Hologenome analysis reveals dual symbiosis in the deep-sea hydrothermal vent snail *Gigantopelta aegis*

Yi Lan1,2, Jin Sun1,2, Chong Chen3, Yanan Sun1, Yadong Zhou4, Yi Yang1,2, Weipeng Zhang5, Runsheng Li6, Kun Zhou1, Wai Chuen Wong1,2, Yick Hang Kwan1,2, Aifang Cheng1,2, Salim Bougouffa7,8, Cindy Lee Van Dover9, Jian-Wen Qiu10 & Pei-Yuan Qian1,2✉

Animals endemic to deep-sea hydrothermal vents often form obligatory symbioses with bacteria, maintained by intricate host–symbiont interactions. Most genomic studies on holobionts have not investigated both sides to similar depths. Here, we report dual symbiosis in the peltospirid snail *Gigantopelta aegis* with two gammaproteobacterial endosymbionts: a sulfur oxidiser and a methane oxidiser. We assemble high-quality genomes for all three parties, including a chromosome-level host genome. Hologenomic analyses reveal mutualism with nutritional complementarity and metabolic co-dependency, highly versatile in transporting and using chemical energy. *Gigantopelta aegis* likely remodels its immune system to facilitate dual symbiosis. Comparisons with *Chrysomallon squamiferum*, a conidential snail with a single sulfur-oxidising gammaproteobacterial endosymbiont, show that their sulfur-oxidising endosymbionts are phylogenetically distant. This is consistent with previous findings that they evolved endosymbiosis convergently. Notably, the two sulfur-oxidisers share the same capabilities in biosynthesising nutrients lacking in the host genomes, potentially a key criterion in symbiont selection.
A
imals live with microorganisms in and around them, forming ecological units referred to as holobionts. Some metazoans, such as reef-forming corals, have established intricate obligatory symbiotic relationships with single-celled organisms that enable the holobiont to utilise available energy sources more effectively. In deep-sea chemosynthetic environments such as hydrothermal vents, many endemic animals harbour endosymbiotic bacteria within their cells to access energy and nutrients released through microbial oxidation of reducing substances such as hydrogen sulfide, hydrogen, thiosulfate, and methane. Most of these, such as the peltospirid Scaly-foot Snail Chrysomallon squamiferum, the giant vesicomyid clam Turneroconcha magna, as well as the siboglinid tubeworms Riftia pachyptila and Paracarcinia echinospina, have evolved obligate symbiotic relationships with one dominant phylotype of microbial endosymbiont. By contrast, several species of annelids and molluscs, including the thyasirid clam Maorithyas hadalís, the provannid snail genus Alviniconcha, some Bathymodiolus mussels, and the shallow-water oligochaete worm Olavius, are capable of hosting multiple co-occurring endosymbionts.

Dracogyra gigantea is a family of gastropods endemic to deep-sea hydrothermal vents. It currently has 11 recognised genera, but only Gigantopelta and Chrysomallon are known to have enlarged oesophageal glands hosting endosymbionts for nutrition, while the rest are grazers, filter-feeders, or deposit feeders. Snails in both genera appear to obtain their endosymbionts from the environment through horizontal transmission in each generation. Previous multi-gene phylogenetic analyses have revealed that Gigantopelta and Chrysomallon are not the most closely related peltospirid genera, and the two are separated by a number of other much smaller, non-endosymbiotic genera such as Peltopsira, Nodopelta, Rhynochopela, and Dragocysta. Despite their superficial similarity, Gigantopelta and Chrysomallon are also quite different in their internal anatomy. A crucial difference between the two is that while Chrysomallon relies on endosymbionts for nutrition throughout its post-settlement life, Gigantopelta initially adopts grazing for nutrients and only later (at around 5–7 mm shell length) rapidly shifts to hosting endosymbionts. This process, dubbed ‘cryptometamorphosis’, occurs with a dramatic reorganisation of its digestive system where the oesophageal gland is greatly enlarged while the other parts of the digestive system effectively stop growing. Thus, these two genera have evolved an endosymbiotic lifestyle independently and convergently, rather than from a common ancestor with symbionts.

The fact that representatives of the two genera, Gigantopelta aegis and the Scaly-foot Snail C. squamiferum, co-occur in great abundance in the Longqi vent field on the ultra-slow spreading Southwest Indian Ridge makes them ideal candidates to compare molecular adaptations in the two holobiont systems. From previous studies, we know that C. squamiferum hosts only one phylotype of sulfur-oxidising Gammaproteobacteria whose genome has been sequenced. This genome shows that the symbionts of C. squamiferum are capable of sulfur oxidation, carbon fixation, and nutrient biosynthesis. The whole genome of C. squamiferum was recently sequenced and assembled to chromosome-scale, but the study focused on the origin of its unique biomineralisation, leaving adaptations to symbiosis mostly unexamined. Previously, 16S rRNA sequencing of the closely related Gigantopelta chesaui only revealed a single sulfur-oxidising Gammaproteobacteria symbiont, although transmission electron microscopy (TEM) micrographs of the oesophageal gland hinted a second, rarer, likely methane-oxidising symbiont.

Here, we reveal that G. aegis exhibits dual endosymbiosis housing two Gammaproteobacteria endosymbionts, one sulfur-oxidising and one methane-oxidising, in its oesophageal gland. We use high-quality genome assemblies of both the G. aegis snail host and its two physiologically distinct endosymbionts to elucidate interactions and co-operations among this ‘triptych’ of parties in this holobiont. Furthermore, we compare the hologenome of G. aegis with that of C. squamiferum to understand how G. aegis benefits from maintaining symbiosis with more than one endosymbiont.

Results and discussion
Features of a holobiont with three parties. Members of the two peltospirid genera, Gigantopelta aegis and the Scaly-foot Snail Chrysomallon squamiferum, occur side-by-side at Tiatam chimney in the Longqi vent field (Fig. 1a). We obtained high-quality genome assemblies of all three parties in the G. aegis holobiont using an adult snail for in-depth analyses of its symbiosis. The host genome is 1.15 Gb in size, comprising 15 pseudo-chromosomes with 94.6% genome completeness (Fig. 1b, Supplementary Table 1, and Supplementary Note 1). This is one of the few chromosome-level genome assemblies currently available in Gastropoda, or even across Mollusca. The genome size of G. aegis is more than double that of C. squamiferum (~444.4 Mb), mainly due to the larger proportion of repetitive regions (Supplementary Fig. 1 and Supplementary Table 2). A total of 21,438 genes (Supplementary Data 1) were predicted from the G. aegis genome, of which 1782 genes were highly expressed in the symbiont-hosting oesophageal gland (Supplementary Data 2). No particular chromosomal distribution bias was observed among these genes (Fig. 1b).

TEM micrographs of the oesophageal gland from G. aegis confirmed the existence of intracellular endosymbionts densely packed inside bacteriocytes (Fig. 2a and Supplementary Fig. 2). Two types of endosymbionts with distinct morphological features were found within a single bacteriocyte (Fig. 2a). Sequences of the 16S ribosomal RNA from the oesophageal gland of three G. aegis individuals were consistent in revealing the two types of symbionts belonging to Gammaproteobacteria, with the more abundant one being a sulfur-oxidising bacteria (SOB hereafter) and the less common one being a methane-oxidising bacteria (MOB hereafter) in the genus Methylomarinum of the family Methylococcaceae (Supplementary Fig. 3). The SOB clustered...
with uncultured bacteria phylogenetically (Supplementary Fig. 4), and thus its taxonomic affinity remains unclear. Of the two symbionts identified from TEM, one exhibited stacked intracellular membranes (Fig. 2a), a key characteristic of Type I methanotrophs, while the other lacked these membranes. Fluorescence in situ hybridisation (FISH) carried out on transverse sections of the oesophageal gland from *G. aegis* yielded positive signals of the probe targeting the SOB, and 1.67% in MOB; Supplementary Table 3). In the metaproteomic analyses of the three parties, 704 proteins were identified in the MOB (Supplementary Data 5), providing additional protein-based evidence to trace the metabolism of the *G. aegis* holobiont.

