


BRIEF REPORT **OPEN ACCESS**

Hepatincolaceae (Alphaproteobacteria) are Distinct From *Holosporales* and Independently Evolved to Associate With Ecdysozoa

Michele Castelli¹  | Leandro Gammuto¹ | Diona Podushkina¹ | Matteo Vecchi^{2,3} | Tiziana Altiero⁴ | Emanuela Clementi¹ | Roberto Guidetti² | Lorena Rebecchi² | Davide Sassera^{1,5}

¹Department of Biology and Biotechnology, University of Pavia, Pavia, Italy | ²Dipartimento di Scienze Della Vita, Università degli Studi di Modena e Reggio Emilia, Modena, Italy | ³Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Krakow, Poland | ⁴Dipartimento Educazione e Scienze Umane, Università degli Studi di Modena e Reggio Emilia, Modena, Italy | ⁵Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

Correspondence: Michele Castelli (michele.castelli@unipv.it)

Received: 7 June 2024 | **Revised:** 11 November 2024 | **Accepted:** 3 December 2024

Funding: This work was supported by European Union—NextGenerationEU, Italian NRPP Mission 4 Component 2 Investment 1.4.

Keywords: bacterial parasites | gut bacteria | prophages | protists | *Rickettsiales* | symbiosis | tardigrades

ABSTRACT

The *Hepatincolaceae* (*Alphaproteobacteria*) are a group of bacteria that inhabit the gut of arthropods and other ecdysozoans, associating extracellularly with microvilli. Previous phylogenetic studies, primarily single-gene analyses, suggested their relationship to the *Holosporales*, which includes intracellular bacteria in protist hosts. However, the genomics of *Hepatincolaceae* is still in its early stages. In this study, the number of available *Hepatincolaceae* genomes was increased to examine their evolutionary and functional characteristics. It was found that the previous phylogenetic grouping with *Holosporales* was incorrect due to sequence compositional biases and that *Hepatincolaceae* form an independent branch within the *Hepatincolaceae*. This led to a reinterpretation of their features, proposing a new evolutionary scenario that involves an independent adaptation to host association compared to the *Holosporales*, with distinct specificities. The *Hepatincolaceae* exhibit greater nutritional flexibility, utilising various molecules available in the host gut and thriving in anaerobic conditions. However, they have a less complex mechanism for modulating host interactions, which are likely less direct than those of intracellular bacteria. In addition, representatives of *Hepatincolaceae* show several lineage-specific traits related to differences in host species and life conditions.

1 | Introduction

Associations between prokaryotes and eukaryotes are widespread and diverse (McFall-Ngai et al. 2013; Drew, Stevens, and King 2021; Husnik et al. 2021). Many phylogenetically and physiologically diverse bacteria have evolved metabolic dependence on their hosts, accompanied by marked genome size reduction (McCutcheon and Moran 2012; Bennett and Moran 2015; Wernegreen 2017).

In particular, the *Alphaproteobacteria* encompass multiple obligate host-associated lineages (Hördt et al. 2020), including the *Rickettsiales* and the *Holosporales* (Muñoz-Gómez et al. 2019; Castelli et al. 2024). Curiously, early molecular phylogenies on 16S rRNA gene sequences lumped those two lineages together, with the *Holosporales* as early diverging among the *Rickettsiales* (e.g., Wang et al. 2004). More recent analyses based on multiple markers allowed to establish that this grouping was artefactual (Georgiades et al. 2011; Ferla

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Environmental Microbiology* published by John Wiley & Sons Ltd.

et al. 2013; Schulz et al. 2014; Wang and Wu 2015), due to convergently evolved sequence compositional biases (i.e., low GC content) as a consequence of obligate associations with their hosts (Muñoz-Gómez et al. 2019). Specifically, the *Rickettsiales* are an ancient and early diverging lineage within *Alphaproteobacteria* (Wang and Luo 2021), while the *Holosporales* have originated independently and more recently within the *Rhodospirillales sensu* Hördt et al. (2020). Based on this phylogenetic position, it was even proposed to lower *Holosporales* at the family rank within *Rhodospirillales* (Muñoz-Gómez et al. 2019). Here, we will follow the definition of the order *Holosporales* by Szokoli et al. (2016) since it allows for the focus on the diversity and evolution of its main sub-lineages, which are accordingly ranked as families.

The *Holosporales* include four families (Hess, Suthaus, and Melkonian 2016; Szokoli et al. 2016; Schrallhammer, Castelli, and Petroni 2018). Three of them, namely *Holosporaceae*, ‘*Caedimonadaceae*’, and ‘*Candidatus Paracaedibacteraceae*’ (from now on, *Candidatus* will be abbreviated as *Ca.*), encompass bacteria that are intracellularly associated with their eukaryotic hosts, typically protists (Horn et al. 1999; Birtles et al. 2000; Baker et al. 2003; Dirren and Posch 2016; Hess, Suthaus, and Melkonian 2016; Szokoli et al. 2016; Potekhin et al. 2018; Boscaro et al. 2019; George et al. 2020; Midha et al. 2021; Zilio et al. 2021; Castelli et al. 2022; Shiohama et al. 2022; Lanzoni et al. 2024), with few exceptions of arthropods (Jones, McCormick, and Martin 2008; Nunan et al. 2013; Konecka and Olszanowski 2019). These protist-associated *Holosporales* include very peculiar representatives such as infectious ones (Schulz et al. 2014; Fokin et al. 2019; Beliavskaia et al. 2020; Schrallhammer and Potekhin 2020), and others able to kill uninfected hosts (Schrallhammer and Schweikert 2009; Schrallhammer, Castelli, and Petroni 2018). In the last few years, the representatives of these families have been the subject of multiple genomic and phylogenomic studies (Georgiades et al. 2011; Wang and Wu 2015; George et al. 2020; Midha et al. 2021; Castelli et al. 2022; Shiohama et al. 2022; Giovannini, Petroni, and Castelli 2024).

On the other hand, members of the fourth family (named ‘*Ca. Hepatincolaceae*’ (Szokoli et al. 2016) or ‘*Ca. Tenuibacteraceae*’ (Kroer et al. 2016); from now on, *Hepatincolaceae*) have been found so far hosted only by ecdysozoan animals. Such hosts are mostly arthropods, including insects (Ramírez-Puebla et al. 2010; Koch et al. 2013), crustaceans (Wang et al. 2004; Li et al. 2007; Bauermeister, Ramette, and Dattagupta 2012; Dittmer et al. 2023), arachnids (Qu et al. 2015) and myriapods (Chipman et al. 2014), as well as tardigrades (Vecchi et al. 2018; Guidetti et al. 2020) and priapulids (Kroer et al. 2016). Based on the available data, the *Hepatincolaceae* dwell within the host gut lumen (Li et al. 2007; Koch et al. 2013; Bouchon, Zimmer, and Dittmer 2016), being extracellularly associated with host microvilli (Wang et al. 2004; Kroer et al. 2016) and likely exploiting this condition to scavenge nutrients (Dittmer et al. 2023).

To sum up, *Hepatincolaceae* are quite sharply distinguished from the other *Holosporales*, in particular, for host lineages and extracellular location. Phylogenetic analyses on 16S rRNA gene sequences most frequently placed them as early diverging

within the *Holosporales* (Kroer et al. 2016; Guidetti et al. 2020; Lanzoni et al. 2024). However, they are still poorly investigated at the genomic and phylogenomic level, as the first three genome sequences of this family, all belonging to the genus ‘*Ca. Hepatincola*’ (from now on, *Hepatincola*), were only recently published (Dittmer et al. 2023).

Here, we present an extended genome sampling of the *Hepatincolaceae*, which allowed us to infer a novel phylogenetic and evolutionary framework. Specifically, we newly sequenced the symbiont of the tardigrade *Richtersius cf. coronifer* (for a better definition of the tardigrade taxon and its symbionts see Guidetti et al. 2016, 2020; Vecchi et al. 2018), and assembled the genome of the symbiont of the myriapod *Strigamia maritima* from reads that were previously obtained for sequencing the genome of its host (Chipman et al. 2014). Moreover, we determined that a previously published genomic assembly of a bacterial symbiont of the termite *Labiatermes labralis* (Hervé et al. 2020) can be ascribed to the *Hepatincolaceae*, so we included it in our analyses. Leveraging this extended sampling and accounting for compositional heterogeneity biases, we re-assessed the phylogenetic positioning of the *Hepatincolaceae*, determining that they are phylogenetically distinct from the ‘true’ *Holosporales sensu* Szokoli et al. (2016), thus representing an evolutionarily independent obligatorily host-associated lineage within the *Rhodospirillales sensu* Hördt et al. (2020).

2 | Materials and Methods

2.1 | Genome Sequencing, Assembly and Annotation of the Symbiont of *R. cf. coronifer*

The starting point was a *R. cf. coronifer* population originally isolated from the moss in Öland, Sweden (Vecchi et al. 2018). DNA extraction was conducted from multiple single individuals, as previously described (Guidetti et al. 2020). Then, semi-quantitative PCR (i.e., on original samples and the respective serial 1:10 dilutions) was performed with previously defined primers and cycling conditions for the SSU rRNA genes of the bacterium (Guidetti et al. 2020) and of its host (Bertolani et al. 2014), respectively. Accordingly, the specimen with the highest estimated bacterium/host ratio was subjected to whole-genome amplification with the REPLI-g Single Cell Kit (QIAGEN). Then, total DNA sequencing was carried out through a Nextera XT library on an Illumina HiSeq X machine by Admera Health (South Plainfield, NJ, USA), producing 44,276,484 pairs of 150 bp reads.

After an assessment with FastQC (Andrews 2010) that confirmed the high quality of the reads, these were directly assembled with SPAdes 3.6 (Bankevich et al. 2012), thus obtaining 86,693 contigs (296,744,432 bp). The choice of SPAdes over other software more tailored for metagenomes (e.g., metaSPAdes [Nurk et al. 2017]) was due to a previous successful experience in alike host-symbiont datasets (e.g., Castelli et al. 2019, 2022, 2024; Giovannini, Petroni, and Castelli 2024). After this preliminary assembly, a multi-step procedure was applied (Castelli et al. 2019), to select the sequences belonging to the bacterial symbiont based on the blobology pipeline (Kumar et al. 2013). This approach, feasible in low-scale studies such as herein, allows an effective

and accurate genome assembly of bacteria from low-complexity metagenomes, such as the analysed herein host metagenomes (see also Section 3 for estimates on genome completeness and contamination). Accordingly, the contigs were classified as per their length, GC% content, NCBI taxonomy of the best megablast hit on the NCBI nucleotide database, and sequencing coverage based on reads mapped on the novel assembly with Bowtie2 (Langmead and Salzberg 2012). Considering also the rRNA genes identified with barrnap (Table S1) (Seemann 2013), a selection of contigs with coverage higher than 100 was performed (Figure S1), and the respective mapping reads (Langmead and Salzberg 2012) were reassembled separately with SPAdes, using the option -k 21,33,55,77,99,121. The selected contig sequences were manually examined and revised prior to and after reassembly by examining blastp hits on NCBI nr protein database after annotation with Prokka 1.10 (Seemann 2014), and using Bandage (Wick et al. 2015) for visualisation of SPAdes assembly graphs.

Moreover, to rule out the presence of additional sequences belonging to the symbiont, the annotated ORFs on contigs with coverage 10–100 were queried on the NCBI nr protein database with DIAMOND (Buchfink, Xie, and Huson 2015), applying an e-value threshold of $1e-5$. DIAMOND results for contigs with at least one ORF with the best hit on *Bacteria* were inspected manually.

