PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Perez M, Breusing C, Angers B, Beinart RA, Won Y-J, Young CR. 2022 Divergent paths in the evolutionary history of maternally transmitted clam symbionts. *Proc. R. Soc. B* **289**: 20212137. https://doi.org/10.1098/rspb.2021.2137

Received: 29 September 2021
Accepted: 14 February 2022

Subject Category:
Evolution

Subject Areas:
ecology, evolution, genomics

Keywords:
symbiosis, reductive genome evolution, deep-sea, niche partitioning, selection

Author for correspondence:
Maëva Perez
e-mail: maeva.perez@umontreal.ca

†Authors contributed equally.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5882707.

THE ROYAL SOCIETY PUBLISHING

Divergent paths in the evolutionary history of maternally transmitted clam symbionts

Maëva Perez[1,†], Corinna Breusing[2,†], Bernard Angers[1], Roxanne A. Beinart[2], Yong-Jin Won[3] and C. Robert Young[4]

[1]Department of Biological Sciences, Université de Montréal, Montreal, Canada
[2]Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA
[3]Division of EcoScience, Ewha Womans University, Seoul, South Korea
[4]National Oceanography Centre, Southampton, UK

MP, 0000-0003-4532-8299; CB, 0000-0001-6845-0188; BA, 0000-0002-0992-140X; RAB, 0000-0003-1672-5957; Y-JW, 0000-0003-1343-8418

Vertical transmission of bacterial endosymbionts is accompanied by virtually irreversible gene loss that results in a progressive reduction in genome size. While the evolutionary processes of genome reduction have been well described in some terrestrial symbioses, they are less understood in marine systems where vertical transmission is rarely observed. The association between deep-sea vesicomyid clams and chemosynthetic Gammaproteobacteria is one example of maternally inherited symbioses in the ocean. Here, we assessed the contributions of drift, recombination and selection to genome evolution in two extant vesicomyid symbiont clades by comparing 15 representative symbiont genomes (1.017–1.586 Mb) to those of closely related bacteria and the hosts' mitochondria. Our analyses suggest that drift is a significant force driving genome evolution in vesicomyid symbionts, though selection and interspecific recombination appear to be critical for maintaining symbiont functional integrity and creating divergent patterns of gene conservation. Notably, the two symbiont clades possess putative functional differences in sulfide physiology, anaerobic respiration and dependency on environmental vitamin B12, which probably reflect adaptations to different ecological habitats available to each symbiont group. Overall, these results contribute to our understanding of the eco-evolutionary processes shaping reductive genome evolution in vertically transmitted symbioses.

1. Introduction

Heritable symbioses with intracellular bacteria are observed across the eukaryotic domain of life [1]. These symbioses have profound consequences for both host and symbiont, by altering sex ratios in a population, providing nutrients that are otherwise unavailable in the host's habitat, or enhancing resistance to predators and pathogens [1,2]. Vertical transmission of bacterial lineages from parent to offspring inevitably leads to reductive genome evolution (RGE) in the symbionts [2,3]. This process results from successive bottleneck events during transovarial transmission, which decrease the effective population size and genetic diversity of endosymbiont populations [4]. The genetic homogeneity of vertically transmitted symbionts is further amplified by reduced rates of horizontal gene transfer (i.e. homologous recombination between bacterial lineages), which decrease with higher degrees of host restriction [5]. Consequently, genetic drift increases relative to selection in these taxa, favouring the accumulation of slightly deleterious mutations (Muller's ratchet) [6]. The pea aphid–*Buchnera* symbiosis and several other well-studied insect–bacteria models support this neutral hypothesis [7], whereas other metazoan/microbial symbioses highlight the importance of selection in shaping RGE. For instance, Red Queen–King dynamics are predicted to maintain specificity and the functioning of cytonuclear interactions between host and symbiont [2,8].

Furthermore, symbiont traits that are beneficial for the host might experience increased selective pressures for maintenance of function and/or diversification, while selection may be relaxed on genes that are functionally redundant or no longer necessary for symbiont survival. Thus, differences in gene content among related symbionts can reveal how host–symbiont pairs diverged in their ecological niches over evolutionary time [9]. Ultimately, niche differentiation mediated by differential gene loss has the potential to influence host community structure through habitat partitioning and host evolution through ecological speciation. Despite its importance for ecological and evolutionary processes, there is still a significant gap in our understanding of the selective processes influencing patterns of genome reduction in vertically transmitted bacteria. This is especially true for the heritable endosymbionts of marine organisms, since vertical transmission is less common in aquatic symbioses [1].

Relatively strict vertical transmission of bacterial endosymbionts has been observed in deep-sea clams of the family Vesicomyidae (subfamily Pliocardiinae) [10], providing an opportunity to examine neutral and selective processes shaping RGE in the marine environment. Vesicomyid clams represent the most diverse group of deep-sea bivalves, with 173 described species present in reducing habitats ranging from hydrocarbon seeps on continental margins to hydrothermal vents on mid-ocean ridges [11,12]. All symbiont-bearing taxa are nutritionally dependent on their chemosynthetic gammaproteobacterial partners, which derive chemical energy from the oxidation of reduced sulfur compounds to produce nutrition for their hosts [13,14]. Symbiont capture was likely a single event that happened before their radiation about 45 Ma [11,15], an acquisition that is much more recent than that of well-studied terrestrial symbioses (approx. 100–200 Ma) [16,17]. Based on ribosomal sequence data, vesicomyid symbionts are classified into two divergent clades: Clade I (associated with hosts of the *gigas*-group), and Clade II (associated with all other vesicomyid hosts) [11,18]. Topological congruences between host mitochondrial and symbiont phylogenies indicate that symbionts co-evolve with their hosts [10], although disruptions of these relationships occur through infrequent horizontal transmission events that allow for recombination between bacterial lineages [19–23]. Previous analyses of one representative symbiont lineage from each clade (*Candidatus* Ruthia magnifica for Clade I and *Ca.* Vesicomyosocius okutanii for Clade II) suggest that RGE is ongoing in vesicomyid symbionts and that Clade I is in a more advanced state of genome reduction than Clade II [24]. *Ca.* Ruthia magnifica and *Ca.* Vesicomyosocius okutanii possess intermediate genome sizes (1.16 Mbp versus 1.02 Mbp) and levels of AT enrichment (66% versus 68%) compared to other host–restricted symbionts, while contrasting levels of gene decay and GC content for 10 housekeeping genes were observed across their respective clades [25].Variations in host affiliation and genome reduction between symbiont clades do not appear to be driven by adaptation to different broad-scale habitat types, as host species of both clades have been found at hydrothermal vents and hydrocarbon seeps and often co-occur at the same locality [11,14,26,27]. However, limited genetic data suggest that the two symbiont clades differ in physiological characteristics related to nitrate reduction and sulfur metabolism, which may affect microhabitat exploitation [14,26], and could, thus, influence patterns of gene conservation. In fact, niche partitioning has been linked to patterns of gene loss in a variety of marine and freshwater bacteria [28,29].

In this study, we assessed the contributions of neutral and selective processes to RGE in vesicomyid symbionts by comparing their genome characteristics to those of outgroup bacterial relatives and the hosts' mitochondria. We tested the hypothesis that genetic drift is a significant force driving RGE in these symbionts and determined to what extent selection has shaped their genetic makeup over evolutionary times.

2. Material and methods

Detailed methods are available in the electronic supplementary material.

(a) Genome analyses

New mitochondrial and symbiont genomes for 11 lineages of vesicomyid clams were assembled and annotated in this study, while genomes for another four species were retrieved from previous publications [5,30–37] (figure 1; electronic supplementary material, figure S1). Bacterial relatives of the SUP05 clade that comprised lower degrees of host restriction (*Bathymodiolus thermophilus* symbiont (Y.-J. Won 2018, data deposited in GenBank under accession no. NZ_CP024634), *Ca.* Thioglobus autotrophicus [38]) were selected as outgroups (figure 1; electronic supplementary material, tables S1 and S2). Similarities and taxonomic affiliations among genomes were assessed with FastANI [39] and GTDB-Tk [40].

