize because in the ancestor of the angiosperms, the entire chlorophyll biosynthetic pathway was relocated to the nucleus.<sup>22</sup> Without both the telltale photosystems and chlorophyll genes in the same small genome, it becomes much harder to prove the former is absent while the latter persists.

Altogether, these findings reinforce how unparsimonious major evolutionary transitions may be, and events surrounding the origins of parasitism and especially the numerous loss-of-function events that can accompany this transition are characterized by many independent parallel occurrences in different descendent lineages. This is clearly the case for many aspects of plastid functional reduction and loss (Figure 3), and the finding that even tightly linked functional losses like photosynthesis and chlorophyll biosynthesis can be decoupled over long periods of evolutionary time are just the most extreme examples of this. Although they are a technically challenging system, further work on corallicolids will hopefully illuminate the function of chlorophyll and, by extension, why it was lost in so many other apicomplexan lineages.

## RESOURCE AVAILABILITY

### Lead contact

Requests for further information, resources, and reagents can be directed to the lead contact, Victoria K.L. Jacko-Reynolds ([victoriak.jacko-reynolds@botany.ubc.ca](mailto:victoriak.jacko-reynolds@botany.ubc.ca)).

### Materials availability

New unique reagents were not generated in this study.

### Data and code availability

- Raw transcriptome sequencing reads have been deposited in the NCBI Short Read Archive (SRA) under BioProject PRJNA1199006.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

This work was supported by grants to P.J.K. from the Gordon and Betty Moore Foundation (<https://doi.org/10.37807/GBMF9201>) and to W.K.K. from the

European Molecular Biology Organization (IG-5045-2022) and Fundação para a Ciência e a Tecnologia (CEECIND/01358/2021).

## AUTHOR CONTRIBUTIONS

V.K.L.J.-R. and P.J.K. designed and wrote the original draft. V.K.L.J.-R. carried out the investigation and representation of data with assistance from W.K.K. and A.M.B. Field methods were designed and developed by N.A.T.I. and W.K.K. with support from M.J.A.V. Sample collection was performed by V.K.L.J.-R., P.J.K., W.K.K., M.T., and G.L. with support from M.J.A.V. Transcriptome processing was performed by V.K.L.J.-R. and W.K.K. with the assistance of J.S., and metagenomic searches were performed by A.M.B. and J.d.C. Electron microscopy processing and imaging were performed by S.J.L. and M.T.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Sample isolation and corallicolid enrichments
  - Electron microscopy
  - Sequencing and transcriptome assembly
  - Orthology identification and phylogenomic analysis
  - Plastid protein search and de novo plastid genome assembly
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2025.01.028>.

Received: September 10, 2024

Revised: December 3, 2024

Accepted: January 14, 2025

Published: February 18, 2025

## REFERENCES

1. Lim, L., and McFadden, G.I. (2010). The evolution, metabolism and functions of the apicoplast. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 749–763. <https://doi.org/10.1098/rstb.2009.0273>.
2. Kwong, W.K., del Campo, J., Mathur, V., Vermeij, M.J.A., and Keeling, P.J. (2019). A widespread coral-infecting apicomplexan with chlorophyll biosynthesis genes. *Nature* 568, 103–107. <https://doi.org/10.1038/s41586-019-1072-z>.
3. Kwong, W.K., Irwin, N.A.T., Mathur, V., Na, I., Okamoto, N., Vermeij, M.J.A., and Keeling, P.J. (2021). Taxonomy of the Apicomplexan Symbionts of Coral, including Corallicolida ord. nov., Reassignment of the Genus *Gemmocystis*, and Description of New Species *Corallicola aquarius* gen. nov. sp. nov. and *Anthozoaphila gnarlus* gen. nov. sp. nov. *J. Eukaryot. Microbiol.* 68, e12852. <https://doi.org/10.1111/jeu.12852>.
4. Janoušková, J., Horák, A., Barott, K.L., Rohwer, F.L., and Keeling, P.J. (2012). Global analysis of plastid diversity reveals apicomplexan-related lineages in coral reefs. *Curr. Biol.* 22, R518–R519. <https://doi.org/10.1016/j.cub.2012.04.047>.