Then, primers were designed on selected contig ends (Table S2), and PCR reactions were performed with the Takara ExTaq (Takara Bio, Japan) or with the PCR BIO VeriFi polymerase (PCR Biosystems, London, UK). Products were Sanger sequenced at Eurofins Genomics (Ebersberg, Germany).

The final genome assembly was annotated with Prokka, and results were manually curated by inspecting the blastp hits of the predicted proteins on NCBI nr and on a custom database of *Holosporales*, as well as on NCBI conserved domains (Lu et al. 2020).

2.2 | Selection and Assembly of the *Hepatincolaceae* Symbiont of *S. maritima*

We aimed to expand the set of genome assemblies of *Hepatincolaceae* from published sequencing data besides the three available *Hepatincola* genomes (Dittmer et al. 2023). The first strategy involved inspection of the SILVA taxonomy (release 138.1) (Quast et al. 2013) and the reference 16S rRNA gene tree in ARB (Ludwig et al. 2004), looking for relatives of *Hepatincola porcellionum* (AY188585). Accordingly, a 16S rRNA gene sequence was identified (AFFK01003480) as being derived from the genome assembly project of the myriapod *S. maritima* (Chipman et al. 2014), suggesting the presence of a putative *Hepatincolaceae* symbiont. The corresponding reads (SRX326837, SRX326839, SRX326840 and SRX326841) were downloaded, combined and processed to selectively assemble the bacterial genome as described above for the symbiont of *R. cf. coronifer*, with some differences, presented below. The preliminary assembly counted 600,069 contigs (248,668,256 bp). In this case, to increase the specificity and sensitivity of the taxonomic

assignment, a DIAMOND blastx search of the whole set of preliminary contigs on NCBI nr was used instead of a blastn search on NCBI nucleotide. The contigs with \log_{10} coverage lower than 1.875, excluding those with best hits on eukaryotic sequences, were selected (Figure S2), and the corresponding mapped reads were reassembled with default SPAdes settings. The resulting reassembly was manually curated as described above. The second strategy consisted in testing the affiliated of published assemblies, and is described in the section 2.3.

2.3 | Phylogenomic Analyses

A comprehensive phylogenomic analysis was conducted to investigate the phylogeny of the *Hepatincolaceae*, including assigning potential published Metagenome-Assembled Genomes (MAGs) to this lineage and clarifying the position within *Alphaproteobacteria* of the *Hepatincolaceae* as a whole. For this purpose, the two assemblies obtained in this study and the three *Hepatincola* genomes (Dittmer et al. 2023) were merged into a representative set of *Alphaproteobacteria*. This set was obtained by taking genome assemblies from described alphaproteobacterial lineages (plus other *Proteobacteria* as outgroup) from a previous study (Castelli et al. 2024) (with the exclusion of *Rickettsiales* and ‘*Ca. Pelagibacterales*’, known to be involved in phylogenetic artefacts with *Holosporales* due to compositional biases and fast sequence evolution rates; Muñoz-Gómez et al. 2019) and a selection of additional assemblies (in particular MAGs) belonging to yet undescribed alphaproteobacterial lineages based on the GTDB taxonomy (r207 release) (Parks et al. 2022) (Table S3). The final set counted 179 organisms. This approach also allowed us to recognise that an additional published assembly (Hervé et al. 2020) is ascribable to the *Hepatincolaceae* (see Section 3 for details). The completeness and contamination degree of all *Hepatincolaceae* and alphaproteobacterial MAGs were verified with CheckM 1.2.3 (Parks et al. 2015).

Annotated protein sequences were predicted for each assembly in the dataset with Prokka. Phylogenomics was performed with a previously determined dataset of 179 alphaproteobacterial orthologs (Castelli et al. 2024), selected from eggNOG orthogroups (Huerta-Cepas et al. 2019) predicted with eggNOG mapper (Cantalapiedra et al. 2021) as previously described (Castelli et al. 2024) (Table S4). The sequences of each of the identified 179 orthologs of interest were aligned with MAFFT 7.475 L-INS-i (Katoh and Standley 2013), and trimmed with BMGE 1.12 (Criscuolo and Gribaldo 2010) selecting the BLOSUM30 matrix, as recommended in the BMGE manual for distantly related sequences. Trimmed orthologs were concatenated together (Borowiec 2016). To account for potential artefacts due to compositional heterogeneity, the 10%, 20%, 30%, 40% or 50% most biased sites, identified as previously described (Muñoz-Gómez et al. 2019), were removed from the concatenated alignments. Maximum likelihood phylogenies were inferred on the original concatenated alignment and each compositionally trimmed alignment with IQ-TREE 1.6.12 (Nguyen et al. 2015) with the LG + C60 + F + R6 model as in (Muñoz-Gómez et al. 2019), performing 1000 ultra-fast bootstraps (Minh, Nguyen, and von Haeseler 2013) and SH-aLRT with 1000 replicates.

Average nucleotide identity (ANI) values among the *Hepatinoceae* were calculated with the EZBioCloud tool (Yoon et al. 2017). Moreover, the assignment of the *Hepatinoceae* to the GTDB taxonomic groups was obtained with GTDB-Tk 2.1.1 (Chaumeil et al. 2020).

2.4 | Investigations on *Hepatinoceae* Bacteria in the Microbiome of Tardigrades

To screen for the presence of relatives of the symbiont of *R. cf. coronifer* in previously published tardigrade microbiomes, we respectively downloaded already clustered OTUs, when available (Tibbs-Cortes, Tibbs-Cortes, and Schmitz-Esser 2022; Zawierucha et al. 2022), or sequencing reads (Kaczmarek et al. 2020; Mioduchowska et al. 2021; Boscaro et al. 2022) (Table S5). Reads were processed with QIIME2 (Bolyen et al. 2019) and de novo clustered into OTUs with 99% identity. OTU sequences aligned >90% of their length with at least 97% identity to the 16S rRNA gene of the symbiont of *R. cf. coronifer* were further manually verified by comparison with the NCBI nr database to assess their affiliation. For the phylogenetic analyses, 56 sequences of *Hepatinoceae* plus 10 *Rhodospirillales* as outgroups were automatically aligned with ARB 5.5 on the SSU ref. NR99 SILVA database 138.1, manually edited to optimise base pairing in the predicted rRNA structure, and trimmed at the shortest sequence length at both sides (final 1128 sites). The optimal substitution model was identified using jModelTest 2.1 (Darriba et al. 2012) according to the Akaike information criterion. A maximum likelihood tree was inferred with PHYML 2.4 (Guindon and Gascuel 2003) performing 1000 pseudo-replicates. Due to its short size, the retrieved OTU sequence related to the symbiont of *R. cf. coronifer* (see Section 3 for details) was added later using the quick-add marked function of the ARB package.

2.5 | Functional and Metabolic Predictions and Comparisons

Clusters of orthologous groups (COGs) was identified with the NCBI pipeline (Galperin et al. 2015) in the *Hepatinoceae* (the newly assembled symbiont of *R. cf. coronifer* and symbiont of *Strigamia*, the three *Hepatinoceae* spp., and the symbiont of *L. labralis* GCA_009780035.1), their closest relatives based on phylogeny (*Thalassospira profundimaris*, *Terasakiella pusilla*, symbiont of *Haliotis discus hannai*, GCA_001510075.1, GCA_001830425.1, GCA_002327565.1, GCA_002687515.1, GCA_009694195.1, GCA_009649675.1, GCA_013204045.1, GCA_014859895.1, GCA_018662225.1) and a representative set of *Holosporales* (*Holospira undulata*, ‘*Ca. Hepatobacter penaei*’, ‘*Ca. Bealeia paramacronuclearis*’, ‘*Caedimonas varicaedens*’, ‘*Ca. Nucleicultrix amoebiphila*’, ‘*Ca. Paracaedibacter acanthamoebae*’, ‘*Ca. Odyssella thessalonicensis*’, ‘*Ca. Finniella inopinata*’). The COG repertoire for each organism (or group of organisms) is herein defined as the number of unique COGs annotated in the respective gene set(s), regardless of how many genes were annotated to each COG or if some gene was annotated to multiple COGs. Such repertoires were then compared manually, in particular among *Hepatinoceae* and between *Hepatinoceae* and the other organisms (see Section 3 and Text S1 for details). Reference

metabolic pathways for such comparisons were taken from BioCyc (Karp et al. 2019) and KEGG (Kanehisa et al. 2016).

3 | Results

3.1 | Novel Genomes of *Hepatinoceae* Bacteria

In this work, we newly sequenced and assembled the genome of the *Hepatinoceae* symbiont of the tardigrade *R. cf. coronifer* (Guidetti et al. 2020), and assembled the genome of the *Hepatinoceae* symbiont of the centipede *S. maritima* from previously published reads (Chipman et al. 2014). Both assemblies had sizes and GC content comparable to the previously published genomes of *Hepatinoceae* (~1.3 Mb, and ~30 GC%; Table S14,S6). The genome assembly of the symbiont of *Richtersius* was quite contiguous (15 contigs, N50=996,816 bp, L50=1), with no unconnected contig ends in the assembly graph, and few long repeats (thousands of bp), concentrated in two genome areas (Figure S3), which likely represent as many putative prophages (see below). This suggests that this assembly is likely complete and clean, consistent with the high completeness (98.9%) and no contamination (0.0%) estimated by CheckM (Table S7). On the other hand, the assembly of the symbiont of *Strigamia* is more fragmented (227 contigs, N50=7795 bp, L50=49), with a slightly lower completeness score (86.8%) and a slightly higher, though negligible, contamination (1.1%) (Table S7). This potential incompleteness can be explained by the relatively lower sequencing coverage in sequencing reads, which were obtained in a study focused on the genomics of the arthropod host (Chipman et al. 2014). This could also explain why some genes common in the other *Hepatinoceae* could not be found in this bacterium (see Text S1).

3.2 | Phylogenomics

We then aimed to infer the phylogenetic relationship among the five *Hepatinoceae* with genome sequences available (i.e., the two novel ones and the three *Hepatinoceae* spp.; Dittmer et al. 2023), as well as of the *Hepatinoceae* with respect to the *Holosporales* and other *Alphaproteobacteria*.

In the phylogeny obtained from the original concatenated alignment, all the *Hepatinoceae* clustered together with full support (Figure S4). Moreover, a previously published MAG (GCA_009780035.1) was found to be nested within the same clade and was thus identified as a sixth additional member of the *Hepatinoceae*. The respective sample originated from the gut of the soil-feeding termite *L. labralis* (Hervé et al. 2020), which is comparable with the typical provenance of the members this bacterial lineage (i.e., gut of arthropods and other Ecdysozoa) (Wang et al. 2004; Kroer et al. 2016; Szokoli et al. 2016; Dittmer et al. 2023). This MAG belongs to the GTDB order and family WRAU01. Interestingly, GTDB-Tk assigned to the same lineage also the other five *Hepatinoceae* assemblies herein analysed, suggesting that WRAU01 could correspond to the *Hepatinoceae*, and may be renamed accordingly in the future. The inner relationships of the *Hepatinoceae* were fully supported as well, with *Hepatinoceae* Av and Pdp as closest relatives with respect to *Hepatinoceae* Pp, consistent with (Dittmer et al. 2023), the

symbiont of *Labiotermes* as the sister group of *Hepatincola* spp., and the symbionts of *Richtersius* and *Strigamia* in progressive sequential branching order. The high sequence divergence of each of the latter three bacteria (ANI always below 68%) suggests that they should be ascribed to as many new separate genera and species (Table S8) (Barco et al. 2020). For the symbiont of *Richtersius*, this is consistent with previous analyses on the 16S rRNA gene (Guidetti et al. 2020). Screening of tardigrade microbiomes for relatives of the symbiont of *Richtersius* allowed the identification of a single OTU phylogenetically related to this bacterium (OTU000635: length: 257 bp, 16S rRNA gene identity: 98.81%), indicating the probable presence of a clade of tardigrade-associated *Hepatincolaceae* (Figure S5). This OTU was retrieved in several samples from the same study (Tibbs-Cortes, Tibbs-Cortes, and Schmitz-Esser 2022) involving multiple tardigrade species (Table S9). Accordingly, the symbiont of *Richtersius* will be from now on referred to as *Tardigradibacter bertolanii* (see taxonomic description at the end of Section 4).