**Evolution of symbiosis in deep-sea peltospirid snails.** Dated and fossil-calibrated phylogenetic reconstruction based on 529 single-copy genes suggest that the two peltospirid snails, *G. aegis* and *C. squamiferum*, diverged 116.06 million years ago (Ma; 95% confidence intervals: 71.4–183.6 Ma) (Fig. 3 and Supplementary Note 3), which was consistent with a previous study that the ancestor of peltospirids extends to the Early Cretaceous, and the order Neomphalida likely had a mid-Jurassic origin. These early fossils of peltospirids from Early Cretaceous seep deposits had small body size and unlikely to have been symbiotic. As such, the fossil history also supports the findings that *Gigantopelta* and *Chrysomallon* probably evolved symbiosis separately after the current radiation of Neomphalida and Peltospiridae began.

The genomes of both *G. aegis* and *C. squamiferum* snails had 11 *hox* gene clusters in the same order (Supplementary Fig. 6), 13 protein-coding genes of the mitogenome in the same order (Supplementary Fig. 7) and highly consistent macro-synteny patterns (Fig. 4). These results indicate that majority of the gene order and synteny were conserved within the family Peltospiridae, and such clear correspondences exemplify the high quality of both genome assemblies. In addition, the genomic rearrangements revealed by micro-synteny blocks were exclusively intra-chromosomal, rather than inter-chromosomal (Fig. 4). Given the small number of chromosome-level genome assemblies within molluscs, we are unable to test if such a high macro-synteny conservation is the norm within the group for a similar divergence time. Nevertheless, the macro-synteny patterns in these two species are also notably conserved compared with the caenogastropod *Pomacea canaliculata* and the bivalve *Mizuhopecten yessoensis*, distantly related to peltospirid snails within Mollusca. In most molluscan lineages, intra-chromosomal
rearrangements likely underpin the evolution of genetic regulatory networks and related functions, but the underlying mechanism warrants further studies. Among other molluscs with chromosome-level assembly, the oyster *Crassostrea gigas* uniquely has a much reduced genome size and highly derived karyotype30, but it is likely an exception rather than the norm among molluscs.

A third, much smaller, peltospirid snail, *Dracogyra subfuscus*, co-occurs with *G. aegis* and *C. squamiferum* in the Longqi vent field, but anatomical investigation showed that it neither has an enlarged oesophageal gland nor relies on symbionts for nutrition20. Previous phylogeny showed that it is one of many non-symbiotic peltospirid genera separating *Chrysmallon* and...
**Fig. 3** Divergent time of 20 lophotrochozoan taxa. Time-calibrated phylogeny showing the estimated divergence time of 20 lophotrochozoan taxa, with a focus on 18 mussels. Phoronis australis (Phoronida) and Lingula anatina (Brachiopoda) served as outgroups for the molluscs. A total of 529 single-copy orthologues were used. The purple horizontal lines indicate the 95% confidence intervals of the divergence times. Full information on genome resources, fossil records, and geographic events used in the time calibrated nodes (red stars) are provided in Supplementary Note 3. The colour labelling scheme of taxa: Cephalopoda: orange; Bivalvia: purple; Gastropoda: red; Brachiopoda: blue; Phoronida: grey; Mollusca: green; and Neomphalida: black. Ma: millions of years ago.

**Gigantopelta**, and the sister genus of **Gigantopelta**20. We assembled a draft genome of *D. subfuscus* (Supplementary Note 7) and conducted genome binning (Supplementary Fig. 8) to substantiate these claims20. Our results confirmed that *D. subfuscus* is non-symbiotic and much more closely related to *G. aegis* with an average divergent time of 40.13 Ma (Fig. 3) than *C. squamiferum*. These results were in line with previous findings suggesting that *G. aegis* and *C. squamiferum* evolved symbiosis through independent and convergent evolution14,20. Although some *Bathymodiolus* mussels also house dual symbionts of sulfur-oxidising symbionts and methane-oxidising symbionts31, *Bathymodiolus* mussels likely had a common symbiotic ancestor before their radiation in the deep sea32,33.

Despite the phylogenetic divergence between the sulfur-oxidising endosymbionts in the two endosymbiotic peltspird snails *G. aegis* and *C. squamiferum* (Supplementary Fig. 9), the two SOBs were similar in their gene contents encoding for nutritional metabolism, and the same goes for the two host snails (Fig. 5). Phylogenetic divergence of hosts and their SOBs and the nutritional metabolism, and the same goes for the two host snails

**Genetic control of symbiont acquisition.** The host’s immune response is critical for symbiont infection and colonisation. In terms of the gene families shared among *G. aegis* and 18 other lophotrochozoan genomes (Supplementary Fig. 10), several immunity-related gene families were expanded in *G. aegis* (Supplementary Table 4 and Supplementary Note 4). For example, the gene families fucolectin (FUCL) and galectin (Gal) were prominently expanded in *G. aegis* (Fig. 6a); these lectins have been reported as immune-recognition molecules34,35. FUCLs can bind with carbohydrate recognition domains on the bacterial surface in pearl oysters34, and Gals can target microbial non-self glycans and act as recognition receptors in the innate immune response35.

Compared with *C. squamiferum*, *G. aegis* harbours a much greater abundance of pattern recognition receptors, including peptidoglycan recognition proteins (PGRPs), toll-like receptors (TLRs), fibrinogen-related proteins (FBGs), and C-type lectins (CLECs) (Fig. 7). These receptors are thought to help other chemosymbiotic host animals, such as siboglinid tube worms and bathymodiolin mussels, respond to the microbes appropriately, as well as assist in the acquisition and maintenance of their symbiotic populations36. Pattern recognition receptors highly expressed in the oesophageal glands when compared with other organs are likely related to the recognition of the endosymbionts. From gene expression patterns, all PGRPs are highly expressed in the oesophageal glands when compared with other organs in both peltspird snails, suggesting that PGRPs play an important role in the establishment of symbiosis in both snails. A total of 26 genes in *G. aegis* and 53 genes in *C. squamiferum* involved in the lysosome pathway were highly expressed in the oesophageal gland compared with other tissues (Supplementary Figs. 11 and 12). The abundance and high expression of PGRPs and lysosome genes (Fig. 7 and Supplementary Figs. 11 and 12) may help the hosts in regulating the symbiont population36, possibly via the lysosomal digestion process to unlock nutrients in the symbiont biomass.

Many pattern recognition receptors other than PGRPs, particularly those belonging to the TLR gene family, are more highly expressed in the oesophageal gland of *G. aegis* than of *C. squamiferum* (Fig. 7). The TLR proteins serve as key molecular linkages between the host and bacteria in bacterial invasion and host recognition, and they are expanded in the deep-sea
tubeworm *Lamellibrachia luymesi* [37]. These observations indicate that TLR proteins play an important role in symbiosis establishment across different deep-sea animals. Although *G. aegis* occupies the same habitat as *C. squamiferum* in the Longqi vent field, these additional pattern recognition receptors may have enabled *G. aegis* to host two distinct endosymbionts. The host-specific selectivity for different lineages of gammaproteobacterial endosymbionts in peltospirids is again in line with previous evidence indicating that they have independent origins of endosymbioses [14].