- Trznadel, M., Holt, C.C., Livingston, S.J., Kwong, W.K., and Keeling, P.J. (2024). Coral-infecting parasites in cold marine ecosystems. *Curr. Biol.* 34, 1810–1816.e4. <https://doi.org/10.1016/j.cub.2024.03.026>.
- Vohsen, S.A., Anderson, K.E., Gade, A.M., Gruber-Vodicka, H.R., Dannenberg, R.P., Osman, E.O., Dubilier, N., Fisher, C.R., and Baums, I.B. (2020). Deep-sea corals provide new insight into the ecology, evolution, and the role of plastids in widespread apicomplexan symbionts of anthozoans. *Microbiome* 8, 34. <https://doi.org/10.1186/s40168-020-00798-w>.
- Kolisko, M., Boscaro, V., Burki, F., Lynn, D.H., and Keeling, P.J. (2014). Single-cell transcriptomics for microbial eukaryotes. *Curr. Biol.* 24, R1081–R1082. <https://doi.org/10.1016/j.cub.2014.10.026>.
- Bonacolta, A.M., Krause-Massaguer, J., Smit, N.J., Sikkel, P.C., and del Campo, J. (2024). A new and widespread group of fish apicomplexan parasites. *Curr. Biol.* 34, 2748–2755.e3. <https://doi.org/10.1016/j.cub.2024.04.084>.
- Köhler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J.M., Palmer, J.D., and Roos, D.S. (1997). A plastid of Probable Green Algal Origin in Apicomplexan Parasites. *Science* 275, 1485–1489. <https://doi.org/10.1126/science.275.5305.1485>.
- Janouškovec, J., Paskerova, G.G., Miroliubova, T.S., Mikhailov, K.V., Birley, T., Aleoshin, V.V., and Simdyanov, T.G. (2019). Apicomplexan-like parasites are polyphyletic and widely but selectively dependent on cryptic plastid organelles. *eLife* 8, e49662. <https://doi.org/10.7554/eLife.49662>.
- Mathur, V., Wakeman, K.C., and Keeling, P.J. (2021). Parallel functional reduction in the mitochondria of apicomplexan parasites. *Curr. Biol.* 31, 2920–2928.e4. <https://doi.org/10.1016/j.cub.2021.04.028>.
- Mathur, V., Kolisko, M., Hehenberger, E., Irwin, N.A.T., Leander, B.S., Kristmundsson, Á., Freeman, M.A., and Keeling, P.J. (2019). Multiple Independent Origins of Apicomplexan-Like Parasites. *Curr. Biol.* 29, 2936–2941.e5. <https://doi.org/10.1016/j.cub.2019.07.019>.
- Janouškovec, J., Tikhonov, D.V., Burki, F., Howe, A.T., Kolisko, M., Mylnikov, A.P., and Keeling, P.J. (2015). Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proc. Natl. Acad. Sci. USA* 112, 10200–10207. <https://doi.org/10.1073/pnas.1423790112>.
- Mathur, V., Salomaki, E.D., Wakeman, K.C., Na, I., Kwong, W.K., Kolisko, M., and Keeling, P.J. (2023). Reconstruction of Plastid Proteomes of Apicomplexans and Close Relatives Reveals the Major Evolutionary Outcomes of Cryptic Plastids. *Mol. Biol. Evol.* 40, msad002. <https://doi.org/10.1093/molbev/msad002>.
- Reinbothe, S., Reinbothe, C., Apel, K., and Lebedev, N. (1996). Evolution of Chlorophyll Biosynthesis—The Challenge to Survive Photooxidation. *Cell* 86, 703–705. [https://doi.org/10.1016/s0092-8674\(00\)80144-0](https://doi.org/10.1016/s0092-8674(00)80144-0).
- Meskauskiene, R., Nater, M., Goslings, D., Kessler, F., op den Camp, R., and Apel, K. (2001). FLU: A negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 98, 12826–12831. <https://doi.org/10.1073/pnas.221252798>.