In the phylogenomic tree based on the original dataset, the *Hepatincolaceae* clade was retrieved in a quite early-divergent position within *Alphaproteobacteria*, close to a number of assemblies from various origins, but distant from the *Holosporales* (Figure S4). We then aimed to counterbalance the effects of GC/AT compositional biases on the phylogenetic inference by progressively removing the most biased sites from the alignment through an on-purpose site selection method (Muñoz-Gómez et al. 2019). When applying this approach, the *Hepatincolaceae* and their inner phylogenetic branches were stably supported, but the relationships with other *Alphaproteobacteria* drastically changed (Figures 1 and S4). Specifically, each of the alignments with 10% and 20% most biased sites removed produced ‘unique’ topologies, probably indicative of partial and ‘ongoing’ attenuation of compositional biases (with the 20% trimming resulting in a close *Hepatincolaceae*–*Holosporales* relationship, comparable to previous reconstructions for those lineages; Kroer et al. 2016; Szokoli et al. 2016; Dittmer et al. 2023). Whereas, in the three trees obtained with the more pronounced trimming (30%, 40% and 50% most biased sites removed), the *Hepatincolaceae* stably found the same novel phylogenetic ‘neighbourhood’ with high supports (always above 90% ultra-fast bootstrap). Accordingly, the *Hepatincolaceae* branched within the *Rhodospirillales sensu* Hördt et al. (2020), similar to the *Holosporales*, but in a distinct and distantly related position. While the *Holosporales* (which include multiple groups labelled as orders in GTDB; Figures 1 and S4) were found closely related to the *Kiloniellaceae*, *Rhodovibrionaceae*, *Thalassobaculaceae* and *Azospirillaceae*, consistent with (Muñoz-Gómez et al. 2019), the *Hepatincolaceae* were closely related to several genome assemblies originating from marine environments (Figures 1 and S4). In detail, the sister group of the *Hepatincolaceae* can be subdivided into two highly supported branches. The first branch is constituted of *T. pusilla* and a bacterial symbiont from the digestive gland of the mollusc *H. discus hannai* (Huang et al. 2020), while the second branch is constituted of eight marine MAGs. On the other hand, *T. profundimaris* and another marine MAG form the sister group of the lineage composed of the *Hepatincolaceae* and the closely related bacteria mentioned above. These relatives of the *Hepatincolaceae* are ascribed to the family *Thalassospiraceae* (and to the *Rhodospirillales_A* or

MC-1 order groups of GTDB; Figure 1). To sum up, the obtained topology showed a fully supported monophylum encompassing almost all the *Rhodospirillales sensu* Hördt et al. (2020) with monophyletic families (besides the inclusion of *Holosporales* and *Hepatincolaceae*), with the only independent branch represented by the *Geminococcaceae*. These results, in particular the monophyly of this lineage with the exclusion of the latter family, are consistent with previous in-depth analyses (Muñoz-Gómez et al. 2019).

3.3 | General Genome Comparisons

We compared the functional repertoire of the *Hepatincolaceae* with a selected set of other bacteria. These include their closest relatives among *Rhodospirillales*, as based on the novel phylogenomics with most biased sites removed (Figure 1), namely genome assemblies and MAGs of marine bacteria, including the symbiont of *Haliotis*, being a closely related but independently evolved host-associated bacterium. The set for comparison was completed by a representative selection of *Holosporales*, which had been considered the closest relatives of *Hepatincolaceae* until now. According to our analyses, the *Holosporales* and the *Hepatincolaceae* should be regarded as instances of independent evolution of host association involving evolutionarily related bacteria, thus making the comparison of their features useful to evidence consistent or peculiar traits.

In terms of annotated COGs, the size of the functional repertoire of the *Hepatincolaceae* is quite homogeneous and within the typical range of the *Holosporales* (Figures 2 and S6; Table S10). However, it is not as large as one of the symbionts of *Haliotis* (which is consistent with the largest sets among *Holosporales*) or of their other close relatives, which, on average, have roughly twice as large COG repertoires (Figure 2). According to their remarkably larger COG sets, all the latter marine bacteria, including MAGs, were deemed as putatively non-host-dependent and will be referred to as ‘free-living’ from now on. On the other hand, all the analysed herein *Holosporales*, including those with the largest genomes and COG repertoires, belong to well-characterised host-associated bacteria and consistently present a reduction in biosynthetic pathways (Categories E, F and H in Figure 2), comparably to the *Hepatincolaceae* (for a recent in-depth analysis, see Giovannini, Petroni, and Castelli 2024).

We thus performed further, more detailed comparisons of the functional repertoires of the analysed organisms, focusing on differences among *Hepatincolaceae* and between *Hepatincolaceae* as a whole and the other investigated bacteria. Overall, the repertoires of the *Hepatincolaceae* and the other symbiotic bacteria investigated are largely subsets of one of their free-living relatives, each with several lineage-specific features (Figure S6). Such comparisons are summarised in Figure 3 and described in detail in (Text S1), with the main matters presented below and the corresponding COG repertoires reported in (Table S10).

3.4 | Carbohydrate and Energy Metabolism

Hepatincolaceae can perform most of the core reactions in carbohydrate and energy metabolism (Figure 3; Table S10),

as previously shown in *Hepatincola* spp. (Dittmer et al. 2023). They are capable of performing glycolysis, but can exert only a few steps that are specific to gluconeogenesis with respect

to glycolysis (and are shared with the Krebs cycle; see Text S1 for details). Consistent with the symbiont of *Haliotis* and many *Holosporales*, the *Hepatincolaceae* present variable

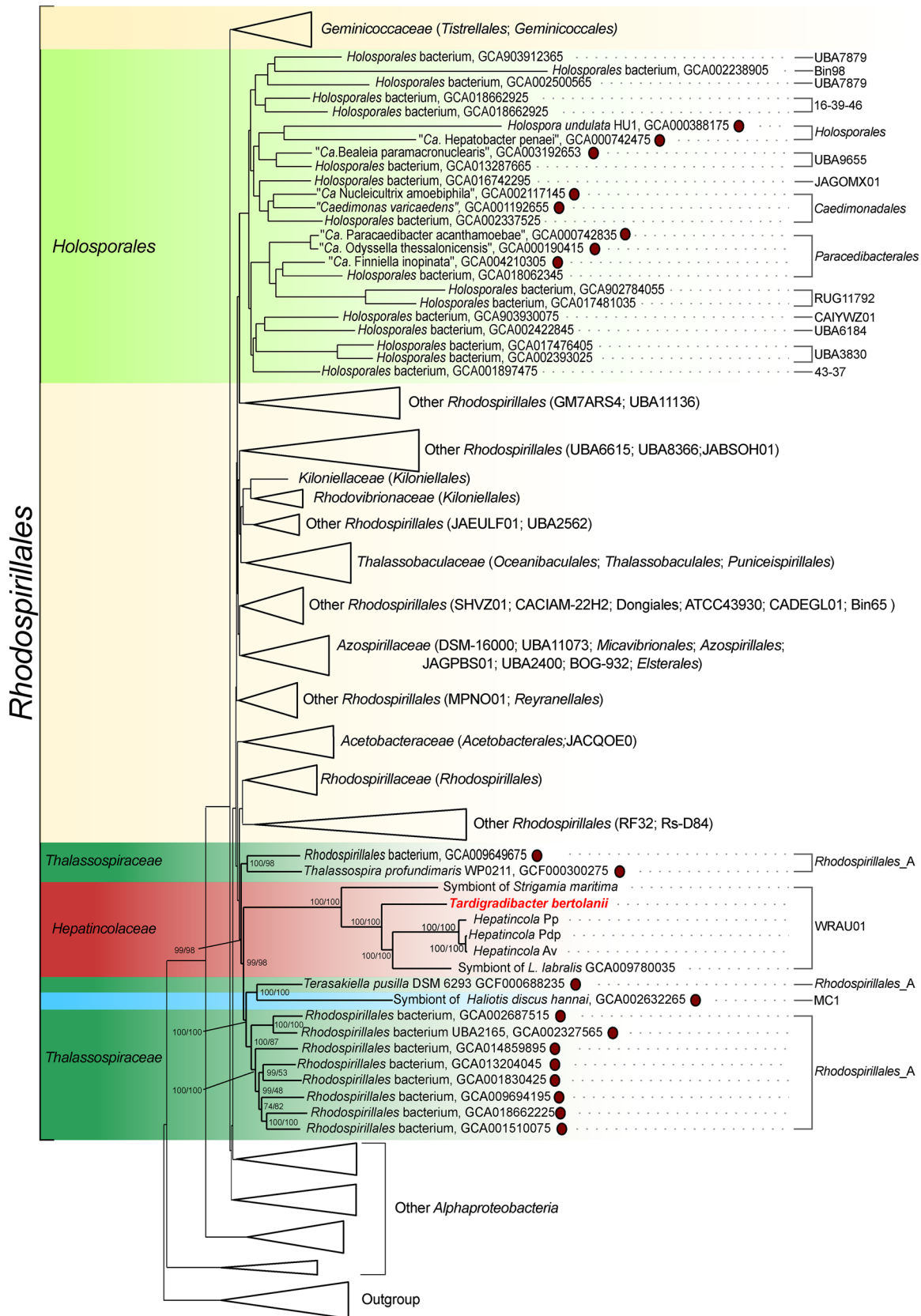


FIGURE 1 | Legend on next page.

levels of reduction of the Krebs cycle, up to full absence in *Tardigradibacter*, and at least a minimal oxidative phosphorylation, which in some representatives includes the cytochrome *o* oxidase, absent in the other symbionts investigated, but is always devoid of cytochrome *c* reductase/oxidase, unlike many *Holosporales* (Figure 3).

Differently from the *Holosporales*, most *Hepatincolaceae* are also capable of getting energy in anaerobic conditions by either lactic or ethanol fermentation, the latter coupled with the ability to employ nitrate as a terminal acceptor (Figure 3).

The *Hepatincolaceae* are also equipped with a quite rich set of transporters that may enable them to obtain several metabolites from their hosts (Figure 3; see also Text S1). In particular, as previously shown in *Hepatincola* spp. (Dittmer et al. 2023), and unlike the free-living relatives or the *Holosporales*, they have several phospho-transferase system (PTS) sugar transporters.

Moreover, all the arthropod-associated *Hepatincolaceae* (i.e., all except *Tardigradibacter*) can metabolise chitin or its derivatives, thanks to diverse enzymes, including chitinases (see Text S1 for details).