(b) Comparative genomics

Sequence homology between symbiont genomes was inferred through assessment of positional orthology and orthogroup identification with OrthoFinder [41] (electronic supplementary material, tables S3 and S4). ProgressiveMauve [42] and Grimm [43] were used to identify large-scale structural differences among mitochondrial and symbiont genomes based on 13 and 716 conserved protein-coding genes, respectively. Phylogenetic trees were produced from these gene sets in MrBayes [44]. Concordance among tree topologies was assessed with Bucky [45]. Pairwise synonymous substitution rates for the mitochondrial and symbiont core genomes were computed following the method in Goldman & Yang [46].

(c) Selection analyses

Branch-specific episodic diversifying selection was identified based on non-recombining core protein-coding genes using aBSRel [47]. Changes in the strength of selection on core protein-coding genes were inferred through quantifications of codon usage bias [48] and phylogenetic hypothesis testing with Relax [49]. To strengthen the inferences from these analyses, we performed all tests with additional metagenome-assembled genomes representative of the diversity of free-living and horizontally transmitted SUP05 bacteria. Fubar [50] and Meme [51] were used to assess signatures of pervasive and episodic site-specific positive selection in 17 candidate genes that showed marked differences in presence/absence or duplication patterns between the two symbiont clades.

3. Results

(a) Host mitochondrial and symbiont genomes and phylogenies

The genome-wide mitochondrial phylogeny is congruent with host phylogenetic relationships based on multilocus

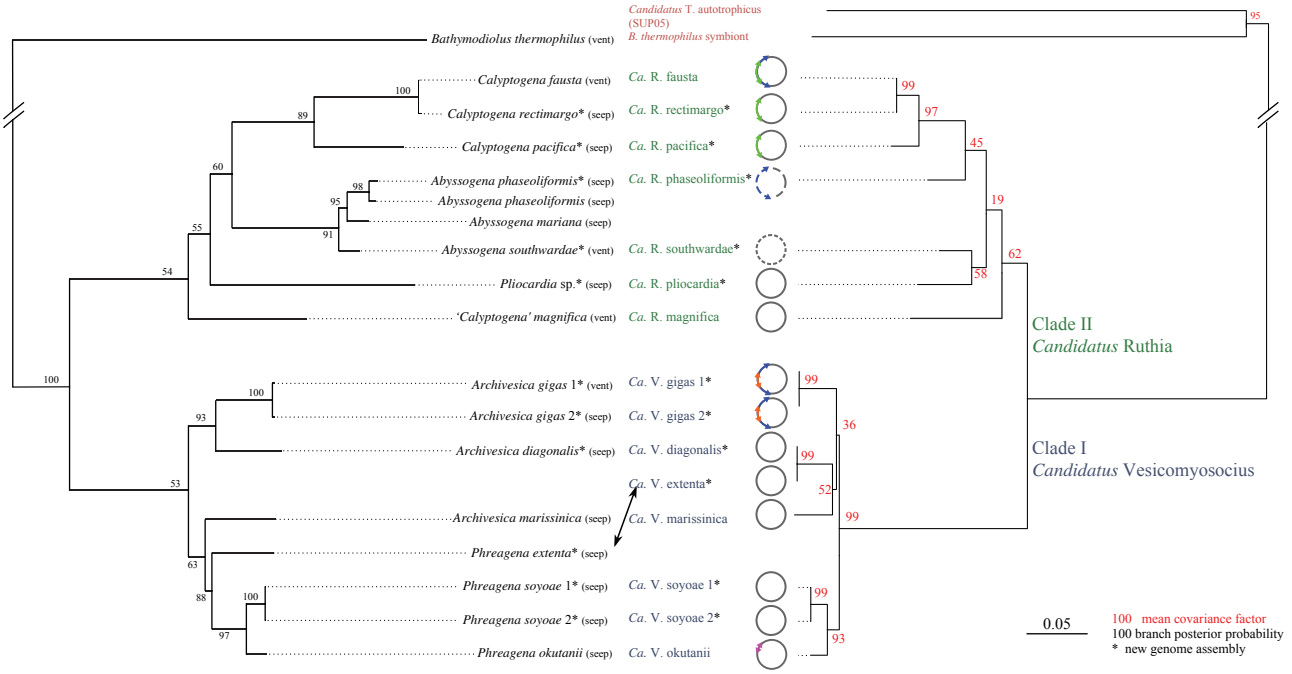

Figure 1. Genome-wide host mitochondrial (left) and symbiont (right) trees. Phylogenies represent the Bayesian majority-rule consensus of 2000 independent trees (GTR + G + I model). Left: consensus tree and branch posterior probabilities from the concatenated alignment of 13 core mitochondrial genes. Right: consensus tree from the concatenated alignment of syntenic blocs shared between symbiont (Clade I: blue; Clade II: green) and outgroup (light red) genomes. Genome inversions and assembly fragmentation are displayed at the end of the branches (blue: inversions between TufA/B paralogues; green, orange and magenta: other inversions). Numbers in red are the genome-wide mean covariance factors, which represent the percentage of non-recombining syntenic blocs supporting each split in the phylogeny. *Genomes newly sequenced in this study. (Online version in colour.)

and *COI* sequence data [11] (figure 1). Host mitochondrial genomes examined in this study possess identical gene orders and contents, though structural variation is evident among taxa between the $tRNA^{Trp}/tRNA^{His\_2}$ and *ND6* loci and in the *COX2* gene (1005–1452 bp; electronic supplementary material, figure S2).

Intra-host symbiont populations were genetically homogeneous with frequency distributions of genetic variants typical of monoclonal populations (electronic supplementary material, table S5 and figure S3). Genome size, GC content and number of intact protein-coding genes for the 15 vesicomyid symbiont assemblies ranged from 1.02–1.59 Mb, 31–37% and 896–1455, respectively (electronic supplementary material, table S5), with Clade I having consistently lower values for these genomic characteristics than Clade II. Following initial nomenclature, the symbiont lineages are referred to by the previously erected genera for the two groups, *Candidatus* Vesicomyosocius for Clade I, and *Candidatus* Ruthia for Clade II, accompanied by host species name [30–32]. This classification at the genus level is consistent with 16S rRNA similarity (less than 95%) [52], clustering based on average nucleotide identity and alignment fraction [53] (electronic supplementary material, figure S4), taxonomic assignment (electronic supplementary material, table S6) and genetic isolation between the two symbiont clades (see below).

Mitochondrial and symbiont phylogenies show good concordance for all lineages except one (figure 1; electronic supplementary material, figure S5). *Ca.* V. diagonalis and *Ca.* V. extenta are nearly identical based on genome size, GC content and gene composition (electronic supplementary material, figures S4–S6 and table S5), whereas the respective host mitochondrial lineages are divergent. The donor lineage in this recent symbiont replacement appears to be *Archivesica diagonalis*, which co-occurs with *Phreagena extenta* at hydrocarbon seeps in Monterey Canyon. Signatures of elevated substitution rates are evident on the branch leading to Clade I, which is notably longer than the corresponding branch for Clade II, in contrast to patterns in the host phylogeny (figure 1). The symbiont pairs across the Clade I–Clade II bipartition are also significantly more divergent than the others even when controlled for host divergence ($1 < dSmito < 2$; electronic supplementary material, figure S7).

(b) Symbiont genome structure and recombination

The *B. thermophilus* symbiont and *Ca.* T. autotrophicus share about 1 Mbp of their genomes with the clam symbionts, with at least 22 and 3 inversion events being present relative to the *Ca.* R. magnifica reference, respectively (electronic supplementary material, figure S8).