- Kauss, D., Bischof, S., Steiner, S., Apel, K., and Meskauskiene, R. (2012). FLU, a negative feedback regulator of tetrapyrrole biosynthesis, is physically linked to the final steps of the Mg<sup>++</sup>-branch of this pathway. *FEBS Lett.* 586, 211–216. <https://doi.org/10.1016/j.febslet.2011.12.029>.
- Falciatore, A., Merendino, L., Barneche, F., Ceol, M., Meskauskiene, R., Apel, K., and Rochaix, J.D. (2005). The FLP proteins act as regulators of chlorophyll synthesis in response to light and plastid signals in *Chlamydomonas*. *Genes Dev.* 19, 176–187. <https://doi.org/10.1101/gad.321305>.
- Strand, A., Asami, T., Alonso, J., Ecker, J.R., and Chory, J. (2003). Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrin IX. *Nature* 421, 79–83. <https://doi.org/10.1038/nature01204>.
- Tanifuji, G., Kamikawa, R., Moore, C.E., Mills, T., Onodera, N.T., Kashiwaga, Y., Archibald, J.M., Inagaki, Y., and Hashimoto, T. (2020). Comparative Plastid Genomics of *Cryptomonas* Species Reveals Fine-Scale Genomic Responses to Loss of Photosynthesis. *Genome Biol. Evol.* 12, 3926–3937. <https://doi.org/10.1093/gbe/evaa001>.
- Barrett, C.F., Freudenstein, J.V., Li, J., Mayfield-Jones, D.R., Perez, L., Pires, J.C., and Santos, C. (2014). Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Mol. Biol. Evol.* 31, 3095–3112. <https://doi.org/10.1093/molbev/msu252>.
- Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Müller, K.F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* 76, 273–297. <https://doi.org/10.1007/s11103-011-9762-4>.
- Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-Length RNA-Seq from Single Cells Using Smart-Seq2. *Nat. Protoc.* 9, 171–181. <https://doi.org/10.1038/nprot.2014.006>.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. <https://doi.org/10.14806/ej.17.1.200>.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Pribelski, A.D., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Laetsch, D.R., and Blaxter, M.L. (2017). BlobTools: interrogation of genome assemblies. *F1000Res* 6. <https://doi.org/10.12688/f1000research.12232.1>.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. <https://doi.org/10.1038/nprot.2013.084>.
- Larralde, M. (2022). Pyrodigal: Python bindings and interface to Prodigal, an efficient method for gene prediction in prokaryotes. *J. Open Source Softw.* 7, 4296. <https://doi.org/10.21105/joss.04296>.
- Seemann, T. (2013). Barnap 0.9: rapid ribosomal RNA prediction. <https://github.com/tseemann/barnap>.
- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>.
- Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- Geneious Prime. 2024. From Data to Discovery—Faster. <https://www.geneious.com>.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., and Teeling, E. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1530–1534. <https://doi.org/10.1093/molbev/msaa015>.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. <https://doi.org/10.1093/molbev/msu300>.
- Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
- Rambaut, A. (2007). FigTree, a graphical viewer of phylogenetic trees.
- Roure, B., Rodríguez-Ezpeleta, N., and Philippe, H. (2007). SCaFoS: A tool for selection, concatenation and fusion of sequences for phylogenomics. *BMC Evol. Biol.* 7, S2. <https://doi.org/10.1186/1471-2148-7-S1-S2>.
- Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and