3.5 | Metabolism of Nucleotides, Amino Acids and Cofactors

The nucleotide and amino acid synthesis abilities of the *Hepatincolaceae* are quite low (Figure 3; Table S10), as previously shown in *Hepatincola* spp. (Dittmer et al. 2023). Specifically, those for nucleotides are intermediate between the symbiont of *Haliotis* and the *Holosporales*. The *Hepatincolaceae* can also get deoxyribonucleotides via aerobic and, only in the symbiont of *Labiotermes*, anaerobic ribonucleotide reductases. Biosynthetic abilities for amino acids are even scarcer than those for nucleotides, similar to the average *Holosporales*, while the symbiont of *Haliotis* can produce almost all amino acids except for proline and tyrosine (Text S1).

Consistent with their poor biosynthetic capabilities, the *Hepatincolaceae* have several transporters for nucleotides, nucleosides and amino acids (Figure 3), which, however, do not include tlc nucleotide translocases, typical in *Holosporales* and other host-associated intracellular bacteria (Major, Embley, and Williams 2017; Castelli et al. 2024; Mies et al. 2024).

The *Hepatincolaceae* also present versatile capabilities of purine salvage, and for degrading imported (deoxy)nucleosides and

amino acids for energy production, with lineage-specific patterns and partly shared traits with the other bacteria here analysed (Figure 3; Text S1).

On the other hand, all the *Hepatincolaceae* can perform *de novo* synthesis or salvage of several cofactors, either from initial precursors or from intermediates (Figure 3; Table S10; Text S1). Some other capabilities are found in a single representative, namely the symbiont of *Strigamia*, able to synthesise riboflavin and biotin, and the symbiont of *Labiotermes*, able to salvage thiamine diphosphate and synthesise the molybdenum cofactor, the latter being consistent with the presence of the molybdoenzyme nitrate reductase in the same bacterium, just like in the symbiont of *Haliotis* and many free-living relatives.

3.6 | Other Genomic Features

The *Hepatincolaceae* possess components of the Sec core translocon and co-translational transport (Figure 3; Table S10). Regarding post-translational transport, they lack the dedicated Sec component, but they do possess the Sec-independent Tat secretion system. However, they do not present a number of secretion/interaction apparatuses that are typical of the *Holosporales* (Figure 3). At the same time, differently from all of the other bacteria analysed herein, *Hepatincola* Av, *Tardigradibacter* and the symbiont of *Strigamia* possess components of the Type IV secretion system, including the virD2 relaxase (Text S1).

We could also identify some protein classes that have been previously indicated as potentially involved in bacterial–host interactions (Text S1), with several lineage-specific variations, in particular as compared to the *Holosporales*. However, *Hepatincolaceae*, the symbiont of *Haliotis* and their free-living relatives do not have any patatin-like phospholipase (Figure 3).

As most other bacteria herein analysed, all the *Hepatincolaceae* can produce the main components of the typical Gram-negative cell membranes and walls, namely lipids, phospholipids, peptidoglycan, and lipopolysaccharide (Table S10; Text S1). Moreover, they all can perform mismatch repair, nucleotide excision repair, base-excision repair, as well as single- and double-strand break homologous recombination, consistent with the symbiont of *Haliotis* and most *Holosporales* (Text S1).

Interestingly, *Tardigradibacter* can synthesise the osmoprotectant glycine betaine from choline (Nau-Wagner et al. 2012).

FIGURE 1 | Phylogenetic tree of the *Alphaproteobacteria* after removing the 30% most compositionally biased sites. Different coloured backgrounds highlight each group, namely the *Hepatincolaceae* (red), their free-living relatives (dark green), the symbiont of *Haliotis* (blue), the *Holosporales* (light green), and the other *Rhodospirillales* (yellow). For space constraints, the families of the *Rhodospirillales* and the other more distantly related lineages are collapsed as triangular shapes (the full tree is shown in Figure S4, together with the tree inferred on the compositionally untreated dataset and those inferred with the other trimming thresholds). The assignments to order level groups from the GTDB taxonomy for each organism individually in the non-collapsed lineages are reported on the right-hand side, while those of the respective representatives are shown under parentheses for the collapsed lineages. The novel *Tardigradibacter bertolanii* is highlighted in bold, and all the non-*Hepatincolaceae* genome assemblies employed in the comparative genomics are flanked by a dark red circle. For the clade encompassing the *Hepatincolaceae* and their relatives, the number of branches that stand for ultra-fast bootstrap and SH-aLRT support values are reported.

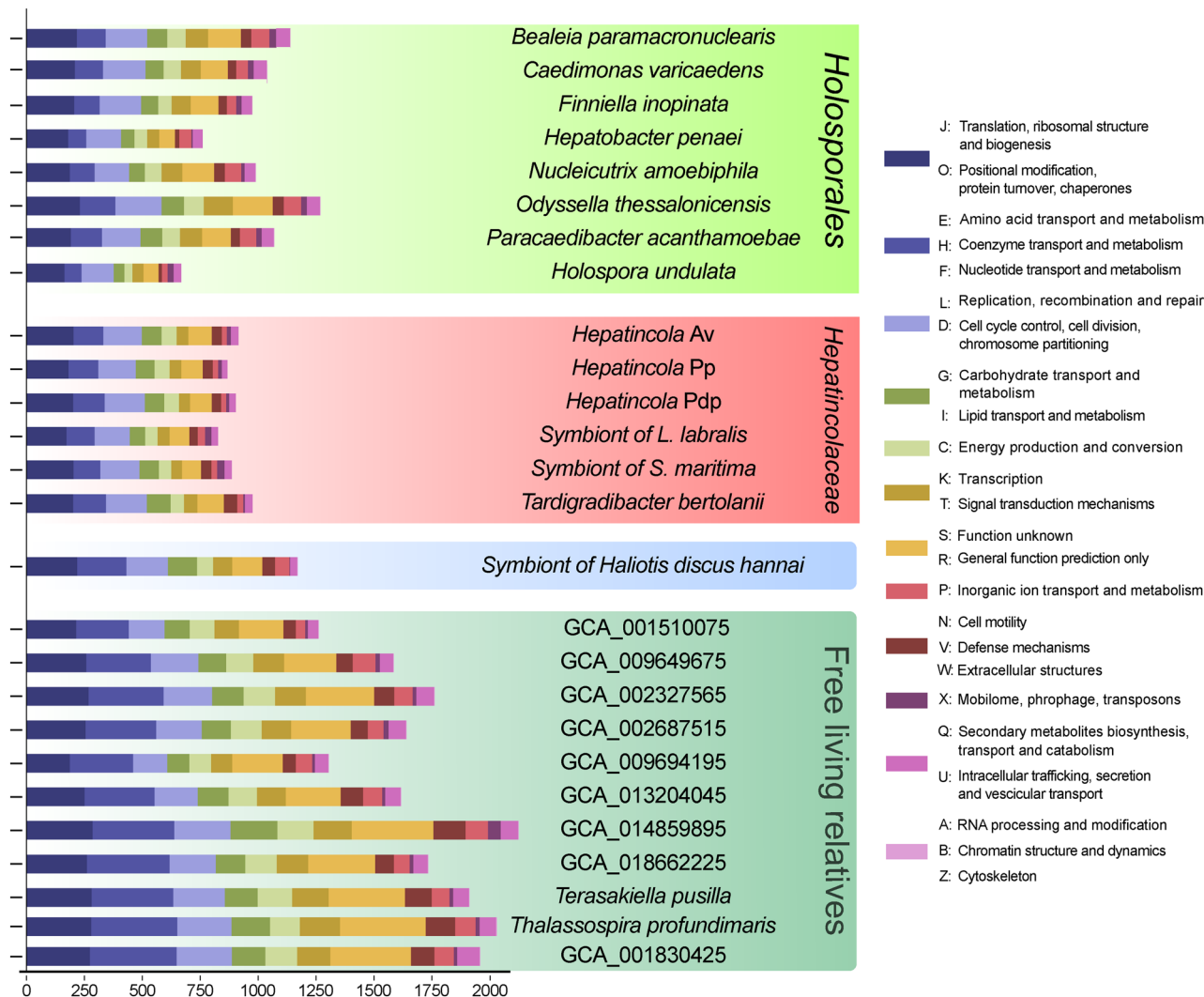


FIGURE 2 | Barplot comparing the functional repertoire in terms of clusters of orthologous groups (COGs) of the *Hepatinoceae* and their relatives. For viewers' clarity, categories were merged into loosely related groups. Bacteria are organised into four groups by their phylogeny and lifestyle, namely the *Holosporales* (light green background), the *Hepatinoceae* (red background), the symbiont of *Haliotis* (blue background), and the free-living relatives of the latter and the *Hepatinoceae* (dark green background). *L. labralis* is an abbreviation for *Labiotermes labralis* and *S. maritima* for *Strigamia maritima*.

Neither gene of this pathway was found in the other *Hepatinoceae* (besides a ~30% truncated *betA* in symbiont of *Strigamia*) or the other symbionts investigated (Figure 3).

Most *Hepatinoceae* possess several phage genes. In *Tardigradibacter*, these are grouped in two putative complete prophages, with high reciprocal sequence identity, each encoding multiple phage components. Many homologues of those genes could be found in the other *Hepatinoceae*, in particular in the 'prophage regions' 1 (in *Hepatincola Av* and *Pdp*) and 3 (in all *Hepatincola* spp.) (Dittmer et al. 2023) as well as in the symbiont of *Strigamia*. The reciprocal sequence identities of phage genes among the *Hepatinoceae* are higher than those of other bacterial lineages, suggestive of a common origin. However, the lack of synteny is also indicative of a non-negligible evolutionary distance (Figure S7). Curiously, a prophage of *Tardigradibacter* encodes for two Type II toxin-antitoxin systems. For only one of the antitoxins, we could

detect a homologue among the *Hepatinoceae*, namely in the symbiont of *Labiotermes*.

4 | Discussion

The *Hepatinoceae* (or 'Ca. Tenuibacteraceae'; Kroer et al. 2016) are a lineage of *Alphaproteobacteria* living in the gut of ecdysozoan hosts, extracellularly associated with microvilli (Kroer et al. 2016; Dittmer et al. 2023). They were until now considered a sublineage of the *Holosporales* (Szokoli et al. 2016; Guidetti et al. 2020), which otherwise include only obligatorily intracellular bacteria prevalently hosted by protists (Giovannini, Petroni, and Castelli 2024). However, due to the paucity of genome sequences, such reconstructions were chiefly based on 16S rRNA gene sequences, which, due to a limited number of sites and compositional biases (low GC content), can be prone to artefacts and in analogous circumstances had

led to incorrect grouping of phylogenetically independent host-associated alphaproteobacterial lineages (e.g., *Rickettsiales* and *Holosporales*) (Ferla et al. 2013; Muñoz-Gómez et al. 2019; Huang et al. 2020).