There is also genomic evidence that the symbionts of *G. aegis* play a role in the invasion of the host. The peptidoglycan-associated lipoprotein (Pal), OmpA family proteins, and OmpH family proteins are widespread in environmental and symbiotic Gram-negative bacteria living in both marine and non-marine environments according to the NCBI protein database. Pal was expressed in the transcriptomes and proteomes of the *G. aegis* SOB (Supplementary Data 3, Supplementary Data 5), which can be recognised by TLR protein families of the host and further trigger an immune signalling cascade of the host in pathogenesis [38]. Furthermore, both the SOB and the MOB of *G. aegis* contained OmpA family proteins and OmpH family proteins (Supplementary Data 3 and 4), which could interact with TLR and scavenger receptors to help the bacteria invade and adapt to the intracellular environment [39]. These symbiont attributes are likely helpful in establishing an endosymbiotic lifestyle.

**Energy resources.** Previous reports showed that the concentrations of reduced components (i.e. CH₄, sulfur substances, and hydrogen) in the endmember fluid from the Longqi field on the Southwest Indian Ridge are comparable or higher than those in the vents along the Central Indian Ridge [24,40,41]. Although methane is enriched in Longqi compared with other vents on the Central Indian Ridge [41], further biogenic methane input from microbes living below and around the community may be sufficient to sustain the MOB. The host snail has a hypertrophied blood vascular system, which has been suggested to be an anatomical adaptation to transport essential dissolved compounds from the vent fluid together with oxygen from the surrounding seawater, to the SOB and MOB, although no transporters have been identified [14]. Among the transporters of *G. aegis* identified herein (Supplementary Data 6), those transporting oxygen (Supplementary Fig. 13) including one myoglobin, five neuroglobins, and one haemocyanin were highly expressed in the oesophageal gland, but whether they help transport hydrogen sulfide and methane remains unclear. The SOB can utilise thiosulfate, sulfite, sulfide, nitrate, nitrite, oxygen, carbon oxide, and

---

**Fig. 4 Chromosome-scale synteny among four Mollusca genomes.** Chromosome-scale synteny is shown in dot plots with comparisons among the genomes of *Gigantopelta aegis* (15 chromosomes), *Chrysomallon squamiferum* (15 chromosomes), *Pomacea canaliculata* (14 chromosomes), and *Mizuhopecten yessoensis* (19 chromosomes). These plots include the comparison a between *G. aegis* and *C. squamiferum* (comparison within Neomphalida: black), b between *G. aegis* and *P. canaliculata* (comparison within Gastropoda: red), c between *G. aegis* and *M. yessoensis* (comparison within Mollusca: green); d between *C. squamiferum* and *P. canaliculata* (comparison within Gastropoda: red), e between *C. squamiferum* and *M. yessoensis* (comparison within Mollusca: green), and f between *P. canaliculata* and *M. yessoensis* (comparison within Mollusca: green). Each dot represents a synteny block. The numbers on the left and the bottom are the accumulation of the genes in the chromosomes. The colour scheme follows the taxon indicator colours in Fig. 3. C: chromosome. Source data of a–f are provided as a Source Data file.
hydrogen, while the MOB can utilise methane, nitrate, nitrite, oxygen, and hydrogen. These two symbionts form a dual symbiosis system that is highly versatile in utilising chemical substances (Fig. 8a). The central metabolism, associated gene expression levels, and protein abundance levels of the symbionts are summarised in Fig. 8.

Our genome data indicate that sulfur metabolism in the SOB of *G. aegis* (Fig. 8a) is similar to that of the SOBs of *C. squamiferum*, the provannid snails *Alviniconcha kojimai* and *A. strummeri* (*the ‘Gamma-1’ symbiont*), and the *P. echinospica* tubeworm4,6,8. The proteins of *soxA* (35th in proteome), *dsrB* (21th in proteome), *sqr* (34th in proteome), *aprA* (6th in proteome), *aprB* (63th in proteome), and *sat* (57th in proteome) (Fig. 8) show high expression levels, indicating their active engagement in oxidising thiosulfate, sulfitre, and sulfate in the environment and in detoxifying sulfide for the holobiont.

Similar to many sulfur-oxidising endosymbionts of vent molluscs4,5,31, the SOB of *G. aegis* relies on the complete Calvin–Benson–Bassham (CBB) cycle to fix CO2 using a form II ribulose-bisphosphate carboxylase (the protein of *cbbM* gene: 2nd in proteome) recovered with the highest protein expression level and lacks a complete rTCA cycle (Fig. 8). Phosphoribulokinase (the protein of *prkB* gene: 15th in proteome), type I glyceraldehyde-3-phosphate dehydrogenase (the protein of *gapA* gene: 16th in proteome), and transketolase (the protein of *tkt* gene: 43th in proteome) play important roles in the CBB cycle for carbon fixation and exhibit high protein expression levels in the SOB (Fig. 8).

The MOB of *G. aegis* has neither a complete CBB cycle nor an rTCA cycle pathway for CO2 fixation. It possesses the methane monoxygenase operon (*pmoA* gene: 5th in transcriptome, 50th in proteome; *pmoB* gene: 6th in transcriptome, 1st in proteome; *pmoC* gene: 1st in transcriptome, 10th in transcriptome) within the top 1% gene expression (Fig. 8b and Supplementary Fig. 14), confirming that methane oxidation is among the most prominent metabolic processes in the MOB. In the tetrahydromethanopterin pathway, the high expression level of the formaldehyde activating enzyme (*fae* gene: 65th in transcriptome, 20th in proteome) indicates an active generation of CO2 via formaldehyde oxidation. The generated CO2 could be redirected to the CBB cycle of the SOB to form an internal carbon cycle among the symbionts, as detected in the internal symbiotic organs (trophosomes) of vent siboglinid tube-worms42. This phenomenon would allow the symbionts of *G. aegis* to maximise carbon efficiency within the trophosome-like oesophageal glands when in the absence of direct interaction with CO2-rich seawater and vent fluids.

Aerobic methanotrophs are classified into two groups based on their formaldehyde assimilation pathways: Type I (part of Gammaproteobacteria) using a ribulose monophosphate (RuMP) cycle and Type II (part of Alphaproteobacteria) using...
a serine cycle. Methane-oxidising symbionts usually belong to Type I and only use the RuMP cycle; examples include the MOBs of deep-sea *Bathymodiolus* mussels, snails, and sponges (refer to genomes in Supplementary Fig. 9). However, the MOB of *G. aegis* surprisingly possesses a complete serine cycle in addition to a complete RuMP cycle (Fig. 8a). This can mitigate the accumulation and toxicity of formaldehyde when formaldehyde is in excess for the RuMP pathway. In the RuMP cycle, high expression of 3-hexulose-6-phosphate synthase (*hxlA* gene: 20th in transcriptome) and 6-phospho-3-hexulose-6-phosphate (*hxlB* gene: 25th in proteome) in the Entner–Doudoroff pathway indicate an important role for formaldehyde oxidation (Fig. 8). Furthermore, the MOB genome possesses an additional Embden–Meyerhof–Parnas pathway for glycolysis (Fig. 8a), absent in other MOBs of chemosynthetic holobionts reported to date (referred in Supplementary Fig. 9). The co-occurrence of multiple methane assimilation pathways was also present in its free-living relatives within the same family Methylomonadaceae, such as...
Methylomonas methanica, Methylomicrobium alcaliphilum, and Methylomicrobium buryatense (Fig. 8b and Supplementary Table 5). These results indicate that the G. aegis MOB and its free-living relatives are highly efficient and versatile in assimilating carbon sources from methane.