- molecular dating. *Bioinformatics* 25, 2286–2288. <https://doi.org/10.1093/bioinformatics/btp368>.
40. Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 52, W78–W82. <https://doi.org/10.1093/nar/gkae268>.
  41. Eddy, S.R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7, e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
  42. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 51, D587–D592. <https://doi.org/10.1093/nar/gkac963>.
  43. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., et al. (2023). InterPro in 2022. *Nucleic Acids Res.* 51, D418–D427. <https://doi.org/10.1093/nar/gkac993>.
  44. Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large data sets. *Bioinformatics* 30, 3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>.
  45. Dierckxsens, N., Mardulyn, P., and Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45, e18. <https://doi.org/10.1093/nar/gkw955>.
  46. Lang, B.F., Beck, N., Prince, S., Sarrasin, M., Rioux, P., and Burger, G. (2023). Mitochondrial genome annotation with MFannot: a critical analysis of gene identification and gene model prediction. *Front. Plant Sci.* 14, 1222186. <https://doi.org/10.3389/fpls.2023.1222186>.
  47. Poux, S., Arighi, C.N., Magrane, M., Bateman, A., Wei, C.H., Lu, Z., Boutet, E., Bye-A-Jee, H., Famiglietti, M.L., Roechert, B., et al. (2017). On expert curation and scalability: UniProtKB/Swiss-Prot as a case study. *Bioinformatics* 33, 3454–3460. <https://doi.org/10.1093/bioinformatics/btx439>.
  48. Burki, F., Kaplan, M., Tikhonenkov, D.V., Zlatogursky, V., Minh, B.Q., Radaykina, L.V., Smirnov, A., Mylnikov, P., and Keeling, P.J. (2016). Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc. R. Soc. Lond. B* 283, 20152802. <https://doi.org/10.1098/rspb.2015.2802>.
  49. Wang, H.C., Minh, B.Q., Susko, E., and Roger, A.J. (2018). Modeling Site Heterogeneity with Posterior Mean Site Frequency Profiles Accelerates Accurate Phylogenomic Estimation. *Syst. Biol.* 67, 216–235. <https://doi.org/10.1093/sysbio/syx068>.
  50. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. <https://doi.org/10.1038/nmeth.4285>.