Genome structure also differs among the clam symbionts (figure 1; electronic supplementary material, figure S8), with intra-host structural variation being particularly evident within *Ca.* R. phaseoliformis and *Ca.* R. southwardae. Three distinct inversions compared to the *Ca.* R. magnifica genome were found in the genomes of *Ca.* V. okutanii, *Ca.* V. gigas, and the monophyletic group composed of *Ca.* R. fausta, *Ca.* R. pacifica and *Ca.* R. rectimargo. Inversions between the *tufA* and *tufB* paralogues, hotspots for chromosomal inversions [54], seem to have happened multiple times throughout the symbiont phylogeny (figure 1; electronic supplementary material, figure S8).

Bayesian concordance analysis detected substantial recombination within, but not among symbiont clades (figure 1). Relatively little topological concordance was found in Clade II, with 37 different topologies being necessary to fully represent the diversity of conflicting phylogenetic signals compared to only 11 in Clade I (electronic supplementary material,

figure S9). Within Clade I, conflict arises from the uncertainty of the position of *Ca.* V. gigas and *Ca.* V. marissinica (figure 1). Within Clade II, only the grouping of *Ca.* R. fausta, *Ca.* R. rectimargo and *Ca.* R. pacifica is supported by all gene trees, while the positions of all other species have low support.

(c) Patterns of gene conservation

The free-living bacterial and environmentally acquired symbiont genomes contained many large (greater than 5 kb) contiguous sections that were absent in the clam symbionts. These genomic islands comprised mostly unannotated genes and mobile elements, but also genes related to heavy metal tolerance and anti-viral defense (in the *B. thermophilus* symbiont) as well as motility and nitrogen metabolism (in *Ca.* T. autotrophicus) (electronic supplementary material, table S4).

The clam symbionts possessed essentially a subset of the genes found in the outgroup lineages, with *Ca.* R. southwardae, *Ca.* R. phaseoliformis and *Ca.* R. pliocardia showing the highest degree of gene conservation. Many genes unique to the vesicomyid symbionts appeared to be pseudogenes resulting from the degeneration of ancestral homologues. Patterns of pseudogenization were relatively prevalent and variable in Clade II (electronic supplementary material, figure S6), while homologous regions within genomes of Clade I were usually characterized by large deletions. Among the Clade II symbionts, gene degeneration was most pronounced in *Ca.* R. magnifica, which possessed a conservation pattern closer to that of Clade I (electronic supplementary material, figures S5 and S6).

Both symbiont clades shared a core genome related to chemoautotrophic metabolism, but showed differences in presence/absence, duplication and degeneration patterns for genes related to a diversity of other metabolic processes (electronic supplementary material, results, figure S6 and table S4). For instance, the genomes of Clade I and Clade II symbionts encoded different types of methionine synthase. While Clade I contained genes for the cobalamin-dependent homocysteine methyltransferase *metH* and associated genes for cobalamin (precursor) transport and conversion (*btuM*, *btuR*/*cobA*), Clade II contained the cobalamin-independent version of this enzyme (*metE*) along with its transcriptional activator (*metR*). However, all symbiont lineages lacked pathways for de novo cobalamin biosynthesis. Genomes of both symbiont clades also differed in the presence of operons for dissimilatory (*narGHIJ*: Clade I) and assimilatory (*nasA*: Clade II) nitrate reductases, genes for putative nickel transporters (*hupE*) and nickel-dependent enzymes (*gloA*), as well as genes involved in glyoxylate regeneration (*icl*) and transcriptional repression of certain ribonucleotide reductases (*nrdR*) (only in Clade II). Surprisingly, *nasA* was annotated as pseudogene in almost all Clade II lineages and *Ca.* T. autotrophicus. This is possibly a misclassification as functional expression of *nasA* is observed in deep-sea SUP05 populations [55]. Alternatively, this gene might be in an early stage of pseudogenization as all variants encompassed over 74% of the intact protein length. An operon encoding cysteine dioxygenase type I (*cdo*) and an aspartate aminotransferase superfamily protein, which has homology to cysteine sulfinic acid decarboxylase (*csad*) from *B. azoricus* (GenBank: SEH86284), was exclusively found in Clade I. Unlike their Clade II congeners, the genomes of almost all Clade I symbionts were characterized by a duplication of the sulfide : quinone oxidoreductase type I gene (*sqrI*).

(d) Genome-wide patterns of relaxed and intensified selection

Both symbiont clades showed reduced codon usage bias (figure 2*a*,*b*) and d*N*/d*S* rate-class extremes (figure 2*c*) compared to the outgroup (figure 2*c*), indicating a genome-wide decline in the efficacy of natural selection, i.e. a reduction in selective constraint (electronic supplementary material, tables S7 and S8). Codon usage bias and selection intensity analyses including incomplete genomes from free-living SUP05 bacteria and closely related horizontally transmitted symbionts associated with deep-sea mussels (*Bathymodiolus* sp.) and sponges (*Suberites* sp.) further support the inference of drift-driven RGE in the vesicomyid symbionts (electronic supplementary material, figure S10 and results). Relaxation was comparable to that observed in insect endosymbionts and appeared to be exacerbated in Clade I (figure 2*b*,*c*) [49]. Genes exhibiting intensified and relaxed selection in the clam symbionts represented a multitude of metabolic categories, although some functions were predominantly affected by directional shifts in selection regimes (electronic supplementary material, figure S11). Genes under relaxed selection were mostly involved in protein, amino acid and nucleoside/nucleotide metabolism, cell division and cell cycle, whereas genes under intensified selection were largely associated with respiration and sulfur metabolism.

(e) Patterns of positive selection in core and clade-specific genes

One hundred and fourteen protein-coding core genes exhibited evidence for episodic diversifying selection along branches in the phylogeny (electronic supplementary material, table S9). Selection appears to be distributed throughout the evolutionary history of the symbionts (figure 3), acting mostly on the outgroup branches and the branches discriminating the outgroup, Clade I, and Clade II. Eighty-five per cent of loci that showed signs of selection were classified into SEED categories (figure 3). These loci were equally represented among cellular functions of the core genome except for a few categories (e.g. nucleotide synthesis and defense) along the branches separating the two symbiont clades (figure 3).

Apart from *gloA*, *narI* and *narJ*, all investigated metabolic genes that were differentially preserved between clades showed evidence of pervasive or episodic site-specific diversifying selection that affected structural or functional regions in the encoded proteins (electronic supplementary material, table S10). Pervasive positive selection was observed at one to three sites across the entire phylogeny in ten of the 17 genes tested: *btuM*, *btuR*, *csad*, *hupE*, *icl*, *metR*, *narG*, *narH*, *nasA*, *sqrI*. In addition, episodic positive selection was detected at one to seven sites along a proportion of branches in all tested genes except for *btuR*, *gloA*, *narI* and *narJ*. In the case of *cdo*, *csad* and *nasA*, these episodes of site-specific selection seemed to have mostly occurred in the ancestral lineages as no evidence for selection was found along the extant symbiont branches (electronic supplementary material, table S10).