Herein, we took advantage of an extended genomic dataset of *Hepaticolaceae* and of methods to attenuate compositional biases (Muñoz-Gómez et al. 2019) to reassess the phylogenetic affiliation of this lineage, showing that the grouping with the

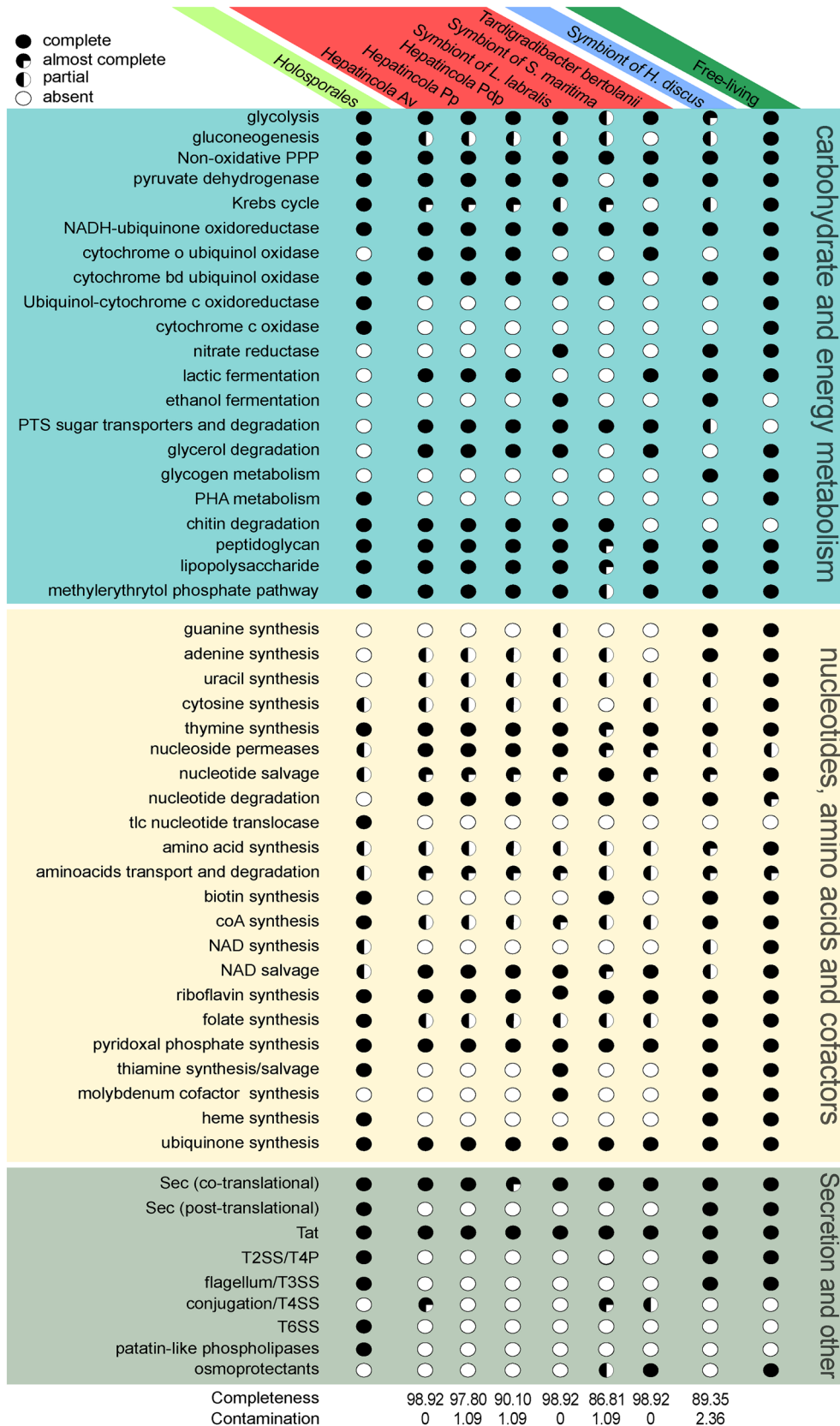


FIGURE 3 | Legend on next page.

Holosporales was artefactual (Figure 1). According to our improved reconstruction, the *Hepatincolaceae* branched within the *Rhodospirillales* sensu Hördt et al. (2020), close to free-living marine bacteria such as *Terasakiella* and *Thalassospira*, as well as to the symbiont of the mollusc *H. discus hannai* (Huang et al. 2020), which however branched within free-living relatives independently from the *Hepatincolaceae*. Interestingly, and consistent with previous studies (Muñoz-Gómez et al. 2019; Castelli et al. 2024), the *Holosporales* branched within the *Rhodospirillales* as well, but apart from the *Hepatincolaceae*, namely close to the families *Kiloniellaceae*, *Rhodovibrionaceae*, *Thalassobaculaceae* and *Azospirillaceae* (Figures 1 and S4). The actual phylogenetic proximity of the *Hepatincolaceae* and the *Holosporales* may have contributed to strengthening their artefactual grouping.

The *Hepatincolaceae* phylogeny suggests that they are the descendants of bacteria that ancestrally engaged in an association with a marine arthropod (or another ecdysozoan), possibly after being accidentally ingested, resisting digestion and later adapting to thrive therein. This seems different from the origin of the *Holosporales*, which has been tentatively linked to freshwater environments (Lanzoni et al. 2024). Such reconstruction of the origin of the *Hepatincolaceae* is also consistent with the fact that several representatives, including early diverging ones, were found in association with marine hosts (Kroer et al. 2016; Dittmer et al. 2023). The colonisation of hosts inhabiting different environments, such as freshwater and terrestrial, could have been favoured by associations with versatile hosts, such as certain crabs (Li et al. 2007), able to survive in both freshwater and seawater (Dittel and Epifanio 2009). Given their extracellular location, the *Hepatincolaceae* are most likely transmitted from host to host via a faecal–oral route (e.g., by egg-smearing; Kikuchi et al. 2009) or by other feeding behaviours observed in known hosts, such as ingestion of exuviae (Tong et al. 2023) or carnivory (Lewis 1961). These mechanisms lead to an indirect vertical transmission, but may also favour horizontal transmission, including host species shift. Considering the inability to produce carbon storage forms such as glycogen or polyhydroxyalkanoate granules (Figure 3), differently from respectively the symbiont of *Haliotis* and several *Holosporales*, the *Hepatincolaceae* are probably engaged only in quite fast stages of passive transmission, a hypothesis also supported by the lack of flagellum (Figure 3). However, considering that all the *Hepatincolaceae* with sequenced genomes are associated with terrestrial hosts, as well as the environment-dependent presence of flagella in other host-associated alphaproteobacterial lineages (Castelli et al. 2024), it would be interesting to check whether aquatic ones bear flagella instead.

According to their genome features, the *Hepatincolaceae* present typical traits of obligatorily host-associated bacteria, in particular, small genome size and absence of essential metabolic pathways, clearly indicative of marked metabolic dependence on their hosts. These features bring them together with the *Holosporales* and the closely related symbiont of *Haliotis*, representing cases of evolutionary convergence. However, significant specificities could be found in the *Hepatincolaceae* as a whole, which, besides being lineage-specific traits, may also shed light on differences in lifestyle between these lineages (Figure S8). For example, we found that all the *Hepatincolaceae* display complex and robust surface structures (peptidoglycan and lipopolysaccharide), differently from some *Holosporales* (George et al. 2020; Castelli et al. 2022), as well as from representatives of the *Rickettsiales*, another group of intracellular *Alphaproteobacteria* (Lin and Rikihisa 2003; Min et al. 2008). They also lack patatin-like phospholipases, which is implied in bacterial evasion from host intracellular vacuoles (Borgo et al. 2022) and is common among the *Holosporales*.

Relevant differences were also found in nutrient uptake and energy metabolism. The *Holosporales* live in quite stable environments, namely host cells, and thus can take advantage of constant metabolite supply. Accordingly, similarly to the *Rickettsiales* (Driscoll et al. 2017), their metabolism (and transporters) can accommodate and exploit a limited number of input molecules to be either consumed directly for ATP synthesis or converted to metabolic precursors, for example, by gluconeogenesis to get building blocks of peptidoglycan and lipopolysaccharide. On the other hand, the host gut where the *Hepatincolaceae* thrive, while potentially quite nutrient-rich as well, is more variegated and variable, being characterised by undigested compounds from each different food source of the host, as well as their multiple intermediates of digestion. To fully exploit these conditions, it seems not surprising that the *Hepatincolaceae* are more flexible in terms of input molecules, bearing transporters and processing enzymes for multiple sugars and other carbon sources, including amino acids and nucleosides, as well as being able to either recycle or degrade them (Figure 3; Text S1). Conversely, the *Holosporales* mostly import nucleotides for biosynthesis purposes only. The presence of full glycolysis in the *Hepatincolaceae* seems consistent with this scenario (Figure 3), together with the significant reduction of gluconeogenesis, which is likely dispensable due to the adequate carbohydrate supply. It is also noteworthy that, while the *Holosporales* are strictly aerobic, the *Hepatincolaceae* have a quite reduced or even absent Krebs cycle, but, thanks to fermentation and/or alternative terminal electron acceptors, many of them can thrive in anaerobic conditions that can be encountered in host guts (Figure 3) (Engel and Moran 2013).

FIGURE 3 | Graphical representation of the presence and completeness degree of selected metabolic pathways and functions in the *Hepatincolaceae* and the other bacteria analysed herein (*Holosporales*, symbiont of *Haliotis*, free-living relatives). Pathways/functions are grouped by the coloured background in three broad sets, namely ‘carbohydrate and energy metabolism’ (teal), ‘nucleotides, amino acids and cofactors’ (yellow), and ‘secretion and other’ (olive green). The filling degree of each circle indicates the completeness of the corresponding function in the organism examined, namely complete (fully coloured), almost complete (three-quarters coloured), partial (half coloured), absent (empty). For the *Holosporales* and the free-living relatives, each circle stands for the completeness of the respective richest representative (for detailed information, see Supplementary Table S10 and Text S1). For reference, the completeness and contamination values of the individual genome assemblies according to CheckM (Table S7) are reported at the bottom lines.

It is also interesting to observe that, besides basic systems such as Sec and Tat, the *Hepatincolaceae* are almost devoid of protein secretion systems, differently from the *Holosporales*, which typically have type II and type VI secretion systems (Figure 3), most probably implied in the interaction with host cells of these intracellular bacteria (George et al. 2020; Giovannini, Petroni, and Castelli 2024). The only partial exception could be the finding, only in some representatives of the *Hepatincolaceae*, of several components of the type IV secretion system (Figure 3), which may be enough for a functional secretory apparatus. At the same time, the presence of the relaxase component virD2 could indicate the role of this apparatus in DNA uptake and transfer (Christie 2016). Considering also the direct and indirect contact with diverse arrays of other bacteria in the host guts (Engel and Moran 2013; Schapheer, Pellens, and Scherson 2021), this may favour HGT events. Interestingly, we identified hints of potential HGTs in a number of cases, mostly involving genes related to the uptake and digestion of various nutrients, in particular multiple carbohydrates, including chitin, nucleosides, and amino acids (Table S11). This suggests a certain degree of genomic plasticity among the *Hepatincolaceae*, which may be related to the lineage-specific adaptations among the *Hepatincolaceae* (see also below). In any case, putative effector molecules, in particular bearing ankyrin repeats, could be found in the *Hepatincolaceae* more frequently than in their free-living relatives (Text S1), which is consistent with a possible modulation of host interactions.

Several highly related phage genes were found in most *Hepatincolaceae*, that is, in all except the symbiont of *Labioterme*s, although this assembly was published after metagenome binning (Hervé et al. 2020), and may thus lack accessory genomic components such as phages. These genes are probably functional, in particular in *Tardigradibacter*, where they are all organised in two putative complete prophages with high reciprocal identities (Figure S7). These findings are consistent with an ancestral presence and/or a lineage-specific adaptation of such prophages in the *Hepatincolaceae*. Considering their conservation despite the quite reduced genomes, it is intriguing to wonder whether those prophages could exert some role in the lifestyle of *Hepatincolaceae* bacteria, particularly in interactions with the hosts. Along this line of thought, the presence of toxin-antitoxin genes in one prophage of *Tardigradibacter* is noteworthy, being reminiscent of prophage-linked toxins and antitoxins implied in the cytoplasmic incompatibility caused by *Wolbachia* in arthropod hosts (Beckmann et al. 2019).