**STAR★METHODS**

**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Chemicals, peptides, and recombinant proteins</b>		
Triton X-100	Sigma-Aldrich	Cat# T9284
dNTP mix	ThermoFisher	Cat# R0192
First-strand buffer	ThermoFisher	Cat# 18064-014
Superscript II reverse transcriptase	ThermoFisher	Cat# 18064-014
Recombinant Ribonuclease Inhibitor	ThermoFisher	Cat# 10777019
Betaine	Sigma-Aldrich	Cat# B0300
Magnesium chloride	Sigma-Aldrich	Cat# M8266
KAPA HiFi HotStart ReadyMix (2x)	Fisher Scientific	Cat# 50-196-5299
Ampure XP beads	ThermoFisher	Cat# Q32854
UltraPure DNase/RNase-Free Distilled Water	ThermoFisher	Cat# 10977023
Ethyl alcohol anhydrous	Greenfield Global	Cat# P006EAAN
DTT	ThermoFisher	Cat# 18064-014
TRIzol Reagent	Invitrogen	Cat# 15596026
Chloroform	Fisher Scientific	Cat# C607-1
RNAlater stabilization solution	ThermoFisher	Cat# AM7020
Percoll	Sigma-Aldrich	Cat# P1644
Glutaraldehyde	Electron Microscopy Sciences	Cat# 16220
Piperazine-N,N'-bis(2-ethansulfonic acid) (PIPES)	ThermoFisher Scientific	Cat# 172615000
Spurr's resin	Electron Microscopy Sciences	Cat# 14300
Osmium tetroxide	Ted Pella	Cat# 18459
Formvar	Electron Microscopy Sciences	Cat# 15800
Uranyl acetate	Electron Microscopy Sciences	Cat# 22400
Lead citrate	Agar Scientific	Cat# AGR1210
<b>Critical commercial assays</b>		
Nextera XT	Illumina	Cat# FC-131-1024
TruSeq stranded mRNA	Illumina	Cat# 20020594
Illumina stranded mRNA	Illumina	Cat# 20040532
<b>Deposited data</b>		
Raw sequencing reads	This paper	[NCBI BioProject]:[PRJNA1199006]
SSU sequences	This paper	[GenBank]: [PQ863690-PQ863710]
LSU sequences	This paper	[GenBank]: [PQ871568-PQ871573, PQ872228-PQ872240]
Experimental models: Organisms/strains	This paper	N/A
Corallicolid ex. <i>Madracis mirabilis</i>	This paper	N/A
<i>Anthozoaphila gnarlus</i> ex. <i>Parazoanthus swiftii</i>	This paper	N/A
<b>Oligonucleotides</b>		
Oligo-dT30VN (50-AAGCAGTGGTATCAAC GCAGAGTACT30VN-30)	Picelli et al. <sup>23</sup>	N/A
IS-PCR oligo (50-AAGCAGTGGTATCAACG CAGAGT-30)	Picelli et al. <sup>23</sup>	N/A
TSO (50-AAGCAGTGGTATCAACGCAGAGT ACATrGrG+G-30)	Picelli et al. <sup>23</sup>	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Cutadapt	Martin <sup>24</sup>	<a href="http://code.google.com/p/cutadapt/">http://code.google.com/p/cutadapt/</a>
rnaSPAdes	Bankevich et al. <sup>25</sup>	<a href="http://bioinf.spbau.ru/spades">http://bioinf.spbau.ru/spades</a>
BLAST	Altschul et al. <sup>26</sup>	<a href="https://blast.ncbi.nlm.nih.gov/">https://blast.ncbi.nlm.nih.gov/</a>
Blobtools	Laetsch and Blaxter <sup>27</sup>	<a href="https://blobtools.readme.io/docs/what-is-blobtools">https://blobtools.readme.io/docs/what-is-blobtools</a>
TransDecoder	Haas et al. <sup>28</sup>	<a href="https://github.com/TransDecoder/TransDecoder/wiki">https://github.com/TransDecoder/TransDecoder/wiki</a>
Pyrodigal	Larralde <sup>29</sup>	<a href="https://pypi.org/project/pyrodigal/0.4.0/">https://pypi.org/project/pyrodigal/0.4.0/</a>
Barnap	Seemann <sup>30</sup>	<a href="https://github.com/tseemann/barnap">https://github.com/tseemann/barnap</a>
MAFFT	Katoh and Standley <sup>31</sup>	<a href="https://mafft.cbrc.jp/alignment/software/">https://mafft.cbrc.jp/alignment/software/</a>
trimAl	Capella-Gutiérrez <sup>32</sup>	<a href="http://trimal.cgenomics.org/">http://trimal.cgenomics.org/</a>
Geneious	Prime <sup>33</sup>	<a href="https://www.geneious.com">https://www.geneious.com</a>
IQ-Tree2 and IQ-Tree	Minh et al. <sup>34</sup> and Nguyen et al. <sup>35</sup>	<a href="http://www.iqtree.org/">http://www.iqtree.org/</a>
RAxML	Stamatakis <sup>36</sup>	<a href="https://github.com/stamatak/standard-RAxML">https://github.com/stamatak/standard-RAxML</a>
FigTree	Rambaut <sup>37</sup>	<a href="https://beast.community/figtree">https://beast.community/figtree</a>
SCaFoS	Roure et al. <sup>38</sup>	<a href="http://megasun.bch.umontreal.ca/Software/scafos/scafos.html">http://megasun.bch.umontreal.ca/Software/scafos/scafos.html</a>
Phylobayes	Lartillot et al. <sup>39</sup>	<a href="http://www.atgc-montpellier.fr/phylobayes/">http://www.atgc-montpellier.fr/phylobayes/</a>
iTOL	Letunic and Bork <sup>40</sup>	<a href="https://itol.embl.de">https://itol.embl.de</a>
HMMER	Eddy <sup>41</sup>	<a href="http://hmmer.org">http://hmmer.org</a>
KEGG KASS	Kanehisa et al. <sup>42</sup>	<a href="https://www.genome.jp/kegg/kaas/">https://www.genome.jp/kegg/kaas/</a>
InterPro	Paysan-Lafosse et al. <sup>43</sup>	<a href="https://www.ebi.ac.uk/interpro/">https://www.ebi.ac.uk/interpro/</a>
AliView	Larsson et al. <sup>44</sup>	<a href="https://ormbunkar.se/aliview/">https://ormbunkar.se/aliview/</a>
NOVOplasty	Dierckxsens et al. <sup>45</sup>	<a href="https://github.com/ndierckx/NOVOplasty">https://github.com/ndierckx/NOVOplasty</a>
MFannot	Lang et al. <sup>46</sup>	<a href="https://github.com/BFL-lab/Mfannot">https://github.com/BFL-lab/Mfannot</a>

**EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**

All samples from coral hosts *Madracis mirabilis* and *Parazoanthus swiftii* (and sponge host) were collected off the shores of Curaçao between October 2021 and February 2023. Hosts were isolated from 18–23 meters depth and maintained with aeration at room temperature in seawater. Specific dates for each sample are documented in [Table S1](#).