4. Discussion

(a) Reductive genome evolution is ongoing and driven by neutral processes

Current insights into the evolutionary processes shaping RGE in maternally inherited symbionts stem mostly from

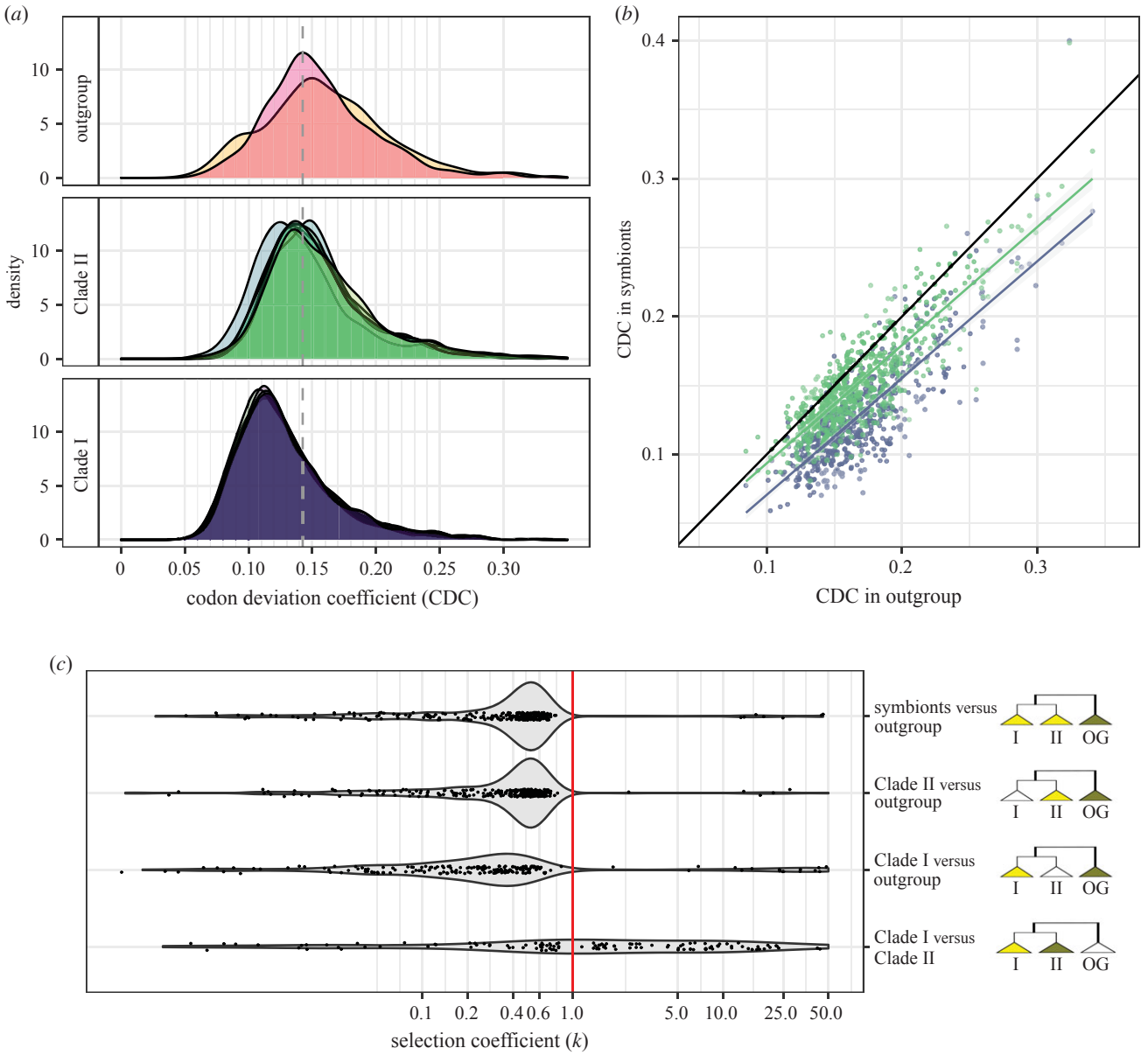

Figure 2. Codon usage bias in clam symbionts and outgroup bacterial relatives. (*a*) Codon deviation coefficient (CDC) spectra for each genome within the outgroup; yellow: *B. thermophilus* symbiont; red: *Ca.* T. autotrophicus. (*b*) Correlation between the average CDC of the outgroup, Clade I (blue) and Clade II (green) based on 555 core genes. CDC values vary from 0 (no bias) to 1 (maximum bias). (*c*) Log-scaled selection parameter (*k*) spectra of core genes for which a significant change in selection was detected. CDC values were significantly lower in Clade I than Clade II, and CDC and *k* values were significantly lower in both symbiont clades than the outgroup (paired Wilcoxon–Mann–Whitney test p-value < 0.01). (Online version in colour.)

well-studied terrestrial insect–bacteria associations, where genetic drift has been shown to be the dominant force driving patterns of endosymbiont gene loss [7,56]. Our analyses of 15 vesicomyid symbiont genomes suggest that neutral processes play an equally important role in marine vertically transmitted symbioses. As in other models of recently acquired bacteria [57,58], gene content differed substantially between vesicomyid symbiont genomes, indicating that the different lineages are independently losing genes. The presence of structural variation and varying degrees of gene degeneration imply that vesicomyid symbionts have not yet reached a stable streamlined state compared to many insect endosymbionts [59], as suggested previously [24]. All clam symbionts exhibited a reduced GC%, decrease in codon usage bias, and a genome-wide trend of relaxation in selective pressures relative to the outgroup. Overall, these observations support the nearly neutral theory of RGE, driven by a reduction of effective population size in these taxa [4].

In agreement with previous findings [19–21,23], we detected no recombination between Clade I and Clade II, even though some of the host taxa co-occur [21,26]. This implies that there is enough molecular and ecological divergence between the two clades for clonal interference and/or strong host–symbiont epistatic interactions to constrain symbiont exchange. Clade I and Clade II are also discriminated based on measures of genomic relatedness and functional genomic traits, all of which support a classification of these symbionts into two distinct bacterial genera, *Ca.* Vesicomyosocius and *Ca.* Ruthia [29,33].

(b) Reductive genome evolution is exacerbated in non-recombining symbionts

Symbionts of Clade I appear to be in a more advanced state of RGE than those of Clade II, as their genomes are smaller and lower in GC%, possess fewer genes and pseudogenes,