Besides the common trends among the *Hepatincolaceae* presented above, some noteworthy lineage-specific traits were identified, tentatively relatable to host-dependent features. In particular, only *Tardigradibacter* can synthesise the osmoprotectant glycine betaine (Nau-Wagner et al. 2012). This may enable the bacterium to survive during the cryptobiosis of its host (Møbjerg et al. 2011), in particular as a consequence of desiccation, and, in principle, may also contribute to the host's survival in such conditions. Other features of individual *Hepatincolaceae* bacteria can be linked to nutrient uptake in relation to different feeding behaviours of the hosts (Text S1). These include the capability to metabolise derivatives of plant pectins (Renard, Crépeau, and Thibault 1999) by the symbiont of *Labioterme*s, the import of chitobiose by the same bacterium, and, in general, the potential to metabolise chitin by all the arthropod-associated

Hepatincolaceae. These may involve the digestion of exoskeleton from ingested exuviae or preys (Lewis 1961; Tong et al. 2023), or, at least in termites, of the fungal cell wall (Marynowska et al. 2023). This diversity (including likely non-orthologous genes for chitin degradation) and other variations (Text S1) are also consistent with a possible role exerted by potential HGT events in the evolutionary adaptations of the *Hepatincolaceae* (Table S11).

With regard to the symbiont of *Haliotis*, although it has been observed as an intracellular bacterium (Horwitz, Mouton, and Coyne 2016), it shares many genomic traits with the *Hepatincolaceae* rather than with the *Holosporales*. These include anaerobic metabolism, capability to salvage nucleotides and their degradation for energy supply, biosynthetic capabilities for nucleotides and amino acids (the latter much richer than the *Hepatincolaceae*), and lack of secretion systems or patatin-like phospholipases (Figure 3; Text S1). This resemblance might even go beyond the high phylogenetic proximity, hinting at a more complex life cycle of this bacterium.

In summary, our study provides evidence that the phylogenetic grouping of *Hepatincolaceae* with the *Holosporales* was artefactual, similar to other known cases among host-associated and AT-rich *Alphaproteobacteria* (Martijn et al. 2018; Muñoz-Gómez et al. 2019), thereby suggesting that such artefacts may be more common than currently recognised. Our findings allow us to re-think the features of the *Hepatincolaceae* from a novel evolutionary perspective. In particular, thanks also to the significantly increased number of genome sequences made herein available, we identified multiple traits involved in their tailored adaptation to life inside the gut of arthropods and other ecdysozoans. Besides still conjectural indications of a possible contribution of *Tardigradibacter* to host cryptobiosis, our data overall confirm previous indications that the *Hepatincolaceae* most likely behave as nutrient scavengers of gut nutrients rather than being beneficial for their hosts (Dittmer et al. 2023). This parasitic behaviour brings them close to the actual *Holosporales* (Garushyants et al. 2018; George et al. 2020; Castelli et al. 2022; Giovannini, Petroni, and Castelli 2024), which, given the updated phylogenetic scenario, however represents a case of convergent evolution, just like the shared traits among *Holosporales* and *Rickettsiales* (Muñoz-Gómez et al. 2019). Further genomic studies with extended datasets, as well as direct experimental investigations, will be necessary to shed further light on the evolution of the *Hepatincolaceae* and their interactions with their hosts.

4.1 | Description of ‘*Candidatus Tardigradibacter bertolanii*’ gen. nov. sp. nov.

‘*Candidatus* Tar.di.gra.di.bac’ter ber.to.la’ni.i’ (N.L. masc. n. *Tardigradibacter*, ‘tardus’, slow, ‘gradus’, pace, gait, and ‘bacter’, rod; N.L. adj. *bertolanii*, in honour of Professor emeritus Roberto Bertolani).

A bacterium found in association with the tardigrade *R. cf. coronifer*, originating from the moss in Öland, Sweden. According to genome features, devoid of flagella or pili, and capable of both aerobic and anaerobic growth. It may survive during host cryptobiosis. Basis of assignment: SSU rRNA gene sequence

(accession number: MK028537) and complete genome sequence (accession number: JBEULC000000000).

Author Contributions

Michele Castelli: conceptualization, investigation, methodology, formal analysis, supervision, data curation, writing – original draft. **Leandro Gammuto:** methodology, data curation, investigation, formal analysis, visualization. **Diona Podushkina:** investigation, methodology. **Matteo Vecchi:** resources, visualization, writing – review and editing. **Tiziana Altiero:** writing – review and editing, resources. **Emanuela Clementi:** writing – review and editing, methodology. **Roberto Guidetti:** writing – review and editing, investigation, resources. **Lorena Rebecchi:** resources, writing – review and editing, investigation. **Davide Sasser:** writing – review and editing, funding acquisition, conceptualization, resources, validation.

Acknowledgements

This project was funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4—Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of the Italian Ministry of University and Research funded by the European Union—NextGenerationEU to DS. The authors would like to thank Nathaniel Everett for assistance in the genome assembly procedures of *Tardigradibacter bertolanii*, Jessica Dittmer and Didier Bouchon for providing genome assemblies of the three *Hepatincola* spp. Open access publishing facilitated by Università degli Studi di Pavia, as part of the Wiley - CRUI-CARE agreement.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Sequences from this study are available under NCBI BioProject PRJNA1112326: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1112326>, including the assembly for *Tardigradibacter bertolanii* (symbiont of *Richtersius* cf. *coronifer*: JBEULC000000000) and symbionts of *Strigamia maritima* (JBEXBW000000000). All the [Supporting Information](#) files on Zenodo are now accessible under the following doi: [10.5281/zenodo.14286905](https://doi.org/10.5281/zenodo.14286905).

References

Andrews, S. 2010. “FastQC: A Quality Control Tool for High Throughput Sequence Data.” <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.

Baker, B. J., P. Hugenholtz, S. C. Dawson, and J. F. Banfield. 2003. “Extremely Acidophilic Protists From Acid Mine Drainage Host Rickettsiales-Lineage Endosymbionts That Have Intervening Sequences in Their 16S rRNA Genes.” *Applied and Environmental Microbiology* 69: 5512–5518.

Bankevich, A., S. Nurk, D. Antipov, et al. 2012. “SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.” *Journal of Computational Biology* 19: 455–477.

Barco, R. A., G. M. Garrity, J. J. Scott, J. P. Amend, K. H. Nealson, and D. Emerson. 2020. “A Genus Definition for Bacteria and Archaea Based on a Standard Genome Relatedness Index.” *mBio* 11: e02475-19. <https://doi.org/10.1128/mbio.02475-19>.

Bauermeister, J., A. Ramette, and S. Dattagupta. 2012. “Repeatedly Evolved Host-Specific Ectosymbioses Between Sulfur-Oxidizing Bacteria and Amphipods Living in a Cave Ecosystem.” *PLoS One* 7: e50254.

Beckmann, J. F., M. Bonneau, H. Chen, et al. 2019. “The Toxin-Antidote Model of Cytoplasmic Incompatibility: Genetics and Evolutionary Implications.” *Trends in Genetics* 35: 175–185.

Beliavskaia, A. Y., A. V. Predeus, S. K. Garushyants, et al. 2020. “New Intranuclear Symbiotic Bacteria from Macronucleus of *Paramecium putrinum*—‘Candidatus Gortzia Yakutica’.” *Diversity* 12: 198.

Bennett, G. M., and N. A. Moran. 2015. “Heritable Symbiosis: The Advantages and Perils of an Evolutionary Rabbit Hole.” *Proceedings of the National Academy of Sciences of the United States of America* 112: 10169–10176.

Bertolani, R., R. Guidetti, T. Marchioro, T. Altiero, L. Rebecchi, and M. Cesari. 2014. “Phylogeny of Eutardigrada: New Molecular Data and Their Morphological Support Lead to the Identification of New Evolutionary Lineages.” *Molecular Phylogenetics and Evolution* 76: 110–126.

Birtles, R. J., T. J. Rowbotham, R. Michel, et al. 2000. “‘Candidatus Odysella thessalonicensis’ gen. nov., sp. nov., an Obligate Intracellular Parasite of *Acanthamoeba* species.” *International Journal of Systematic and Evolutionary Microbiology* 50 Pt 1: 63–72.

Bolyen, E., J. R. Rideout, M. R. Dillon, et al. 2019. “Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2.” *Nature Biotechnology* 37: 852–857.

Borgo, G. M., T. P. Burke, C. J. Tran, N. T. N. Lo, P. Engström, and M. D. Welch. 2022. “A Patatin-Like Phospholipase Mediates *Rickettsia parkeri* Escape From Host Membranes.” *Nature Communications* 13: 3656.

Borowiec, M. L. 2016. “AMAS: A Fast Tool for Alignment Manipulation and Computing of Summary Statistics.” *PeerJ* 4: e1660.

Boscaro, V., C. C. Holt, N. W. L. Van Steenkiste, et al. 2022. “Microbiomes of Microscopic Marine Invertebrates Do Not Reveal Signatures of Phyllosymbiosis.” *Nature Microbiology* 7: 810–819.

Boscaro, V., F. Husnik, C. Vannini, and P. J. Keeling. 2019. “Symbionts of the Ciliate *Euplotes*: Diversity, Patterns and Potential as Models for Bacteria-Eukaryote Endosymbioses.” *Proceedings of the Biological Sciences* 286: 20190693.

Bouchon, D., M. Zimmer, and J. Dittmer. 2016. “The Terrestrial Isopod Microbiome: An All-In-One Toolbox for Animal-Microbe Interactions of Ecological Relevance.” *Frontiers in Microbiology* 7: 1472.

Buchfink, B., C. Xie, and D. H. Huson. 2015. “Fast and Sensitive Protein Alignment Using DIAMOND.” *Nature Methods* 12: 59–60.

Cantalapiedra, C. P., A. Hernández-Plaza, I. Letunic, P. Bork, and J. Huerta-Cepas. 2021. “eggNOG-Mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale.” *Molecular Biology and Evolution* 38: 5825–5829.

Castelli, M., O. Lanzoni, M. Giovannini, et al. 2022. “‘Candidatus Gromoviella Agglomerans’, a Novel Intracellular Holosporaceae Parasite of the Ciliate *Paramecium* Showing Marked Genome Reduction.” *Environmental Microbiology Reports* 14: 34–49.

Castelli, M., T. Nardi, L. Gammuto, et al. 2024. “Host Association and Intracellularity Evolved Multiple Times Independently in the Rickettsiales.” *Nature Communications* 15: 1093.

Castelli, M., E. Sabaneyeva, O. Lanzoni, et al. 2019. “Deianiraea, an Extracellular Bacterium Associated With the Ciliate *Paramecium*, Suggests an Alternative Scenario for the Evolution of Rickettsiales.” *ISME Journal* 13: 2280–2294.

Chaumeil, P.-A., A. J. Mussig, P. Hugenholtz, and D. H. Parks. 2020. “GTDB-Tk: A Toolkit to Classify Genomes With the Genome Taxonomy Database.” *Bioinformatics* 36: 1925–1927.