**METHOD DETAILS**

**Sample isolation and corallicolid enrichments**

To isolate corallicolids, the coral hosts were washed with sterile 0.22 μm filtered seawater and gently crushed with a pestle and mortar. Debris and liquid from the crushed tissue were filtered through consecutive 500 μm, 300 μm, 100 μm and then 50 μm cell strainers (pluriStrainer) under gentle pressure using a 50ml conical tube syringe adapter. Individual corallicolid cells were isolated from the slurry by micropipetting under microscopy and washed using 0.2 μm filter-sterilized seawater. Pooled single cells were preserved in 2 μl cell lysis buffer (as described in Picelli et al.<sup>23</sup>) and stored at -70°C for single cell transcriptomics.<sup>23</sup>

To prepare the coral enrichments, Percoll (Sigma-Aldrich) gradients consisting of three layers were created as follows, from bottom to top: 1) 70% Percoll layer (7:1:2 of Percoll:10× saltwater:dH<sub>2</sub>O), 2) low-salinity 58% Percoll layer (7:1:4 of Percoll:10× saltwater:dH<sub>2</sub>O), 3) 40% Percoll layer (4:1:5 of Percoll:10× saltwater:dH<sub>2</sub>O). The 10× saltwater was made by adding 41% w/v reef salt (Instant Ocean Reef Crystals) to dH<sub>2</sub>O. The crushed coral slurry filtrate was layered on top of the Percoll layers and centrifuged at 1,500g for 15 min. After centrifugation, coral cell debris were largely trapped above the low-salinity layer, and corallicolid cells were isolated from the interface of layers 1 and 2 and within layer 1. The cells were then spun down at 300g for 5 min, washed and resuspended in RNAlater and stored at -70°C. Additionally, whole polyps of *P. swiftii* and fragments of the host sponge both were preserved separately in 1.5 ml of RNAlater and stored at -70°C.

**Electron microscopy**

Electron microscopy samples were prepared from density-gradient enriched corallicolid pellets and fixed in 25% glutaraldehyde:0.2M PIPES buffer:dH<sub>2</sub>O (1:5:4) for 20 min. The sample was washed 3 times using 0.1 M PIPES buffer at 400g for 5 min then resuspended and stored in 0.1M PIPES buffer at 4°C. Fixed and pelleted corallicolids were post-fixed in 1% OsO<sub>4</sub> (w/v) for 30 min at

room temperature. The osmicated pellet was rinsed three times in room temperature 0.05M PIPES buffer, twice in dH<sub>2</sub>O, then dehydrated with an ascending graded ethanol series of 30%, 50%, 70%, 95% and 100% for 10 min each, with two additional 100% ethanol exchanges. Spurr's resin was gradually infiltrated in an ascending series of 10%, 20%, 40%, 60%, 80%, 95% and 100% resin for at least 1h each. Two additional 100% resin exchanges were performed for 12 h each, then samples were transferred to Beem capsules (Ted Pella) containing fresh 100% Spurr's resin and polymerized at 60°C for 48 h. Silver-gold sections (60 nm) were cut using a Leica UC7 ultramicrotome and mounted on copper slot grids coated with 0.3% formvar (w/v). Sections were post-stained with 2% aqueous uranyl acetate (w/v) and lead citrate for 12 and 6 min, respectively. Imaging of sections was performed using a Tecnai Spirit transmission electron microscope operating at 80 kV. Images were acquired on an AMT XR51 CCD camera or an FEI Eagle CCD camera. Images were collected from two independent replicates of chemical fixation for each species.

### Sequencing and transcriptome assembly

Samples containing density-gradient enriched corallicolids and whole *P. swiftii* polyps and host sponge were thawed on ice and centrifuged at 14,000g for 5 min to remove RNAlater supernatant. The resulting enriched cell pellet was processed using the TRIzol (Invitrogen) method for RNA purification. For samples labelled ZCRNA and ZCDNA, RNA and DNA were co-extracted from four respective cell enrichments using the TRIzol method (Table S1). The host sponge, whole polyp, and the enriched *P. swiftii* RNA samples were processed for mRNA sequencing with Illumina TruSeq mRNA stranded protocol, whereas the enriched *M. mirabilis* RNA samples were processed using the Illumina Stranded mRNA protocol.

Samples preserved in lysis buffer for single cell transcriptomics underwent three freeze-thaw cycles prior to mRNA extraction and cDNA synthesis following the SmartSeq2 protocol.<sup>23</sup> Resulting cDNA were processed into libraries using the Nextera XT kit. DNA and cDNA libraries were sequenced on MiSeq and/or NextSeq platforms (raw sequencing reads available in NCBI BioProject: PRJNA1199006) at the Sequencing and Bioinformatics Consortium, University of British Columbia.