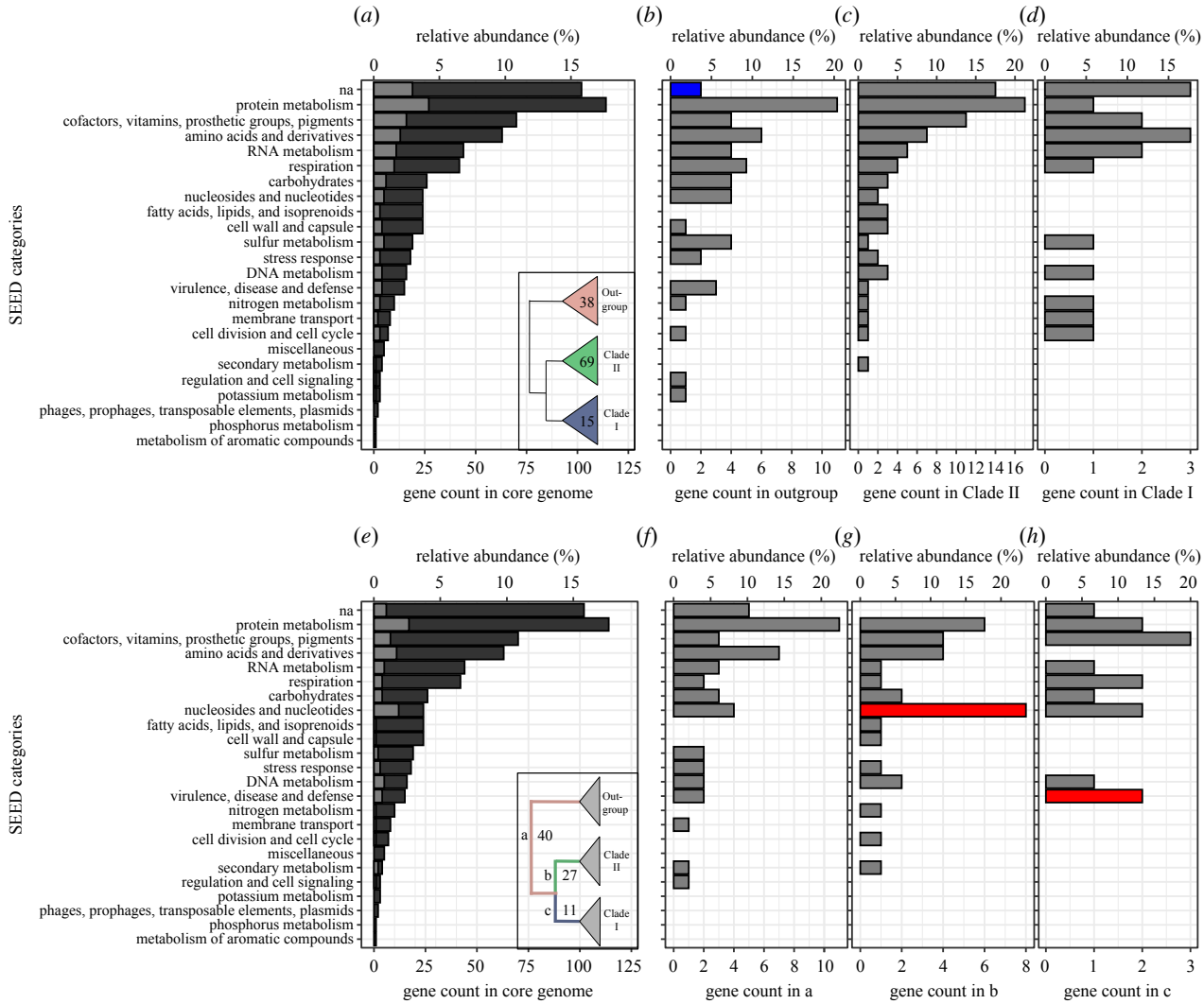

Figure 3. SEED category distribution of core genes under episodic diversifying selection within phylogenetic clades (*a–d*), and on partitioning branches (*e–h*). (*a*) Distribution of all non-recombining core genes (dark grey, 555 loci) and loci under selection within the outgroup, Clade I and Clade II (light grey, 114 loci). (*b*) Genes under selection within the outgroup. (*c*) Genes under selection within Clade II. (*d*) Genes under selection within Clade I. (*e*) Distribution of all non-recombining core genes (dark grey, 555 loci) and loci under selection on all partitioning branches (light grey, 71 loci). (*f*) Genes under selection on branch a. (*g*) Genes under selection on branch b. (*h*) Genes under selection on branch c. Note that genes may be present in multiple functional categories, clades and/or branches. Insets show number of loci selected within clades or along branches. SEED categories significantly overrepresented (in red) and underrepresented (in blue) compared to the core genome are highlighted. NA, no functional annotation. (Online version in colour.)

exhibit less codon usage bias and are in general more homogeneous. Patterns of gene conservation suggest that much of the loss in this group happened after its speciation but before its radiation, a period of roughly 20 Myr [10,11]. Fossil informed phylogenetic inference places the radiation of vesicomyid clams at 47 ± 2 Ma and that of Clade I at 22 ± 5 Ma [11], resulting in a rate of genome reduction of about 20 Kbp/Myr in the Clade I symbionts. Together with increased substitution rates on its diverging branch, these results imply that the ancestral Clade I lineage experienced an acute episodic acceleration of RGE. Based on genome-wide levels of topological disagreement, interspecific homologous recombination is widespread among symbionts of Clade II but almost absent in Clade I. A reduction of the rate of infection by environmental symbionts and/or drift-driven loss of the recombination machinery [18] may have strongly reduced the rate of genetic exchange across Clade I symbionts, thereby setting this genus on a divergent evolutionary path.

Recombination can alter rates of evolution due to Hill–Robertson interference [60] by randomizing the associations between mutations that otherwise would be in linkage disequilibrium. In small populations, deleterious alleles fix through drift, reducing the mean fitness of the population [6]. Additionally, background selection against deleterious alleles can purge linked beneficial alleles from the population [61], whereas hitchhiking effects can retain linked deleterious mutations [62]. Low rates of recombination typically increase Hill–Robertson effects [60], whereas absence of recombination can reduce the rate of adaptation through clonal interference [63].

Strong linkage disequilibrium forces whole genomes to sweep in populations that lack capabilities for genetic exchange. Hence, loss of recombination should favour symbiont replacement in cases where the divergence between native and foreign symbionts is low enough to avoid host–symbiont incompatibilities. We find multiple examples of symbiont replacement among lineages of Clade I. For instance, individual *P. extenta* clams have acquired the symbionts of the sympatric species *A. diagonalis*. Likewise, some *A. gigas* populations have been found to carry the symbionts of the host species *P. soyoae* [19]. Symbiont replacement

occurs in several vertically transmitted symbioses [64] and is speculated to constitute a mechanism for escaping the evolutionary rabbit hole caused by Muller's ratchet [2].

Despite the lack of recombination machinery in Clade I, *Ca.* V. gigas and *Ca.* V. marissinica showed signs of genetic exchange. Perhaps recombination in these species is mediated via symbiont-derived host-encoded proteins. Evidence for transfer of ancestral symbiont gene homologues to the host nuclear genome was recently found in *A. marissinica* clams [31]. Overall, these observations support a crucial role of recombination in maintaining symbiont genome integrity [5] and moderating the ecological consequences of increased clonality.

(c) Selective processes might be tied to genetic and environmental contexts

Although our data indicate that genetic drift is a major force mediating RGE in vesicomyid symbionts, significant fractions of the symbiont genomes are affected by natural selection. Selection might act on genes involved in host–symbiont interactions, as these genes are expected to experience reciprocal adaptations through speciation and niche exploitation. Diversifying selection affecting genes that play a role in host–symbiont interactions, such as lipopolysaccharides and peptidoglycans, is observed in several terrestrial obligate and facultative endosymbionts [65,66]. Surprisingly, our data do not confirm these predictions and instead suggest a pervasive pattern of positive selection affecting a broad range of cellular functions. This could indicate that the accumulation of slightly deleterious mutations in the symbiont genomes enhances selective pressures for compensatory mutations, as described for cellular organelles and other bacterial endosymbionts [67,68]. Alternatively, these patterns might reflect ongoing adaptations to the host intracellular environment [56].

Strong functional contrasts in gene loss between outgroup and clam symbionts as well as between both symbiont clades further suggest a role of niche differentiation in shaping symbiont genome composition through RGE, which has consequences for ecological processes, like habitat use, and evolutionary processes, like host speciation. Genes enabling bacteria to face the challenges of a free-living environment, such as metal detoxification, anti-viral defense and inter-species competition, were not conserved in the clam symbionts, while both clades differed in retention of genes affecting diverse metabolic traits, including the dependency on enzyme cofactors, anaerobic respiration and sulfur physiology.

(d) Symbiont clades show putative differences in physiology and ecological niche

Clade I and Clade II symbionts encoded different, convergently evolved types of methionine synthase [69], which vary in their requirement for vitamin B12. Measurements of cobalamin in the deep-sea are challenging, and we, therefore, do not currently have informative data on vitamin B12 concentrations experienced by the clams. However, as both clades appeared unable to synthesize cobalamin de novo, these findings indicate that the environmental availability of vitamin B12 has the potential to be an important factor influencing the distribution of these taxa. Cobalamin independence in Clade II may offer a selective advantage by allowing these symbioses to exploit habitats that would otherwise be inaccessible. By contrast, the requirement for exogeneous vitamin B12 (or its derivatives) might limit the range of (micro)habitats Clade I-based associations can colonize, unless cobalamin is acquired from a secondary symbiont. Despite this potential cost, the retention of a cobalamin-dependent methionine synthase in Clade I probably provides an evolutionary benefit, given that MetH has a 50-fold higher catalytic rate constant than MetE and thus enables faster growth [69]. Comparative measurements of vesicomyid growth rates suggest that species hosting Clade I symbionts typically grow faster than species with symbionts of Clade II, despite a less efficient sulfur physiology [70]. Since growth is influenced by a variety of factors [70], it is possible that the enzymatic differences in methionine biosynthesis among symbiont clades contribute to an accelerated anabolism in Clade I-based associations, although this remains to be experimentally tested. The preservation of cobalamin-dependent enzymes as a result of conferred physiological advantages appears to be common across the eubacterial domain [71]. About 86% of bacterial lineages seem to have at least one cobalamin-dependent enzyme despite the existence of a cobalamin-independent alternative, and many of these lineages rely on vitamin B12 production from other microbes in their environment [71]. The importance of vitamin B12 for the biology of the two symbiont groups is also evident in the fact that only Clade II symbionts encode a transcriptional repressor (NrdR) for the ribonucleotide reductase NrdAB, a key enzyme that controls the synthesis of DNA [72]. In Clade I, expression of NrdAB is probably regulated by cobalamin, which has been shown to repress NrdAB transcription through riboswitches [73]. There is evidence that the two symbiont clades differ in their requirements for other enzyme cofactors, such as nickel. The genomes of Clade II symbionts encoded a specific transporter for nickel uptake, and most of these lineages contained at least one confirmed nickel-dependent enzyme (glyoxalase I [74]), all of which were absent in Clade I.