Chipman, A. D., D. E. K. Ferrier, C. Brena, et al. 2014. “The First Myriapod Genome Sequence Reveals Conservative Arthropod Gene

- Content and Genome Organisation in the Centipede *Strigamia maritima*." *PLoS Biology* 12: e1002005.
- Christie, P. J. 2016. "The Mosaic Type IV Secretion Systems." *EcoSal Plus* 7. <https://doi.org/10.1128/ecosalplus.ESP-0020-2015>.
- Criscuolo, A., and S. Gribaldo. 2010. "BMGE (Block Mapping and Gathering With Entropy): A New Software for Selection of Phylogenetic Informative Regions From Multiple Sequence Alignments." *BMC Evolutionary Biology* 10: 210.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. "jModelTest 2: More Models, New Heuristics and Parallel Computing." *Nature Methods* 9: 772.
- Dirren, S., and T. Posch. 2016. "Promiscuous and Specific Bacterial Symbiont Acquisition in the Amoeboid Genus *Nuclearia* (Opisthokonta)." *FEMS Microbiology Ecology* 92: fiw105.
- Dittel, A. I., and C. E. Epifanio. 2009. "Invasion Biology of the Chinese Mitten Crab *Eriocheir sinensis*: A Brief Review." *Journal of Experimental Marine Biology and Ecology* 374: 79–92.
- Dittmer, J., M. Bredon, B. Moumen, M. Raimond, P. Grève, and D. Bouchon. 2023. "The Terrestrial Isopod Symbiont '*Candidatus Hepatincola porcellionum*' Is a Potential Nutrient Scavenger Related to Holosporales Symbionts of Protists." *ISME Communications* 3: 18.
- Drew, G. C., E. J. Stevens, and K. C. King. 2021. "Microbial Evolution and Transitions Along the Parasite–Mutualist Continuum." *Nature Reviews Microbiology* 19: 623–638.
- Driscoll, T. P., V. I. Verhoeve, M. L. Guillotte, et al. 2017. "Wholly Rickettsia! Reconstructed Metabolic Profile of the Quintessential Bacterial Parasite of Eukaryotic Cells." *mBio* 8: e00859-17.
- Engel, P., and N. A. Moran. 2013. "The Gut Microbiota of Insects—Diversity in Structure and Function." *FEMS Microbiology Reviews* 37: 699–735.
- Ferla, M. P., J. C. Thrash, S. J. Giovannoni, and W. M. Patrick. 2013. "New rRNA Gene-Based Phylogenies of the *Alphaproteobacteria* Provide Perspective on Major Groups, Mitochondrial Ancestry and Phylogenetic Instability." *PLoS One* 8: e83383.
- Fokin, S. I., V. Serra, F. Ferrantini, L. Modeo, and G. Petroni. 2019. "*Candidatus Hafkinia Simulans* gen. nov., sp. nov., a Novel Holospora-Like Bacterium From the Macronucleus of the Rare Brackish Water Ciliate *Frontonia salmastra* (Oligohymenophorea, Ciliophora): Multidisciplinary Characterization of the New Endosymbiont and Its Host." *Microbial Ecology* 77: 1092–1106.
- Galperin, M. Y., K. S. Makarova, Y. I. Wolf, and E. V. Koonin. 2015. "Expanded Microbial Genome Coverage and Improved Protein Family Annotation in the COG Database." *Nucleic Acids Research* 43: D261–D269.
- Garushyants, S. K., A. Y. Beliavskaia, D. B. Malko, M. D. Logacheva, M. S. Rautian, and M. S. Gelfand. 2018. "Comparative Genomic Analysis of *Holospira* spp., Intranuclear Symbionts of Paramecia." *Frontiers in Microbiology* 9: 738.
- George, E. E., F. Husnik, D. Tashyreva, et al. 2020. "Highly Reduced Genomes of Protist Endosymbionts Show Evolutionary Convergence." *Current Biology* 30: 925–933.e3.
- Georgiades, K., M.-A. Madoui, P. Le, C. Robert, and D. Raoult. 2011. "Phylogenomic Analysis of *Odyssella thessalonicensis* Fortifies the Common Origin of *Rickettsiales*, *Pelagibacter ubique* and *Reclinomonas americana* Mitochondrion." *PLoS One* 6: e24857.
- Giovannini, M., G. Petroni, and M. Castelli. 2024. "Novel Evolutionary Insights on the Interactions of the *Holosporales* (*Alphaproteobacteria*) With Eukaryotic Hosts From Comparative Genomics." *Environmental Microbiology* 26: e16562.
- Guidetti, R., L. Rebecchi, R. Bertolani, K. I. Jönsson, R. Møbjerg Kristensen, and M. Cesari. 2016. "Morphological and Molecular Analyses on *Richtersius* (Eutardigrada) Diversity Reveal Its New Systematic Position and Lead to the Establishment of a New Genus and a New Family Within Macrobiotodea." *Zoological Journal of the Linnean Society* 178: 834–845.
- Guidetti, R., M. Vecchi, A. Ferrari, I. L. G. Newton, M. Cesari, and L. Rebecchi. 2020. "Further Insights in the Tardigrada Microbiome: Phylogenetic Position and Prevalence of Infection of Four New Alphaproteobacteria Putative Endosymbionts." *Zoological Journal of the Linnean Society* 188: 925–937.
- Guindon, S., and O. Gascuel. 2003. "A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood." *Systematic Biology* 52: 696–704.
- Hervé, V., P. Liu, C. Dietrich, et al. 2020. "Phylogenomic Analysis of 589 Metagenome-Assembled Genomes Encompassing all Major Prokaryotic Lineages From the Gut of Higher Termites." *PeerJ* 8: e8614.
- Hess, S., A. Suthaus, and M. Melkonian. 2016. "*Candidatus Finniella* (*Rickettsiales*, *Alphaproteobacteria*), novel Endosymbionts of Viridiraptorid Amoeboflagellates (Cercozoa, Rhizaria)." *Applied and Environmental Microbiology* 82: 659–670.
- Hördt, A., M. G. López, J. P. Meier-Kolthoff, et al. 2020. "Analysis of 1,000+ Type-Strain Genomes Substantially Improves Taxonomic Classification of *Alphaproteobacteria*." *Frontiers in Microbiology* 11: 468.
- Horn, M., T. R. Fritsche, R. K. Gautom, K. H. Schleifer, and M. Wagner. 1999. "Novel Bacterial Endosymbionts of *Acanthamoeba* spp. Related to the *Paramecium caudatum* Symbiont *Caedibacter caryophilus*." *Environmental Microbiology* 1: 357–367.
- Horwitz, R., A. Mouton, and V. E. Coyne. 2016. "Characterization of an Intracellular Bacterium Infecting the Digestive Gland of the South African Abalone *Haliotis midae*." *Aquaculture* 451: 24–32.
- Huang, Z., J. M. Petersen, J. Martijn, T. J. G. Ettema, and Z. Shao. 2020. "A Novel Alphaproteobacterium With a Small Genome Identified From the Digestive Gland of Multiple Species of Abalone." *Environmental Microbiology Reports* 12: 387–395.
- Huerta-Cepas, J., D. Szklarczyk, D. Heller, et al. 2019. "eggNOG 5.0: A Hierarchical, Functionally and Phylogenetically Annotated Orthology Resource Based on 5090 Organisms and 2502 Viruses." *Nucleic Acids Research* 47: D309–D314.
- Husnik, F., D. Tashyreva, V. Boscaro, E. E. George, J. Lukeš, and P. J. Keeling. 2021. "Bacterial and Archaeal Symbioses With Protists." *Current Biology* 31: R862–R877.
- Jones, R. T., K. F. McCormick, and A. P. Martin. 2008. "Bacterial Communities of Bartonella-Positive Fleas: Diversity and Community Assembly Patterns." *Applied and Environmental Microbiology* 74: 1667–1670.
- Kaczmarek, Ł., M. Roszkowska, I. Poprawa, et al. 2020. "Integrative Description of Bisexual *Paramacrotus experimentalis* sp. nov. (Macrotiotidae) From Republic of Madagascar (Africa) With Microbiome Analysis." *Molecular Phylogenetics and Evolution* 145: 106730.
- Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. 2016. "KEGG as a Reference Resource for Gene and Protein Annotation." *Nucleic Acids Research* 44: D457–D462.
- Karp, P. D., R. Billington, R. Caspi, et al. 2019. "The BioCyc Collection of Microbial Genomes and Metabolic Pathways." *Briefings in Bioinformatics* 20: 1085–1093.
- Katoh, K., and D. M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30: 772–780.
- Kikuchi, Y., T. Hosokawa, N. Nikoh, X.-Y. Meng, Y. Kamagata, and T. Fukatsu. 2009. "Host-Symbiont Co-Speciation and Reductive Genome Evolution in Gut Symbiotic Bacteria of Acanthosomatid Stinkbugs." *BMC Biology* 7: 2.