Sulfur-oxidising symbionts of Bathymodiolus mussels, provannid snails in the genera Alviniconcha and Ifremeria, siboglinid tubeworms, as well as the MOBs of Bathymodiolus mussels and Ifremeria snails possess group 1 NiFe hydrogenase (hyaABC)/uptake hydrogenase hupL (Supplementary Table 6), a key gene for hydrogen oxidation43, further confirming that these symbionts are capable of using hydrogen as an additional energy source4,6,8,31. In the present study, the hupL gene was also identified in the genomes of both the SOB and the MOB of G. aegis (Fig. 8 and Supplementary Data 3 and 4). The hupL of both G. aegis symbionts were expressed in the transcriptome but at lower abundance compared with genes responsible for sulfur oxidation and methane oxidation, and they were not found in the metaproteome data possibly due to low abundance. Nevertheless, these results still indicate that both the SOB and the MOB of G. aegis can utilise hydrogen. This study further corroborate the existing hypothesis that the capability of utilising hydrogen for energy is widespread in the symbionts of macrobenthos living in deep-sea chemosynthesis-based ecosystems43.

Fig. 6 Heat maps showing the expanded gene families in Gigantopelta aegis and its two endosymbionts. a Compared with other 15 molluscan taxa (systematic affinity at the class level shown on the right side), Phoronis australis (Phoronida), and Lingula anatina (Brachiopoda) with available genomes (used in the molecular clock analysis), the gene families expanded in G. aegis mainly include immunity-related gene families (Immune), including ankyrin repeat protein (ANK), BTB/POZ domain-containing protein 6 (BTBD6), carcinoma embryonic antigen-related cell adhesion molecule 5 (CEACAMS), E3 ubiquitin-protein ligase (E3), fucolectin (FUCL), galectin (Gal), 5-hydroxytryptamine receptor 4 (HTR4), inhibitor of apoptosis protein (IAP), KRAB-A domain-containing protein 2 (KRBA2), multiple epidermal growth factor-like domain protein (MEGF), thyrotrypsin-releasing hormone receptor (TRHR), and tripartite motif-containing protein (TRIM), as well as transposable elements (TE), including RNA-directed DNA polymerase from mobile element jockey (jockey), retrovirus-related Pol polyprotein from transposon (Pol), reverse transcriptase (RT), transposases, and transposons. b Both the sulfur-oxidising endosymbionts (SOB) and the methane-oxidising endosymbionts (MOB) exhibit transposase expansion. References are 30 symbionts in Gammaproteobacteria (see Supplementary Table 13) from invertebrate taxa living in deep-sea chemosynthetic environments. The systematic affinity of the host taxa is shown on the right side. The colour represents the normalised gene numbers. Gene family abbreviations: NPRA: atrial natriuretic peptide receptor, PTPRK: receptor-type tyrosine-protein phosphatase kappa, ZFHX4: zinc finger homeobox protein 4, Hypothetical: hypothetical protein, Integrase: integrase catalytic region, MTase: N-6 DNA methylase, IS family transposase, IS4: IS4 family transposase, IS4/IS5: IS4/IS5 family transposase, IS5: IS5 family transposase, IS66: IS66 family transposase, IS1595: IS1595 family transposase, IS630: IS630 family transposase, TnpB: TnpB transposase, Hypothetical T: hypothetical transposase. The colour labelling scheme of taxa: Cephalopoda: orange; Bivalvia: purple; Gastropoda: red; Polychaeta: blue; and Porifera: grey. Source data of a and b are provided as a Source Data file.

Fig. 7 Heat maps showing gene abundances and expression levels of pattern recognition receptors in Gigantopelta aegis and Chrysomallon squamiferum. Two heat maps showing gene abundances and expression levels of pattern recognition receptors, including peptidoglycan recognition protein (PGRP), toll-like receptor (TLR), scavenger receptor (SR), fibrinogen-related protein (FBG), galectin (Gal), C-type lectin (CLEC), and fucolectin (FUCL), in Gigantopelta aegis (n = 4) and Chrysomallon squamiferum (n = 3)17 (OsG: oesophageal gland; M: mantle; Ct: ctenidium; FI: internal tissue of foot). The numbers near the lines indicate the gene numbers of the gene families. The colour represents the gene expression level (normalised CPM value). Red: high expression level. Light: low expression level; n means biologically independent animals. Source data are provided in a Source Data file.
Fig. 8 Central metabolism of the dual symbionts of *Gigantopelta aegis*. a Central metabolism pathways in the sulfur-oxidising endosymbiont (SOB) and the methane-oxidising endosymbiont (MOB). The boxes represent genes involved in the respective process, and they are colour coded according to the log-transformed normalised transcripts per kilobase million (TPM) value ($n = 3$) showing the gene expression level in metatranscriptome analysis (colour from red to light white represents high expression level to low expression level; dash arrow indicates that the gene is missing in the genome). b Protein abundances involved in the central metabolism of the SOB (left) and the MOB (right). Protein abundances are reported as mean emPAI values ($n = 3$) assessed in metaproteome analysis; $n$ means biologically independent animals. The description of abbreviations are provided in Supplementary Note 8.
Respiration. Both the SOB and the MOB housed in the oesophageal gland of *G. aegis* can use oxygen and nitrate as electron acceptors for respiration. They are both aerobic and possess succinate dehydrogenase, cytochrome c reductase, and cytochrome oxidase to utilise oxygen (Supplementary Data 3 and 4). Their genomes also encode genes for nitrate reduction, such as nitrate reductase (*napAB* and *nasA* in the SOB and *nargGH* in the MOB), nitrite reductase (*nirBD* in the SOB and *nirBDK* in the MOB), and nitric oxide reductase (*norBC* in both symbionts) (Fig. 8). The symbionts of *C. squamiferum*, deep-sea sulphidinid tubeworms, bathymodulin mussels, vesicomyid clams, provannid snails, and sponges also have the capacity for nitrate respiration (Supplementary Table 7). As the *G. aegis* symbionts rely on the host’s blood vessels for oxygen supply while input from the hydrothermal vent fluctuates due to the constant interplay between the oxygen-poor vent fluid and the oxygen-rich seawater, there may be periods when the oxygen supply runs low in the oesophageal gland. Haemocyanin, haemoglobin, and globin-like genes which play a role in oxygen binding and transport have expression in the foot and mantle, while five sialin genes used for transporting nitrate were highly expressed in the oesophageal gland (Supplementary Fig. 13). Therefore, these symbionts may have developed abilities to avoid oxygen competition with the host, especially under potential periodic hypoxic conditions.

Holobiont nutritional interdependence. Unlike the SOB of *C. squamiferum* which has a complete TCA cycle, the SOB of *G. aegis* lacks the 2-oxoglutarate dehydrogenase gene in the TCA cycle while their homologous genes are present and highly expressed in the MOB. The SOB of *G. aegis* has a C4-dicarboxylate TRAP transport system (DstPQM) that allows the uptake of four-carbon compounds involved in the TCA cycle, indicating that the SOB likely replenishes the missing intermediates of the TCA cycle from one of the other symbiotic partners.