Our data extend previous findings that Clade I and Clade II symbionts show differences in encoded gene clusters for nitrate reduction [14], confirming that these patterns are truly clade-specific characteristics. The use of NarGHIJ for nitrate reduction in Clade I probably enables these symbioses to inhabit hypoxic environments [14], since the use of nitrate as an electron acceptor would reduce the symbiont's requirement for oxygen and, consequently, avoid competition with the host. These assumptions agree with field observations showing that clam species hosting Clade I symbionts typically occupy microhabitats with higher levels of hydrogen sulfide (and, thus, presumably lower oxygen) than those hosting Clade II symbionts [26]. Niche partitioning based on environmental sulfide levels has also been suggested by physiological comparisons of *P. soyoae* and *C. pacifica*, which imply that *P. soyoae* symbionts require higher H_2S concentrations for chemosynthesis [26]. This could be due to a less efficient sulfide metabolism in *Ca.* V. soyoae (Clade I) resulting from an increased load of deleterious mutations in accordance with its more advance state of genome reduction compared to *Ca.* R. pacifica (Clade II). The presence of two tandem copies of *sqrI* displaying evidence of concerted evolution suggests increasing gene dosage as compensating mechanism in Clade I. Our data indicate that there might be additional adaptations (or ancestral restrictions) to contrasting sulfide environments between symbiont clades. Only Clade I symbionts encode genes for CDO and

putatively CSAD, which are key enzymes in the biosynthesis of taurine/hypotaurine. These non-proteinogenic amino acids are important for sulfide detoxification in many symbiont-bearing invertebrates that inhabit sulfidic environments [75]. Since Clade I-based associations appear to be prevalent in high-sulfide habitats, it is possible that these symbionts directly or indirectly contribute to H_2S tolerance of their hosts. Another function of CDO and CSAD could involve replenishment of metabolic intermediates. Vesicomyid symbionts do not possess a complete TCA cycle and must recycle succinate through other means [14]. In Clade II symbionts succinate regeneration occurs via the glyoxylate shunt, while the mechanism in Clade I symbionts is unclear [14]. Perhaps taurine is further metabolized via taurine dioxygenase, which would generate succinate as end product. If this pathway can be confirmed through physiological experiments, autonomous recycling of succinate by the symbiont could make important contributions to the holobiont's carbon budget.

5. Conclusion

Our analyses show that patterns of genome reduction in vesicomyid symbionts are mostly shaped by genetic drift and that factors affecting symbiont clonality strongly influence the rate of RGE. A first period of intensified RGE probably occurred during the shift from horizontal to vertical transmission mode in the Early to Mid-Eocene, and was followed by a second period in Clade I after transition to complete host restriction in the Late Eocene/Early Oligocene. The pervasive nature of episodic diversifying selection across functional traits in the vesicomyid symbiont genomes, however, suggests that neutral evolutionary processes (drift, mutation and recombination) are not the sole drivers of molecular evolution in these taxa. Differential patterns of gene loss between Clade I and Clade II reiterate that RGE does not follow a universal trajectory, but is a reflection of the eco-evolutionary context of the respective host–symbiont association. Convergent gene loss and pseudogenization imply common evolutionary pressures for some genes, whereas selection and lineage-specific gene retention imply niche-specific adaptation in others. Future studies linking environmental data with symbiont genomic information at the population level will help to decipher the contributions of host eco-physiology, symbiont fitness, cytonuclear incompatibilities and rates of lateral transfer to symbiont evolution.

Data accessibility. Symbiont genomes (CP060680–CP060688, JACRUR000000000, JACRUS000000000) and raw reads are available at the National Center for Biotechnology Information under BioProject PRJNA641445. Genome annotations and metabolic reconstructions can be found on the RAST webserver. Host mitochondrial *COI* sequences and genomes have been deposited in GenBank under accession numbers MT894120–MT894130 and MT947381–MT947391, respectively. Genome alignment files and all scripts used in this study are available at https://github.com/maepz/VesicSymb_Evolution.

The data are provided in electronic supplementary material [76].

Authors' contributions. M.P.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing—original draft, writing—review and editing; C.B.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing—original draft, writing—review and editing; B.A.: resources, supervision, writing—original draft, writing—review and editing; R.A.B.: investigation, resources, supervision, writing—original draft, writing—review and editing; Y.-J.W.: data curation, writing—original draft, writing—review and editing; C.R.Y.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by grants of the David and Lucile Packard Foundation (to MBARI), the UK Natural Environment Research Council (grant no. NE/N006496/1 to C.R.Y.), the US National Science Foundation (grant no. OCE-1736932 to R.A.B.), the German Research Foundation (grant no. BR 5488/1-1 to C.B.), the Natural Science and Engineering Research Council of Canada (grant no. 238600 to B.A., Alexander Graham Bell graduate scholarship and Michael Smith Foreign Study Supplements to M.P.) and National Capability funding to the National Oceanography Centre (grant no. NE/R015953/1). Sequencing of the *Bathymodiolus thermophilus* symbiont genome was funded by the Korean Ministry of Oceans and Fisheries (grant no. 20170411 to Y.J.W.). Bioinformatic analyses were in part performed on ComputeCanada HPC clusters.

Acknowledgements. We thank the captains, crews and submersible pilots of the R/Vs *Atlantis*, *Western Flyer* and *Point Lobos* for supporting the sample collections, and Bob Vrijenhoek for providing clam specimens for this study. We further thank N. Pratt and A. Baylay for their contributions to library preparation and sequencing at the National Oceanography Centre Genomics Facility.

References

1. Russell SL. 2019 Transmission mode is associated with environment type and taxa across bacteria-eukaryote symbioses: a systematic review and meta-analysis. *FEMS Microbiol. Lett.* **366**, fnz013. (doi:10.1093/femsle/fnz013)
2. Bennett GM, Moran NA. 2015 Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *Proc. Natl Acad. Sci. USA* **112**, 10 169–10 176. (doi:10.1073/pnas.1421388112)
3. Vrijenhoek RC. 2010 Genetic diversity and connectivity of deep-sea hydrothermal vent metapopulations. *Mol. Ecol.* **19**, 4391–4411. (doi:10.1111/j.1365-294X.2010.04789.x)
4. Moran NA. 1996 Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl Acad. Sci. USA* **93**, 2873–2878. (doi:10.1073/pnas.93.7.2873)
5. Russell SL, Pepper-Tunick E, Svedberg J, Byrne A, Castillo JR, Vollmers C, Beinart RA, Corbett-Detig R. 2020 Horizontal transmission and recombination maintain forever young bacterial symbiont genomes. *PLoS Genet.* **16**, e1008935. (doi:10.1371/journal.pgen.1008935)
6. Muller HJ. 1964 The relation of recombination to mutational advance. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **1**, 2–9. (doi:10.1016/0027-5107(64)90047-8)
7. Wernegreen JJ. 2011 Reduced selective constraint in endosymbionts: elevation in radical amino acid replacements occurs genome-wide. *PLoS ONE* **6**, e28905. (doi:10.1371/journal.pone.0028905)
8. Veller C, Hayward LK, Hilbe C, Nowak MA. 2017 The Red Queen and King in finite populations. *Proc. Natl Acad. Sci. USA* **114**, E5396–E5405. (doi:10.1073/pnas.1702020114)
9. Hansen AK, Moran NA. 2014 The impact of microbial symbionts on host plant utilization by herbivorous insects. *Mol. Ecol.* **23**, 1473–1496. (doi:10.1111/mec.12421)
10. Peek AS, Feldman RA, Lutz RA, Vrijenhoek RC. 1998 Cospeciation of chemoautotrophic bacteria and deep sea clams. *Proc. Natl Acad. Sci. USA* **95**, 9962–9966. (doi:10.1073/pnas.95.17.9962)