Adapters on forward and reverse raw reads were trimmed with Cutadapt v3.2<sup>24</sup> and sequences assembled using maSPAdes v3.15.1.<sup>25</sup> The metatranscriptomes of the zoanthid, *P. swiftii* and host sponge were used as queries against NCBI's nr database using BLASTn<sup>26</sup> to curate a database of Cnidaria and Porifera contigs. The *P. swiftii* and sponge sequences were additionally used to remove Cnidaria and Porifera contigs from the single cell transcriptomes of *Anthozoophila gnarlus* ex. *P. swiftii*. Using the same BLAST method, apicomplexan genes from the *P. swiftii* metatranscriptomes were identified and are available on UBC Borealis: <https://doi.org/10.5683/SP3/2LXG4D>. Further contamination was identified using BLASTx and BLASTn against NCBI nt and Uniprot<sup>47</sup> (The UniProt Consortium 2021) databases and Cnidaria, bacteria and Chordata contamination was removed using BlobTools v3.3.4.<sup>27</sup> The cleaned transcripts were then used in Transdecoder v5.5.0 for proteome prediction (<https://github.com/TransDecoder/TransDecoder/wiki>).<sup>28</sup> The longest ORFs with shared similarity (e-value threshold  $\leq 1e10^{-5}$ ) to annotated proteins in the UniProt database were kept and translated into proteins. Assemblies are available on UBC Borealis: <https://doi.org/10.5683/SP3/2LXG4D>. A predicted proteome was generated from the previously assembled metagenome of ichthyocolid-infected *Ophioblennius macclurei* (PRJNA1041302) blood using Pyrodigal.<sup>29</sup>

### Orthology identification and phylogenomic analysis

To confirm the presence of corallicolid transcripts in each transcriptome, Barrnap v0.9<sup>30</sup> was used to isolate nuclear SSU and LSU rRNA sequences. The taxonomic identity of all rRNA sequences was assessed by BLASTn against the NCBI nt database. The longest apicomplexan positive SSU and LSU sequences were collected and added to a curated database of apicomplexan SSU and LSU sequences and aligned using MAFFT v7.471<sup>31</sup> using auto settings and trimmed using trimAl v1.4 (-c 60 -gt 90).<sup>32</sup> Trimmed SSU and LSU alignments were inspected and concatenated into single alignment using Geneious<sup>33</sup> and used to generate a maximum-likelihood (ML) tree under the GTR model and 1,000 non-parametric bootstraps using IQTree2<sup>34</sup> (Figure S2A). To assess the SSU and LSU sequence variability in each transcriptome, a pairwise alignment was run on all apicomplexan-positive sequences and the percent identities to the longest, representative sequence were collected (available on UBC Borealis: <https://doi.org/10.5683/SP3/2LXG4D>).

The multi-protein phylogenetic analysis was performed using a previously published curated dataset of 263 highly conserved protein-encoding genes.<sup>48</sup> The 263 genes were used as queries against the corallicolid and ichthyocolid predicted proteomes using BLASTp. Gene hits were added to their respective fasta files and aligned with auto settings in MAFFT (L-INS-i), trimmed with trimAl (-gt 60) and ML single gene trees were created using RAxML (raxmlHPC-PTHREADS -p 123456789 -m PROTGAMMALG -f a -x 123 -N 100).<sup>34</sup> RAxML trees were visually inspected using FigTree v1.4.4.<sup>37</sup> to identify and remove contaminants, paralogs and isoforms while preserving the single longest corallicolid sequence. For the phylogenomic tree, the cleaned 263 gene dataset was parsed using SCaFoS v1.25<sup>38</sup> to extract genes from only apicomplexan and selected outgroup taxa and genes represented in less than 40% of taxa were removed from the analysis. A maximum-likelihood analysis was run on the resulting concatenated 201 gene dataset using IQTree2 under the model LG+C60+F+G with 1000 ultrafast bootstraps (-B 1000). The resulting tree was used as a guide tree for the Posterior Mean Site Frequency model (PMSF)<sup>49</sup> in IQTree2 with robustness assessed by 100 non-parametric bootstraps (-b 100). Additionally, a Bayesian analysis was run on the cleaned concatenated gene dataset using PhyloBayes-MPI v20201026 using the CAT-GTR model.<sup>39</sup> Four chains were run in parallel, with a sampling frequency every 5 trees until each chain surpassed 8,000 iterations. Only two of the four chains converged with a max difference of 0.03. A posterior consensus tree was generated after removing the first 10% of trees calculated from the four parallel runs (available on UBC Borealis: <https://doi.org/10.5683/SP3/2LXG4D>) (Figure S2B). Due to the lower recovery of ichthyocolid genes from the metagenomic assembly, a second maximum-likelihood and