Both the *G. aegis* and *C. squamiferum* host genomes lack capability to biosynthesise six amino acids (isoleucine, valine, leucine, lysine, histidine, and tryptophan), six vitamins (vitamin B1, vitamin B2, vitamin B6, vitamin B12, vitamin B12), and coenzyme A (Fig. 5). Other molluscs with high-quality genomes available, except the shallow water mussel *Modiolus philippinarum*, exhibited the same biosynthetic capacity of amino acids (Fig. 5). Supplementary Fig. 15 shows the phylogenetic relationship of the symbionts and their free-living relatives with sequenced genomes. The nutrient metabolism of both the SOB and the MOB are likely to differ from that of their free-living relatives (Fig. 5).

Most nutrients that the hosts cannot biosynthesise themselves, except vitamin B6 and coenzyme A, can be obtained from their SOBs through the host’s lysis of symbiont cells, as indicated by the high expression of lysosome genes in the oesophageal glands (Supplementary Figs. 11 and 12). The MOB of *G. aegis* has a complete pathway to synthesise vitamin B3 and coenzyme A, whereas the SOBs of the two endosymbiotic peltospirids lack a key gene, namely, 2-dehydrod pantaoate 2-reductase (*panE*), in the biosynthesis of pantothene and coenzyme A (Supplementary Fig. 16). Although the gene may be missing in the technical genome assembly or gene prediction, considering the high-quality of these two assembled genomes with high sequencing coverage of long reads, it is highly likely that this gene is truly missing from the symbiont genomes. Therefore, we speculate two possibilities: (1) the SOB, particularly of *C. squamiferum*, has a replacement of 2-dehydrod pantaoate 2-reductase or replenishment for the bio- synthesis of pantothene and coenzyme A as observed in a peach holobiont, or (2) in the *G. aegis* holobiont, the MOB serves as the major source for providing panthenoate and coenzyme A to both the SOB and the host. In return, the host can provide its symbionts with vitamin B3 and vitamin B6 through the symbionts’ ATP-binding cassette transporters of amino acids. These host-supplied nutrients are critical for fulfilling the nutritional demands of the endosymbionts housed in an internal organ that lacks direct contact with the outside world.

Transposase expansion in three partners of the *G. aegis* holobiont. Enrichment of transposable elements (TEs) provides beneficial genomic plasticity for adaptation to an intracellular, symbiotic lifestyle by regulating nearby genes. In the results from the gene family analysis, the expansion of TEs was found in all three partners comprising the *G. aegis* holobiont (Fig. 6). For the host snail, these expansions corresponded to 342 transposable genes (1.6% of total host genes) and 510 DNA transposons (2.4% of total host genes); for the SOB, these expansions corresponded to 704 transposable genes (12.8% of total SOB genes); and for the MOB, these expansions corresponded to 345 transposable genes (11.1% of total MOB genes). Previous studies reported that transposable genes were particularly enriched in bacterial lineages that have recently transitioned to a host-associated lifestyle, indicating that enrichment of transposable genes may have a beneficial effect for the symbionts of *G. aegis* in becoming endosymbionts. Moreover, numerous transposable genes were unusually inserted into the flagellar operons and chemotaxis operons that play important roles in the host infection of both the SOB and the MOB of *G. aegis*. These genes may contribute to the regulation of the symbionts’ motility for infection and colonisation in the host environment. To our knowledge, the expansion of transposable genes occurring across multiple symbiotic partners within the same holobiont has not been reported in any marine invertebrate holobiont.

Overall, we revealed that the deep-sea vent peltospirid snail *G. aegis* exhibits a dual symbiotic lifestyle, housing a more abundant sulfur-oxidising Gammaproteobacteria and a less abundant methane-oxidising Gammaproteobacteria (Type I methanotroph). Furthermore, we provide a high-quality hologenomome including a chromosome-level assembly for the snail host and an in-depth analysis of the genomic interdependencies among the ‘tripch’ of symbiotic partners in this holobiont, finding an intimate mutualistic relationship with complementarity in nutrition and metabolic co-dependency. The two endosymbionts live in bacteriocytes of the oesophageal gland, an organ lacking direct interaction with the ambient environment. Both symbionts and the host are highly versatile with regard to the transportation and utilisation of chemical energy, increasing the efficiency of carbon fixation by forming an internal carbon cycle between the two symbionts. Although *G. aegis* and another chemosymbiotic peltospirid snail, the Scaly-foot Snail, *C. squamiferum*, occupy the same habitat in the Longqi vent field, *C. squamiferum* has a single sulfur-oxidising symbiont. Our data support previous findings that the two peltospirid snails evolved symbiosis independently and convergently, but their phylogenetically distinct sulfur-oxidising endosymbionts have the same capabilities in the biosynthesis of specific nutrients. As such, nutrient biosynthesis capacity may be a key constraint in the selection of symbionts by these peltospirids.

Methods

Deep-sea sampling. *Gigantopelta aegis* snails were collected by the manned submersible HON Jiaolong on-board the RV Xiangyanghong 9 cruise 351 from the Longqi hydrothermal vent field (37°73′39″S, 49°65′0″E, 2763 m depth) on the Southwest Indian Ridge in January 2015. Snails were immediately flash-frozen in liquid nitrogen once recovered on the ship. They were then transferred to −80 °C.
until DNA and RNA extraction. Specimens of G. aegis for Hi-C sequencing and FISH experiments were also collected from Longqi (Tiaam’ chimney) during Leg 3 of the COMRA R/V Dayang Yihao expedition 52 in April 2019. Tissue from the foot was immediately dissected from one individual, cut up, washed in phosphate-buffered saline buffer, and stored at −80 °C until Hi-C library preparation. The oesophageal gland tissue was dissected from the same individual and fixed in 4% paraformaldehyde overnight, followed by dehydration through an ethanol series (20%, 50%, 70%, 90%, and 100%, for 15 min each) and stored at −80 °C until experiments. Samples for TEM were immediately fixed in 10% buffered formalin after recovery on board of the research vessel. The identity of the specimens was confirmed by both mitochondrial COI sequences and morphology. Detailed information on how each individual was used is provided in Supplementary Table 8.

**TEM.** The oesophageal gland tissue was dehydrated using a series of graded acetone and then transferred to Epon resin (Sigma-Aldrich) for embedding. An ultramicrotome (Reichert Ultracut S, Leica) was used to slice ultrathin (70 nm) sections that were used for staining in 2% aqueous uranyl acetate with lead stain solution (0.3% lead acetate and 0.3% lead nitrate, Sigma-Aldrich). The presence of intra-cellular symbionts was confirmed using a Tecnai 20 Transmission Electron Microscopy (FEI) at an acceleration voltage of 120 kV. For more detailed methods for thin-section preparation, see ref. 14.