11. Johnson SB, Krylova EM, Audzijonyte A, Sahling H, Vrijenhoek RC. 2017 Phylogeny and origins of chemosynthetic vesicomyid clams. *Syst. Biodivers.* **15**, 346–360. (doi:10.1080/14772000.2016.1252438)
12. Audzijonyte A, Krylova EM, Sahling H, Vrijenhoek RC. 2012 Molecular taxonomy reveals broad trans-oceanic distributions and high species diversity of deep-sea clams (Bivalvia: Vesicomyidae: Pliocardiinae) in chemosynthetic environments. *Syst. Biodivers.* **10**, 403–415. (doi:10.1080/14772000.2012.744112)
13. Childress JJ, Fisher CR, Favuzzi JA, Sanders NK. 1991 Sulfide and carbon dioxide uptake by the hydrothermal vent clam, *Calyptogena magnifica*, and its chemoautotrophic symbionts. *Physiol. Zool.* **64**, 1444–1470. (doi:10.1086/physzool.64.6.30158224)
14. Newton IL, Girguis PR, Cavanaugh CM. 2008 Comparative genomics of vesicomyid clam (Bivalvia: Mollusca) chemosynthetic symbionts. *BMC Genom.* **9**, 585. (doi:10.1186/1471-2164-9-585)
15. Peek AS, Gustafson RG, Lutz RA, Vrijenhoek RC. 1997 Evolutionary relationships of deep-sea hydrothermal vent and cold-water seep clams (Bivalvia: Vesicomyidae): results from the mitochondrial cytochrome oxidase subunit I. *Mar. Biol.* **130**, 151–161. (doi:10.1007/s002270050234)
16. Moran NA, Munson MA, Baumann P, Ishikawa H. 1993 A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc. Biol. Sci.* **253**, 167–171. (doi:10.1098/rspb.1993.0098)
17. Ferri E *et al.* 2011 New Insights into the evolution of *Wolbachia* infections in filarial nematodes inferred from a large range of screened species. *PLoS ONE* **6**, e20843. (doi:10.1371/journal.pone.0020843)
18. Kuwahara H, Takaki Y, Shimamura S, Yoshida T, Maeda T, Kunieda T, Maruyama T. 2011 Loss of genes for DNA recombination and repair in the reductive genome evolution of thioautotrophic symbionts of *Calyptogena* clams. *BMC Evol. Biol.* **11**, 285. (doi:10.1186/1471-2148-11-285)
19. Stewart FJ, Young CR, Cavanaugh CM. 2008 Lateral symbiont acquisition in a maternally transmitted chemosynthetic clam endosymbiosis. *Mol. Biol. Evol.* **25**, 673–687. (doi:10.1093/molbev/msn010)
20. Stewart FJ, Young CR, Cavanaugh CM. 2009 Evidence for homologous recombination in intracellular chemosynthetic clam symbionts. *Mol. Biol. Evol.* **26**, 1391–1404. (doi:10.1093/molbev/msp049)
21. Decker C, Olu K, Arnaud-Haond S, Duperron S. 2013 Physical proximity may promote lateral acquisition of bacterial symbionts in vesicomyid clams. *PLoS ONE* **8**, e64830. (doi:10.1371/journal.pone.0064830)
22. Breusing C, Johnson SB, Vrijenhoek RC, Young CR. 2019 Host hybridization as a potential mechanism of lateral symbiont transfer in deep-sea vesicomyid clams. *Mol. Ecol.* **28**, 4697–4708. (doi:10.1111/mec.15224)
23. Ozawa G, Shimamura S, Takaki Y, Takishita K, Ikuta T, Barry JP, Maruyama T, Fujikura K, Yoshida T. 2017 Ancient occasional host switching of maternally transmitted bacterial symbionts of chemosynthetic vesicomyid clams. *Genome Biol. Evol.* **9**, 2226–2236. (doi:10.1093/gbe/evx166)
24. Kuwahara H, Takaki Y, Yoshida T, Shimamura S, Takishita K, Reimer JD, Kato C, Maruyama T. 2008 Reductive genome evolution in chemoautotrophic intracellular symbionts of deep-sea *Calyptogena* clams. *Extremophiles* **12**, 365–374. (doi:10.1007/s00792-008-0141-2)
25. Shimamura S *et al.* 2017 Loss of genes related to Nucleotide Excision Repair (NER) and implications for reductive genome evolution in symbionts of deep-sea vesicomyid clams. *PLoS ONE* **12**, e0171274. (doi:10.1371/journal.pone.0171274)
26. Goffredi SK, Barry JP. 2002 Species-specific variation in sulfide physiology between closely related vesicomyid clams. *Mar. Ecol. Prog. Ser.* **225**, 227–238. (doi:10.3354/meps225227)
27. Cruaud P *et al.* 2019 Ecophysiological differences between vesicomyid species and metabolic capabilities of their symbionts influence distribution patterns of the deep-sea clams. *Mar. Ecol.* **40**, e12541. (doi:10.1111/maec.12541)
28. Luo H, Huang Y, Stepanauskas R, Tang J. 2017 Excess of non-conservative amino acid changes in marine bacterioplankton lineages with reduced genomes. *Nat. Microbiol.* **2**, 1–9. (doi:10.1038/nmicrobiol.2017.91)
29. Baumgartner M, Roffler S, Wicker T, Pernthaler J. 2017 Letting go: bacterial genome reduction solves the dilemma of adapting to predation mortality in a substrate-restricted environment. *ISME J.* **11**, 2258–2266. (doi:10.1038/ismej.2017.87)
30. Newton ILG *et al.* 2007 The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science* **315**, 998–1000. (doi:10.1126/science.1138438)
31. Ip JC-H *et al.* 2020 Host-endosymbiont genome integration in a deep-sea chemosymbiotic clam. *Mol. Biol. Evol.* **38**, 502–518. (doi:10.1093/molbev/msaa241)
32. Kuwahara H *et al.* 2007 Reduced genome of the thioautotrophic intracellular symbiont in a deep-sea clam, *Calyptogena okutanii*. *Curr. Biol.* **17**, 881–886. (doi:10.1016/j.cub.2007.04.039)
33. Ozawa G, Shimamura S, Takaki Y, Yokobori S-I, Ohara Y, Takishita K, Maruyama T, Fujikura K, Yoshida T. 2017 Updated mitochondrial phylogeny of pteriomorph and heterodont bivalvia, including deep-sea chemosymbiotic *Bathymodiolus* mussels, vesicomyid clams and the thyasirid clam *Conchocele* cf. *bisecta*. *Mar. Genom.* **31**, 43–52. (doi:10.1016/j.margen.2016.09.003)
34. Yang M, Gong L, Sui J, Li X. 2019 The complete mitochondrial genome of *Calyptogena marissinica* (Heterodonta: Veneroida: Vesicomyidae): insight into the deep-sea adaptive evolution of vesicomyids. *PLoS ONE* **14**, e0217952. (doi:10.1371/journal.pone.0217952)
35. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017 GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **45**, W6–W11. (doi:10.1093/nar/gkx391)
36. Liu H, Cai S, Zhang H, Vrijenhoek RC. 2016 Complete mitochondrial genome of hydrothermal vent clam *Calyptogena magnifica*. *Mitochondrial DNA A DNA Mapp. Seq. Anal.* **27**, 4333–4335. (doi:10.3109/19401736.2015.1089488)
37. Lee Y, Kwak H, Shin J, Kim S-C, Kim T, Park J-K. 2019 A mitochondrial genome phylogeny of Mytilidae (Bivalvia: Mytilida). *Mol. Phylogenet. Evol.* **139**, 106533. (doi:10.1016/j.ympev.2019.106533)
38. Shah V, Morris RM. 2015 Genome sequence of '*Candidatus* Thioglobus autotrophica' strain EF1, a chemoautotroph from the SUP05 clade of marine Gammaproteobacteria. *Genome Announc.* **3**, e01156–15. (doi:10.1128/genomeA.01156-15)
39. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018 High throughput ANI analysis of 90 K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114. (doi:10.1038/s41467-018-07641-9)
40. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019 GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 925–1927. (doi:10.1093/bioinformatics/btz848)
41. Emms DM, Kelly S. 2019 OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238. (doi:10.1186/s13059-019-1832-y)
42. Darling AE, Mau B, Perna NT. 2010 progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **5**, e11147. (doi:10.1371/journal.pone.0011147)
43. Tesler G. 2002 GRIMM: genome rearrangements web server. *Bioinformatics* **18**, 492–493. (doi:10.1093/bioinformatics/18.3.492)
44. Ronquist F *et al.* 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542. (doi:10.1093/sysbio/sys029)
45. Larget BR, Kotha SK, Dewey CN, Ané C. 2010 BUCKy: gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics* **26**, 2910–2911. (doi:10.1093/bioinformatics/btq539)
46. Goldman N, Yang Z. 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736. (doi:10.1093/oxfordjournals.molbev.a040153)
47. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015 Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353. (doi:10.1093/molbev/msv022)
48. Zhang Z, Li J, Cui P, Ding F, Li A, Townsend JP, Yu J. 2012 Codon Deviation Coefficient: a novel measure for estimating codon usage bias and its statistical significance. *BMC Bioinf.* **13**, 43. (doi:10.1186/1471-2105-13-43)
49. Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015 RELAX: detecting relaxed