Bayesian analysis was run on a concatenated 115 gene dataset which prioritized genes present in both the corallicolids and ichthyocolid data using the same IQTree2 and Bayesian parameters (Figures S2B and S2C). All 4 chains surpassed 6,000 iterations and converged with a max difference of 0.178. All finalized trees were viewed and exported using iTOL.<sup>40</sup>

### Plastid protein search and de novo plastid genome assembly

We searched our corallicolid transcriptomes for plastid-derived proteins comprising biosynthetic pathways characteristic of apicomplexan cryptic plastids, specifically isoprenoids, FeS clusters, fatty acids, and heme synthesis. To investigate the presence of chlorophyll biosynthesis, corallicolids and apicomplexan relatives, including *Eleutheroschizon duboscqui* (SRR9888047) and ichthyocolids, were searched using BLASTp and HMMER3.<sup>41</sup> Genes from these pathways that were absent in our analyses of *E. duboscqui* were cross referenced with the original publication,<sup>10</sup> and were included in our report if it was originally reported as present. Queries used in the BLASTp search (e-value threshold:  $\leq e10^{-3}$  and -per\_qcov 0.3) were collected from the Kyoto Encyclopedia of Genes and Genomes (KEGG).<sup>50</sup> The HMMER protein search were created using alignments from KEGG queries against a diverse in-house eukaryotic and prokaryotic proteomic dataset. A comprehensive search for chlorophyll biosynthesis genes also included using BLASTp against uncleaned peptide files and translating chlorophyll biosynthesis proteins into their nucleotide sequence and using BLASTn against uncleaned nucleotide sequences. Transcriptome contigs corresponding to positive protein-level hits were retrieved, and trimmed sequencing reads were mapped against these contigs for verification and to determine read coverage.

Resulting genes from plastid metabolic pathways and the chlorophyll biosynthesis pathway were combined with query proteins, aligned and trimmed using auto settings in MAFFT and trimAl followed by ML tree reconstruction using IQTree2<sup>41</sup> under the LG+I+G4, LG+G4, WAG+F+G4, WAG+F+I+G4 models (available at UBC Borealis: <https://doi.org/10.5683/SP3/2LXG4D>) (determined by ModelFinder). Trees were inspected manually using FigTree to remove paralogs, contamination and isoforms. Cleaned, trimmed alignments were rerun using IQTree2 with the same approach. Sequences that were highly divergent and branching outside apicomplexans or chrompodellids were investigated using InterPro and BLASTp against the NCBI's nr database.<sup>43</sup> The N-terminus of nuclear-encoded plastid-targeted proteins was investigated for targeting signals by visually inspecting corallicolid sequences the gene alignment in AliView v1.28.<sup>44</sup>

The *de novo* assembly of the *Eleutheroschizon duboscqui* plastid genome was produced from high coverage transcriptomic data using the *E. duboscqui* plastid-encoded gene *tufA* as a seed query in NOVOplasty.<sup>45</sup> The resulting contig was annotated with MFannot<sup>46</sup> (<https://megasun.bch.umontreal.ca/apps/mfannot/>) and visually inspected using Geneious (Figure S3).

### QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical support for the ML analysis included a PMSF analysis with 100 non-parametric bootstraps in IQTree2. Using the same concatenated 201-gene alignment, a Bayesian analysis was run in PhyloBayes with four parallel runs using the model CAT-GTR. An approximately unbiased (AU) test was run on the 201-gene alignment using IQTree model LG+C60+F+G.<sup>35</sup>