**FISH.** Oesophageal gland tissues of G. aegis were dehydrated in 100% ethanol and embedded in Shandon Cryomatrix Frozen Embedding Medium (6769006; Thermo Fisher Scientific). After full embedding, they were quickly frozen in dry ice. Cryostat sections (10 μm thickness, Thermo Fisher Scientific) were cut and mounted on glass slides. Two 16S rRNA-based probes were designed to target the endosymbionts specifically. The design of FISH probes for the SOB and the MOB of G. aegis was based on their respective 16S rRNA gene sequences. To ensure that the 20 bp probe was unique and specific, the 16S rRNA gene sequence was searched against online standard databases of nucleotide collection in NCBI. The alignments among the 16S rRNA genes of symbionts and their five best hit sequences with highest similarity sequence were manually checked. The aligned region with the largest difference was manually chosen for probe designs. The uniqueness of the probe sequences was further confirmed by searching against the NCBI nt databases. After confirming the specificity of the 20 bp sequence regions of the SOB and the MOB, 20 bp oligonucleotide probes with 100% match to the regions were synthesised, and their 3’ ends were labelled with fluorescent dye Cy3 and Cy5, respectively. The Cy3-labelled SOB1 probe (5’-AGCATATTTAACCTTGACCC-3') was used to target the sulfur-oxidising endosymbiont, and the Cy5-labelled MOB1 probe (5’-CTGGTTGTTTTTCCCTGCT TCT-3') was used to bind the methan-oxidising endosymbiont (Supplementary Table 9). The sections were rehydrated in a decreasing ethanol series (95%, 80%, and 70%) for 15 min each and hybridised at 46 °C with 50 ng/ml of each probe in a hybridisation buffer (0.9 M NaCl, 0.02 M Tris-HCl, 0.01% sodium dodecyl sulfate, and 0.2% IGEPAL CA-630) overnight. The slides were washed twice in a wash buffer (0.1 M Tris-HCl, 0.02 M Tris-HCl, 0.01% sodium dodecyl sulfate, and 5 mM EDTA) at 48 °C for 15 min. Two drops of 4',6-diamidino-2-phenylindole (DAPI) were added on each slide, followed by incubation at room temperature for 3 min. After washing and air drying, the slides were mounted with SlowFade® Diamond antifade medium (Invitrogen, Carlsbad, CA, USA). Images were obtained using a confocal microscope (Leica Microsystems, Wetlazr, Germany) with a 3D model and post-processed by LAS X software version 3.0.13 (Leica Microsystems, Wetlazr, Germany).

**DNA extraction and Illumina sequencing of G. aegis holonome.** The MagAttract High-Molecular-Weight DNA Kit (QIAGEN, Hilden, Netherlands) was used to extract high-molecular-weight genomic DNA from the oesophageal gland and the foot separately for sequencing of the genome of the symbionts and the host, respectively, following manufacturer protocols. The extracted DNA was further purified by a Genomic DNA Clean & ConcentratorTM-10 kit (ZYMOS Research, Irvine, CA, USA). The DNA was finally eluted in 10 mM Tris-HCl buffer (pH 8.5). The quality of the genomic DNA was evaluated using a Bioanalyzer pILTE (BioDevel, Tampa, FL, USA) with the OD 260 vs 280 of the DNA samples ranged from 2.0 to 2.2. The DNA concentration was assessed using a QubitTM 3 Flurometer (Thermo Fisher Scientific, Singapore). The sizes of DNA fragments were assessed by pulsed-field gel electrophoresis. Approximately 1 μg of DNA of each tissue, including foot and oesophageal gland, was used for the short-insert library (350 and 500 bp) in an Illumina NovaSeq 6000 platform to generate 171 and 49 Gb paired-end reads with a length of 150 bp, respectively.

**Oxford Nanopore Technologies (ONT) library preparation and MinION sequencing.** The DNA of the oesophageal gland was used to prepare the long-read library (Oxford Nanopore MinION, UK) for genome sequencing of endosymbionts. To construct the MinION library, a NEBNext Ultra II End-Repair/dA-tailing kit (NEB#7546) was used for preparation of DNA repair and dA-tailing. The resulting library was purified using a MinELute PCR Purification Kit (QIAGEN) and quantified using a Qubit3 Fluorometer. The library was sequenced using the Oxford Nanopore Technologies (ONT) SQK-LSR109 kit. The ONT sequencing was performed by Oxford Nanopore Technologies Limited, UK, following the manufacturer’s protocols. The output was mapped to the genome of G. aegis, the 16S rRNA gene sequence was searched against online standard databases of nucleotide collection in NCBI. The sizes of DNA fragments were assessed by pulsed-field gel electrophoresis. Approximately 1 μg of DNA of each tissue, including foot and oesophageal gland, was used for the short-insert library (350 and 500 bp) in an Illumina NovaSeq 6000 platform to generate 171 and 49 Gb paired-end reads with a length of 150 bp, respectively.
corMaxEvidenceErate = 0.15, correctedErrorRate = 0.085, minReadLength = 8000 and then used in genome assembly using wtdbi2 version 2.1 (ref. 77) with settings: -e 2 -s ugly reads 5000 -S 1 -k 15 -p 0 -r rescue-low-cov-edges -a1 -nulnsok. To improve the accuracy of the initially assembled contigs, two polishing rounds using Racon version 1.3.1 (ref. 66) were performed and followed by an additional round using Pilon version 1.22 (ref. 67) with Illumina short reads of the foot tissue mapped by Bowtie2 version 2.3.5 (ref. 33) in the -sensitive mode. Bacterial contamination in the host genome assembly was further filtered using MaxBin version 2.2.5 (ref. 62).

Raw Hi-C sequencing reads were first trimmed by Trimomatic version 0.36 (ref. 69). To remove the invalid pairs of Hi-C reads without effective ligation, the reads were mapped to the correct assembled contigs using Bowtie2 version 2.3.5 (ref. 53), and Hi-C contact maps were generated based on the mapped reads using HiC-Pro version 2.11.1 (ref. 61). Juicer version 1.5 (ref. 65) was used to filter and deduplicate. The remaining valid reads were used for contig scaffolding using the 3D de novo assembly (3D-DNA) pipeline version 180114 (ref. 60) in diploid mode. The scaffolds were manually corrected based on the Hi-C contact maps (Supplementary Fig. 18). The continuity and completeness of the assembled genome were evaluated with assemblathon_stats.pl (https://github.com/ucdavis-bioinformatics/assemblathon2-analysis/blob/master/assemblathon_stats.pl) and BUSCO version 3.0.2 (ref. 69) using the metazoa_odbl10 database, respectively.

**Gene predictions and functional annotation.** The coding sequences and proteins of symbionts genomes were predicted by Prodigal version 2.6.3 (ref. 65) with default settings. The host genome was first hard-masked in the repetitive regions and then trained on the introns and exons to predict coding genes of *G. aegis* (for details see Supplementary Note 2).

Repeats in the *G. aegis* genome were de novo identified and classified using RepeatModeler version 1.0.11 (http://www.repeatmasker.org/RepeatModeler/) pipeline and masked using RepeatMasker version 4.0.8 (http://www.repeatmasker.org/RMDownload.html) with the parameter -xsmall (for details see Supplementary Note 2).

To predict the gene models, we ran two rounds of the MAKER pipeline version 2.31.10 (ref. 69) on the soft-masked genome. To guide the prediction, we de novo assembled transcripts, Metazoa protein sequences downloaded from the Swiss-Prot database, and protein sequences of *C. squamiferum* (23) (for details see Supplementary Note 2).

The predicted protein sequences were searched against the NCBI Non-Redundant (NR) database using BLASTp with an E-value cutoff of 1e-5. The sequences were also used to search the Kyoto Encyclopedia of Genes and Genomes (KEGG) database usingblast2GO. The genomic sequences were also used to improve the accuracy of the initially assembled contigs, two polishing rounds using Racon version 1.3.1 (ref. 66) were performed and followed by an additional round using Pilon version 1.22 (ref. 67) with Illumina short reads of the foot tissue mapped by Bowtie2 version 2.3.5 (ref. 53), and Hi-C contact maps were generated based on the mapped reads using HiC-Pro version 2.11.1 (ref. 61). Juicer version 1.5 (ref. 65) was used to filter and deduplicate. The remaining valid reads were used for contig scaffolding using the 3D de novo assembly (3D-DNA) pipeline version 180114 (ref. 60) in diploid mode. The scaffolds were manually corrected based on the Hi-C contact maps (Supplementary Fig. 18). The continuity and completeness of the assembled genome were evaluated with assemblathon_stats.pl (https://github.com/ucdavis-bioinformatics/assemblathon2-analysis/blob/master/assemblathon_stats.pl) and BUSCO version 3.0.2 (ref. 69) using the metazoa_odbl10 database, respectively.