- Koch, H., D. P. Abrol, J. Li, and P. Schmid-Hempel. 2013. "Diversity and Evolutionary Patterns of Bacterial Gut Associates of Corbiculate Bees." *Molecular Ecology* 22: 2028–2044.
- Konecka, E., and Z. Olszanowski. 2019. "Detection of a New Bacterium of the Family Holosporaceae (Alphaproteobacteria: Holosporales) Associated With the Oribatid Mite *Achipteria coleoprata*." *Biologia* 74: 1517–1522.
- Kroer, P., K. U. Kjeldsen, J. R. Nyengaard, A. Schramm, and P. Funch. 2016. "A Novel Extracellular Gut Symbiont in the Marine Worm *Priapululus caudatus* (Priapulida) Reveals an Alphaproteobacterial Symbiont Clade of the Ecdysozoa." *Frontiers in Microbiology* 7: 539.
- Kumar, S., M. Jones, G. Koutsovoulos, M. Clarke, and M. Blaxter. 2013. "Blobology: Exploring Raw Genome Data for Contaminants, Symbionts and Parasites Using Taxon-Annotated GC-Coverage Plots." *Frontiers in Genetics* 4: 237.
- Langmead, B., and S. L. Salzberg. 2012. "Fast Gapped-Read Alignment With Bowtie 2." *Nature Methods* 9: 357–359.
- Lanzoni, O., F. Szokoli, M. Schrallhammer, et al. 2024. "'Candidatus Intestinibacterium Parameciophilum'—Member of the 'Candidatus Paracaedibacteraceae' Family (Alphaproteobacteria, Holosporales) Inhabiting the ciliated protist *Paramecium*." *International Microbiology* 27: 659–671.
- Lewis, J. G. E. 1961. "The Life History and Ecology of the Littoral Centipede *Strigamia* (= *Scolioplanes*) *Maritima* (Leach)." *Proceedings of the Zoological Society of London* 137: 221–248.
- Li, K., W. Guan, G. Wei, et al. 2007. "Phylogenetic Analysis of Intestinal Bacteria in the Chinese Mitten Crab (*Eriocheir sinensis*)." *Journal of Applied Microbiology* 103: 675–682.
- Lin, M., and Y. Rikihisa. 2003. "Ehrlichia Chaffeensis and *Anaplasma phagocytophilum* Lack Genes for Lipid A Biosynthesis and Incorporate Cholesterol for Their Survival." *Infection and Immunity* 71: 5324–5331.
- Lu, S., J. Wang, F. Chitsaz, et al. 2020. "CDD/SPARCLE: The Conserved Domain Database in 2020." *Nucleic Acids Research* 48: D265–D268.
- Ludwig, W., O. Strunk, R. Westram, et al. 2004. "ARB: A Software Environment for Sequence Data." *Nucleic Acids Research* 32: 1363–1371.
- Major, P., T. M. Embley, and T. A. Williams. 2017. "Phylogenetic Diversity of NTT Nucleotide Transport Proteins in Free-Living and Parasitic Bacteria and Eukaryotes." *Genome Biology and Evolution* 9: 480–487.
- Martijn, J., J. Vosseberg, L. Guy, P. Offre, and T. J. G. Ettema. 2018. "Deep Mitochondrial Origin Outside the Sampled Alphaproteobacteria." *Nature* 557: 101–105.
- Marynowska, M., D. Sillam-Dussès, B. Untereiner, et al. 2023. "A Holobiont Approach Towards Polysaccharide Degradation by the Highly Compartmentalised Gut System of the Soil-Feeding Higher Termite *Labiotermes labralis*." *BMC Genomics* 24: 115.
- McCutcheon, J. P., and N. A. Moran. 2012. "Extreme Genome Reduction in Symbiotic Bacteria." *Nature Reviews. Microbiology* 10: 13–26.
- McFall-Ngai, M., M. G. Hadfield, T. C. G. Bosch, et al. 2013. "Animals in a Bacterial World, a New Imperative for the Life Sciences." *Proceedings of the National Academy of Sciences of the United States of America* 110: 3229–3236.
- Midha, S., D. J. Rigden, S. Siozios, G. D. D. Hurst, and A. P. Jackson. 2021. "*Bodo saltans* (Kinetoplastida) is Dependent on a Novel *Paracaedibacter*-Like Endosymbiont That Possesses Multiple Putative Toxin-Antitoxin Systems." *ISME Journal* 15: 1680–1694.
- Mies, U. S., V. Hervé, T. Kropp, et al. 2024. "Genome Reduction and Horizontal Gene Transfer in the Evolution of Endomicrobia—Rise and Fall of an Intracellular Symbiosis With Termite Gut Flagellates." *mBio* 15: e0082624.
- Min, C.-K., J.-S. Yang, S. Kim, M.-S. Choi, I.-S. Kim, and N.-H. Cho. 2008. "Genome-Based Construction of the Metabolic Pathways of *Orientia tsutsugamushi* and Comparative Analysis Within the Rickettsiales Order." *Comparative and Functional Genomics* 2008: 623145.
- Minh, B. Q., M. A. T. Nguyen, and A. von Haeseler. 2013. "Ultrafast Approximation for Phylogenetic Bootstrap." *Molecular Biology and Evolution* 30: 1188–1195.
- Mioduchowska, M., B. Nitkiewicz, M. Roszkowska, et al. 2021. "Taxonomic Classification of the Bacterial Endosymbiont *Wolbachia* Based on Next-Generation Sequencing: Is There Molecular Evidence for Its Presence in Tardigrades?" *Genome* 64: 951–958.
- Möbjerg, N., K. A. Halberg, A. Jørgensen, et al. 2011. "Survival in Extreme Environments—On the Current Knowledge of Adaptations in Tardigrades." *Acta Physiologica* 202: 409–420.
- Muñoz-Gómez, S. A., S. Hess, G. Burger, et al. 2019. "An Updated Phylogeny of the *Alphaproteobacteria* Reveals That the Parasitic *Rickettsiales* and *Holosporales* Have Independent Origins." *eLife* 8: e42535.
- Nau-Wagner, G., D. Opper, A. Rolbetzki, et al. 2012. "Genetic Control of Osmoadaptive Glycine Betaine Synthesis in *Bacillus subtilis* Through the Choline-Sensing and Glycine Betaine-Responsive GbsR Repressor." *Journal of Bacteriology* 194: 2703–2714.
- Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies." *Molecular Biology and Evolution* 32: 268–274.
- Nunan, L. M., C. R. Pantoja, S. Gomez-Jimenez, and D. V. Lightner. 2013. "'Candidatus Hepatobacter Penaei,' an Intracellular Pathogenic Enteric Bacterium in the Hepatopancreas of the Marine Shrimp *Penaeus vannamei* (Crustacea: Decapoda)." *Applied and Environmental Microbiology* 79: 1407–1409.
- Nurk, S., D. Meleshko, A. Korobeynikov, and P. A. Pevzner. 2017. "metaSPAdes: A New Versatile Metagenomic Assembler." *Genome Research* 27: 824–834.
- Parks, D. H., M. Chuvochina, C. Rinke, A. J. Mussig, P.-A. Chaumeil, and P. Hugenholtz. 2022. "GTDB: An Ongoing Census of Bacterial and Archaeal Diversity Through a Phylogenetically Consistent, Rank Normalized and Complete Genome-Based Taxonomy." *Nucleic Acids Research* 50: D785–D794.
- Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson. 2015. "CheckM: Assessing the Quality of Microbial Genomes Recovered From Isolates, Single Cells, and Metagenomes." *Genome Research* 25: 1043–1055.
- Potekhin, A., M. Schweikert, I. Nekrasova, et al. 2018. "Complex Life Cycle, Broad Host Range and Adaptation Strategy of the Intranuclear *Paramecium* Symbiont *Preeria caryophila* comb. nov." *FEMS Microbiology Ecology* 94: fiy076.
- Qu, S.-X., H.-P. Li, L. Ma, et al. 2015. "Effects of Different Edible Mushroom Hosts on the Development, Reproduction and Bacterial Community of *Tyrophagus putrescentiae* (Schränk)." *Journal of Stored Products Research* 61: 70–75.
- Quast, C., E. Pruesse, P. Yilmaz, et al. 2013. "The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic Acids Research* 41: D590–D596.
- Ramírez-Puebla, S. T., M. Rosenblueth, C. K. Chávez-Moreno, M. C. C. P. de Lyra, A. Tecante, and E. Martínez-Romero. 2010. "Molecular Phylogeny of the Genus *Dactylopius* (Hemiptera: Dactylopiidae) and Identification of the Symbiotic Bacteria." *Environmental Entomology* 39: 1178–1183.
- Renard, C. M., M. J. Crépeau, and J. F. Thibault. 1999. "Glucuronic Acid Directly Linked to Galacturonic Acid in the Rhamnogalacturonan Backbone of Beet Pectins." *European Journal of Biochemistry* 266: 566–574.

- Schapheer, C., R. Pellens, and R. Scherson. 2021. "Arthropod-Microbiota Integration: Its Importance for Ecosystem Conservation." *Frontiers in Microbiology* 12: 702763.
- Schrallhammer, M., M. Castelli, and G. Petroni. 2018. "Phylogenetic Relationships Among Endosymbiotic R-Body Producer: Bacteria Providing Their Host the Killer Trait." *Systematic and Applied Microbiology* 41: 213–220.
- Schrallhammer, M., and A. Potekhin. 2020. "Epidemiology of Nucleus-Dwelling *Holospira*: Infection, Transmission, Adaptation, and Interaction With *Paramecium*." In *Symbiosis: Cellular, Molecular, Medical and Evolutionary Aspects. Results and Problems in Cell Differentiation*, edited by M. Kloc, vol. 69, 105–135. Cham: Springer.
- Schrallhammer, M., and M. Schweikert. 2009. "The Killer Effect of *Paramecium* and Its Causative Agents." In *Endosymbionts in Paramecium. Microbiology Monographs. Fujishima*, edited by M. Fujishima, 227–246. Berlin, Heidelberg: Springer.
- Schulz, F., I. Lagkouvardos, F. Wascher, K. Aistleitner, R. Kostanjšek, and M. Horn. 2014. "Life in an Unusual Intracellular Niche: A Bacterial Symbiont Infecting the Nucleus of Amoebae." *ISME Journal* 8: 1634–1644.
- Seemann, T. 2013. "barrnap 0.5: Rapid Ribosomal RNA Prediction." Available online at: <http://www.vicbioinformatics.com/>.
- Seemann, T. 2014. "Prokka: Rapid Prokaryotic Genome Annotation." *Bioinformatics* 30: 2068–2069.
- Shiohama, Y., K. Takeshita, Y. Hirakata, M. K. Nobu, M. Ito, and N. Shinzato. 2022. "Complete Genome Sequence of 'Candidatus Hydrogeosomobacter endosymbioticus,' an Intracellular Bacterial Symbiont of the Anaerobic Ciliate Scuticociliate GW7." *Microbiology Resource Announcements* 11: e0115021.
- Szokoli, F., M. Castelli, E. Sabaneyeva, et al. 2016. "Disentangling the Taxonomy of Rickettsiales and Description of Two Novel Symbionts ('Candidatus Bealeia Paramacronuclearis' and 'Candidatus Fokinia Cryptica') Sharing the Cytoplasm of the Ciliate Protist *Paramecium bicaurelia*." *Applied and Environmental Microbiology* 82: 7236–7247.
- Tibbs-Cortes, L. E., B. W. Tibbs-Cortes, and S. Schmitz-Esser. 2022. "Tardigrade Community Microbiomes in North American Orchards Include Putative Endosymbionts and Plant Pathogens." *Frontiers in Microbiology* 13: 866930.
- Tong, R. L., J. S. Patel, J. M. Gordon, S.-B. Lee, T. Chouvenec, and N.-Y. Su. 2023. "Exuviae Recycling Can Enhance Queen Oviposition and Colony Growth in Subterranean Termites (Blattodea: Rhinotermitidae: Coptotermites)." *Environmental Entomology* 52: 254–258.
- Vecchi, M., I. L. G. Newton, M. Cesari, L. Rebecchi, and R. Guidetti. 2018. "The Microbial Community of Tardigrades: Environmental Influence and Species Specificity of Microbiome Structure and Composition." *Microbial Ecology* 76: 467–481.
- Wang, S., and H. Luo. 2021. "Dating Alphaproteobacteria Evolution With Eukaryotic Fossils." *Nature Communications* 12: 3324.
- Wang, Y., U. Stingl, F. Anton-Erxleben, M. Zimmer, and A. Brune. 2004. "'Candidatus Hepatincola porcellionum' gen. nov., sp. nov., a New, Stalk-Forming Lineage of Rickettsiales Colonizing the Midgut Glands of a Terrestrial Isopod." *Archives of Microbiology* 181: 299–304.
- Wang, Z., and M. Wu. 2015. "An Integrated Phylogenomic Approach Toward Pinpointing the Origin of Mitochondria." *Scientific Reports* 5: 7949.
- Wernegreen, J. J. 2017. "In It for the Long Haul: Evolutionary Consequences of Persistent Endosymbiosis." *Current Opinion in Genetics and Development* 47: 83–90.
- Wick, R. R., M. B. Schultz, J. Zobel, and K. E. Holt. 2015. "Bandage: Interactive Visualization of de Novo Genome Assemblies." *Bioinformatics* 31: 3350–3352.
- Yoon, S.-H., S. Ha, J. Lim, S. Kwon, and J. Chun. 2017. "A Large-Scale Evaluation of Algorithms to Calculate Average Nucleotide Identity." *Antonie van Leeuwenhoek* 110: 1281–1286.
- Zawierucha, K., A. Trzebny, J. Buda, et al. 2022. "Trophic and Symbiotic Links Between Obligate-Glacier Water Bears (Tardigrada) and Cryoconite Microorganisms." *PLoS One* 17: e0262039.
- Zilio, G., L. S. Nørgaard, G. Petrucci, et al. 2021. "Parasitism and Host Dispersal Plasticity in an Aquatic Model System." *Journal of Evolutionary Biology* 34: 1316–1325.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.