selection in a phylogenetic framework. *Mol. Biol. Evol.* **32**, 820–832. (doi:10.1093/molbev/msu400)

50. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013 FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* **30**, 1196–1205. (doi:10.1093/molbev/mst030)
51. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. 2012 Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764. (doi:10.1371/journal.pgen.1002764)
52. Stackebrandt E, Goebel BM. 1994 Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* **44**, 846–849. (doi:10.1099/00207713-44-4-846)
53. Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH, Emerson D. 2020 A genus definition for bacteria and archaea based on a standard genome relatedness index. *mBio* **11**, e02475-19. (doi:10.1128/mBio.02475-19)
54. Hughes D. 2000 Co-evolution of the tuf genes links gene conversion with the generation of chromosomal inversions. *J. Mol. Biol.* **297**, 355–364. (doi:10.1006/jmbi.2000.3587)
55. Anantharaman K, Breier JA, Sheik CS, Dick GJ. 2013 Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. *Proc. Natl Acad. Sci. USA* **110**, 330–335. (doi:10.1073/pnas.1215340110)
56. Martínez-Cano DJ, Reyes-Prieto M, Martínez-Romero E, Partida-Martínez LP, Latorre A, Moya A, Delaye L. 2015 Evolution of small prokaryotic genomes. *Front. Microbiol.* **5**, 742. (doi:10.3389/fmicb.2014.00742)
57. Andersson JO, Andersson SG. 1999 Genome degradation is an ongoing process in *Rickettsia*. *Mol. Biol. Evol.* **16**, 1178–1191. (doi:10.1093/oxfordjournals.molbev.a026208)
58. Burke GR, Moran NA. 2011 Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol. Evol.* **3**, 195–208. (doi:10.1093/gbe/evr002)
59. Tamas I, Klasson L, Canbäck B, Näslund AK, Eriksson A-S, Wernegreen JJ, Sandström JP, Moran NA, Andersson SGE. 2002 50 million years of genomic stasis in endosymbiotic bacteria. *Science* **296**, 2376–2379. (doi:10.1126/science.1071278)
60. Hill WG, Robertson A. 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294. (doi:10.1017/S0016672300010156)
61. Charlesworth B. 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet. Res.* **63**, 213–227. (doi:10.1017/S0016672300032365)
62. Gillespie JH. 2000 Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* **155**, 909–919. (doi:10.1093/genetics/155.2.909)
63. Gerrish PJ, Lenski RE. 1998 The fate of competing beneficial mutations in an asexual population. *Genetica* **102**, 127. (doi:10.1023/A:1017067816551)
64. Sudakaran S, Kost C, Kaltenpoth M. 2017 Symbiont acquisition and replacement as a source of ecological innovation. *Trends Microbiol.* **25**, 375–390. (doi:10.1016/j.tim.2017.02.014)
65. Brownlie JC, Adamski M, Slatko B, McGraw EA. 2007 Diversifying selection and host adaptation in two endosymbiont genomes. *BMC Evol. Biol.* **7**, 68. (doi:10.1186/1471-2148-7-68)
66. Dale C, Moran NA. 2006 Molecular interactions between bacterial symbionts and their hosts. *Cell* **126**, 453–465. (doi:10.1016/j.cell.2006.07.014)
67. Howe DK, Denver DR. 2008 Muller's Ratchet and compensatory mutation in *Caenorhabditis briggsae* mitochondrial genome evolution. *BMC Evol. Biol.* **8**, 62. (doi:10.1186/1471-2148-8-62)
68. Lambert JD, Moran NA. 1998 Deleterious mutations destabilize ribosomal RNA in endosymbiotic bacteria. *Proc. Natl Acad. Sci. USA* **95**, 4458–4462. (doi:10.1073/pnas.95.8.4458)
69. González JC, Banerjee RV, Huang S, Sumner JS, Matthews RG. 1992 Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: two solutions to the same chemical problem. *Biochemistry* **31**, 6045–6056. (doi:10.1021/bi00141a013)
70. Barry JP, Kochevar RE. 1998 A tale of two clams: differing chemosynthetic life styles among vesicomyids in Monterey Bay cold seeps. *Cah. Biol. Mar.* **39**, 329–331.
71. Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR, Taga ME. 2019 Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J.* **13**, 789–804. (doi:10.1038/s41396-018-0304-9)
72. Torrents E. 2014 Ribonucleotide reductases: essential enzymes for bacterial life. *Front. Cell Infect. Microbiol.* **4**, 52. (doi:10.3389/fcimb.2014.00052)
73. Borovok I, Gorovitz B, Schreiber R, Aharonowitz Y, Cohen G. 2006 Coenzyme B12 controls transcription of the *Streptomyces* class Ia ribonucleotide reductase nrdABS operon via a riboswitch mechanism. *J. Bacteriol.* **188**, 2512–2520. (doi:10.1128/JB.188.7.2512-2520.2006)
74. Boer JL, Mulrooney SB, Hausinger RP. 2014 Nickel-dependent metalloenzymes. *Arch. Biochem. Biophys.* **544**, 142–152. (doi:10.1016/j.abb.2013.09.002)
75. Brand GL, Horak RV, Bris NL, Goffredi SK, Carney SL, Govenar B, Yancey PH. 2007 Hypotaurine and thiotaurine as indicators of sulfide exposure in bivalves and vestimentiferans from hydrothermal vents and cold seeps. *Mar. Ecol.* **28**, 208–218. (doi:10.1111/j.1439-0485.2006.00113.x)
76. Perez M, Breusing C, Angers B, Beinart RA, Won YJ, Young CR. 2022 Divergent paths in the evolutionary history of maternally-transmitted clam symbionts. Evolution and ecology of maternally-transmitted symbiont genomes. FigShare.