**Draft genome of *D. subfuscus*.** *Dracoctyra subfuscus* were collected from Longqi (Tiantai’ chimney) during Leg 3 of the COMRA R/V Dayang Yihao expedition 52 in April 2019, fixed in RNAlater and transferred to −80 °C freezer until use. The whole tissue of one individual of *D. subfuscus* was collected from Longqi expedition 52. As the species used (*G. aegis*), the 16S rRNA gene alignment generated by MUSCLE version 3.8.31 (ref. 76) was constructed for both the SOB and the MOB of *G. aegis* using IQ-TREE multicore version 1.6.10 (ref. 77) with -m MFP and 1000 ultrafast bootstraps, respectively. To compare the essential metabolism pathway (e.g. nutrient biosynthesis and essential carbon metabolism) with their free-living relatives with available genomes, their genomes were used for phylogenetic and metabolic pathway analyses. To explore the phylogenetic relationship of the symbionts, 30 genomes of bacterial symbionts belonging to Gammaproteobacteria from deep-sea invertebrate taxa were included (for details of referred symbionts, see Supplementary Table 13). Same pipelines of the orthologue cluster of *G. aegis* were applied in the symbiont analyses. MUSCLE version 3.8.31 (ref. 76) was used for protein alignments. The alignments of single-copy orthologues were concatenated for subsequent phylogeny analysis. The IQ-TREE multicore version 1.6.10 (ref. 77) with settings of LG-1+F+G4+I and 1000 ultrafast bootstraps were used to perform the phylogeny analysis of the symbionts.

**Gene expression analysis of the holobiont.** The coding DNA sequences of the predicted genes were used as references to assess gene expression. Kallisto version 0.45.1 (ref. 69) was used to calculate transcripts per million (TPM) value based on the mapped reads count for each gene. In symbionts, the TPM values were directly used as gene expression values. For the host, the highly expressed genes in the oseophageal gland were identified by differential expression analysis versus foot, ctenidium, and mantle (*n* = 4) using edgeR (81) based on the reads counts. If genes had a fold change of reads counts larger than two and a significant FDR *P* value (<0.05) generated from edgeR (81), they were designated as highly expressed genes. To compare the gene expression of *G. aegis* and *C. squamiferum*, the same analysis pipeline was applied to the transcriptome sequencing data of foot tissue, ctenidium, mantle, and oseophageal gland from *C. squamiferum* (22). The heat maps were plotted based on the counts per million (CPM) value normalised by the STANDARDIZE function complemented in Excel.

**Synteny analysis.** Pairwise synteny blocks were searched using LAST version 1080 (83) with a cutoff of 1%, and the blocks were filtered to keep blocks with an overlap of more than 50% of the first 1000 nucleotides of extracted protein from each sample and stained by colloidal Coomassie blue. The peptide for LC-MS/MS was obtained through protein reduction, alkylation and digestion, peptide extraction, and further drying. Dionex UltiMate 3000 BSC-LCn coupled with an Orbitrap Fusion Lumos Mass Spectrometer (Thermo Fisher) was utilised to perform the mass spectrometry analysis. The search database contains the protein sequences predicted from the genome and the corresponding reversed sequences (decoy) of both *G. aegis* and its two endosymbionts. Mascot version 2.3.0 was used to identify and quantify the protein via the raw mass spectrometry data. Proteins were identified with the assigned peptides’ identification confidence level over 0.95 and false discovery rate of less than 2.5% (see Supplementary Note 5).

**Ethics declarations.** As the species used (*G. aegis* and *D. subfuscus*) are invertebrate gastropod molluscs, no ethical approval or guidance was required. No cephalopod samples were used in this study. Research cruises and their collecting activities were authorised and approved by the China Ocean Mineral Resources Research and Development Association.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

All sequencing data and assembly data of *G. aegis* and its two symbionts were deposited to the National Centre for Biotechnology Information (NCBI) database under BioProject PRJA612619. Detailed data information of *G. aegis* holobiont was provided in Supplementary Tables 8 and 11. All sequencing data and assembly data of *G. aegis* were also deposited to the NCBI database under BioProject PRJA6180542. Assembly data and genome annotation data of both *G. aegis* and *D. subfuscus* are available in figshare (https://doi.org/10.6084/m9.figshare.13317932 and https://doi.org/10.6084/m9.figshare.13317932.v1). The genome assembly and predicted gene models of *G. aegis* can also be found at MolluscDB (http://mbase.qlm.cn/gene/download/download). The
metaproteomic data of G. aegis are available in PRIDE via ProteomeXchange under the identifier PXD022852. Publicly available datasets used in the study include the following: NCBI NR database (https://www.ncbi.nlm.nih.gov/nuccore/), KEGG (https://www.genome.jp/kegg/), Repbase (https://www.girinst.org/repbase/), EuKaryotic Orthologous Groups (KOG) database (https://mycocosm.jgi.doe.gov/help/kogbrowser.jsf), Clusters of Orthologous Groups (COG) database (https://www.ncbi.nlm.nih.gov/COG/), Pfam database (https://pfam.xfam.org), RAST server (http://rast.theeb.org/), SwissProt database via UniProt (https://www.uniprot.org/), and the ecocyc database (https://ecocyc.org). Source data are provided with this paper.

Code availability
All computer commands, codes, and intermediate files for the bioinformatics analyses are based on the available software listed in the ‘Methods’ section, which are provided at https://github.com/ylan5740/Commands-for-homologous-analyses under the GNU General Public License v3.0.

Received: 4 May 2020; Accepted: 7 January 2021; Published online: 19 February 2021

References
1. Robbins, S. J. et al. A genomic view of the reef-building coral Porites lutea and its microbial symbionts. Nat. Microbiol. 4, 2090–2100 (2019).
2. Dubilier, N., Bergin, C. & Lott, C. Symbiotic diversity in marine animals: the sea as a habitat for endosymbionts. Nat. Rev. Microbiol. 7, 75–80 (2009).
3. Newton, I. L. G. et al. ThePorites lutea symbiont genome reveals its extraordinary circulatory system. Science 319, 580–583 (2008).
4. Wiwattanapatapee, R. & Kajiwara, K. Endosymbiont diversity in marine animals: the sea as a habitat for endosymbionts. Nat. Rev. Microbiol. 14, 1–12 (2017).
5. Newton, I. L. G. et al. ThePorites lutea symbiont genome reveals its extraordinary circulatory system. Science 319, 580–583 (2008).
6. Newton, I. L. G. et al. ThePorites lutea symbiont genome reveals its extraordinary circulatory system. Science 319, 580–583 (2008).
7. Fujiwara, Y., Kato, C., Masui, N., Fujikura, K. & Kojima, S. Dual symbiosis in a land snail: the sea as a habitat for endosymbionts. Nat. Microbiol. 2, 965–969 (2017).
8. Sun, J. et al. Signatures of divergence, invasiveness, and terrestrialization revealed by four apple snail genomes. Mol. Biol. Evol. 36, 1507–1520 (2019).
9. Wang, S. et al. Scallop genome provides insights into evolution of bilaterian anatomy and development. Nat. Ecol. Evol. 3, 1–12 (2017).
10. Miyazaki, J. I., de Oliveira Martins, L., Fujita, Y., Matsumoto, H. & Fujiwara, Y. E. Endosymbiosis in the chemosynthetic snailGastropoda: Neomphalina) reveals its extraordinary circulatory system. Proc. Natl. Acad. Sci. U.S.A. 116, 5503–5508 (2019).
11. Petersen, J. M. et al. Hydrogen is an energy source for hydrothermal vent mussels. Proc. Natl. Acad. Sci. U.S.A. 116, 9515–9520 (2019).
12. Petersen, J. M. et al. Hydrogen is an energy source for hydrothermal vent mussels. Proc. Natl. Acad. Sci. U.S.A. 116, 9515–9520 (2019).

ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-21450-7 | www.nature.com/naturecommunications

21. Chen, C., Linse, K., Uematsu, K. & Sigwart, J. D. Cryptic niche switching in a chemosymbiotic gastropod. Proc. Biol. Sci. 285, 1099 (2018).
22. Sun, J. et al. The sea as a habitat for endosymbionts. Nat. Rev. Microbiol. 14, 1–12 (2017).
23. Heywood, J. L., Chen, C., Pearce, D. A. & Linse, K. Bacterial communities associated with the Southern Ocean vent gastropod, Gigantopelta choesia: indication of horizontal symbiont transfer. Polar Biol. 40, 3373–3427 (2017).
24. Sun, J. et al. Dual high-temperature hydrothermal circulation in a detachment faulting system on the ultra-slow spreading ridge. Nat. Commun. 11, 1300 (2020).
25. Miyazaki, J. I. et al. Dual energy metabolism of the Campylobacteria endosymbiont in the chemosynthetic snail Achnincola marisindica. ISME J. 14, 1273–1289 (2020).
26. Salerno, J. L. et al. Characterization of symbiont populations in life-history stages of mussels from chemosynthetic environments. Biol. Bull. 208, 145–155 (2005).
27. Kaim, A., Jenkins, R. G., Tanabe, K. & Kiel, S. Molusks from late Mesozoic seep deposits, chiefly in California. Zootaxa 3861, 401–440 (2014).
28. Sun, J. et al. Signatures of divergence, invasiveness, and terrestrialization revealed by four apple snail genomes. Mol. Biol. Evol. 36, 1507–1520 (2019).
29. Wang, S. et al. Scallop genome provides insights into evolution of bilaterian anatomy and development. Nat. Ecol. Evol. 3, 1–12 (2017).
30. Miyazaki, J. I., de Oliveira Martins, L., Fujita, Y., Matsumoto, H. & Fujiwara, Y. Evolutionary process of deep-sea mussels. Proc. Sci. 1243 (2013), 2013.
31. Miyazaki, J. I., de Oliveira Martins, L., Fujita, Y., Matsumoto, H. & Fujiwara, Y. Evolutionary process of deep-sea mussels. Proc. Sci. 1243 (2013), 2013.
32. Lorion, J. et al. Adaptive radiation of chemosymbiotic deep-sea mussels. Proc. R. Soc. B Biol. Sci. 280, 1243 (2013), 2013.
33. Miyazaki, J. I., de Oliveira Martins, L., Fujita, Y., Matsumoto, H. & Fujiwara, Y. Evolutionary process of deep-sea mussels. Proc. Sci. 1243 (2013), 2013.
34. Godlew ska, R., Wizniowska, K., Petras, Z. & Jaguszyn-Krynicka, E. K. Peptidoglycan-associated lipoprotein (Pal) of Gram-negative bacteria: function, structure, role in pathogenesis and potential application in immunophrophylaxis. FEMS Microbiol. Lett. 298, 1–11 (2009).
35. Jeannin, P. et al. Complexity and complementarity of outer membrane protein A recognition by cellular and humoral innate immunity receptors. Immunity 22, 551–560 (2005).
36. Kawaguchi, S. et al. Fluid chemistry in the Solitario and Dodo hydrothermal fields of the Central Indian Ridge. Geofluids 16, 988–1005 (2016).
37. Popp, J. et al. Soil biota from the active Longqi hydrothermal field, Southwest Indian Ridge. Deep Sea Res. Part I Oceanogr. Res. Pap. 109, 41–47 (2015).
38. Fisher, C. R., Pennickett, M. C. & Brooks, J. M. Stable carbon isotopic evidence for carbon limitation in hydrothermal vent vestimentifera. Science 247, 1094–1096 (1990).
39. Petersen, J. M. et al. Hydrogen is an energy source for hydrothermal vent symbiosis. Nature 476, 176–180 (2011).
40. Price, D. R. & Wilson, A. C. A substrate ambiguous enzyme facilitates genome reduction in an intracellular symbiont. BMC Biol. 12, 110 (2014).
41. Kleiner, M., Young, J. C., Shah, M., VerBerkmoes, N. C. & Dubilier, N. Metaproteomics reveals abundant transposase expression in mutualistic endosymbionts. mBio 4, e00223–13 (2013).
42. Newton, I. L. & Bordenstein, S. R. Correlations between bacterial ecology and mobile DNA. Curr. Microbiol. 62, 198–208 (2011).
43. Wick, R. R., Judd, L. M. & Holt, K. E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol. 20, 129 (2019).
44. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293 (2009).
45. Bolger, A. M., Lohse, M. & Usadel, B. Trimmmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).
46. Bankevich, A. et al. SPAdes: a genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477 (2012).
47. Albertsen, M. et al. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat. Biotechnol. 31, 533–538 (2013).
52. Tian, R. M. et al. The deep
NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-21450-7 ARTICLE
53. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2.
56. Boetzer, M. & Pirovano, W. SSPACE-LongRead: scaffolding bacterial draft
58. Kie
61. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial
62. Wu, Y. W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning
63. Servant, N. et al. HiC-Pro: an optimized and
effective stochastic algorithm for estimating maximum-likelihood
phylogenies.
Mol. Biol. Evol. 21, 157 (2015).
65. Dudchenko, O. et al.
69. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C. & Kanehisa, M. KAAS: an
automatic genome annotation and pathway reconstruction server.
Genome Res. 15, 158 (2005).
72. Genaro, M., Frías, M., Cuéllar, C., Soto, T., Marroqui, N. & Zelaya, Y. Genome
Assembly and annotation using BUSCO: assessing the completeness of
assembly of animal genomes. Mol. Biol. Evol. 37, 397–408 (2020).
73. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2.
59. Ruan, J. & Li, H. Fast and accurate long-read assembly with wtdbg2.
60. Vaser, R., Sovi
67. Ragan, B. L., Knight, R., Jakob, C. R. & de la Peña, L. Toward community-level
phylogenetic analysis and classification of microbial communities.
Mol. Ecol. Res. 16, 349 (2016).
74. Kajitani, R. et al. Efficient de novo assembly of highly heterogenous genomes from single cells, and metagenomes.
Genome Res. 27, 737–746 (2017).
75. Ruan, J. & Li, H. Fast and accurate long-read assembly with wtdbg2.
Nat. Methods 17, 155–158 (2019).
76. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and
table editing. All authors contributed to the manuscript. C.C., J.S., J.-W.Q., and C.V.D. contributed to manu-
script production. C.C., J.S., and J.-W.Q. performed the gene prediction of
Gigantopelta aegis and Dracocephalum subfuscum samples. C.C., performed TEM
examination of D. subfuscum. C.C., performed the DNA extraction, RNA extraction, ONT sequencing, gene expression
analysis, family analysis, synteny analysis, and immune response analysis as well as drafted the manuscript. C.C., J.S., J.-W.Q., and C.V.D. contributed to manu-
script editing. All authors contributed to the manuscript and approved it for submission
and publication.
Competing interests
The authors declare no competing interests.
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-21450-7.
Correspondence and requests for materials should be addressed to P.-Y.Q.
Peer review information Nature Communications thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permission information is available at http://www.nature.com/reprints
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2021