Royal Society Publishing
Informing the science of the future

Adaptive radiation of chemosymbiotic deep-sea mussels

Julien Lorion[1], Steffen Kiel[2], Baptiste Faure[3], Masaru Kawato[1], Simon Y. W. Ho[4], Bruce Marshall[5], Shinji Tsuchida[1], Jun-Ichi Miyazaki[6] and Yoshihiro Fujiwara[1]

[1]Marine Ecosystems Research Department, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima, Yokosuka 237-0061, Japan
[2]Geoscience Center, Geobiology Group, University of Göttingen, Goldschmidtstrasse 3, Göttingen 37077, Germany
[3]Department of Biology, Pennsylvania State University, University Park, PA 16802, USA
[4]School of Biological Sciences, University of Sydney, Sydney, New South Wales 2006, Australia
[5]Museum of New Zealand Te Papa Tongarewa, PO Box 467, 169 Tory St., Te Aro, Wellington 6011, New Zealand
[6]Faculty of Education and Human Sciences, University of Yamanashi, Kofu, Yamanashi 400-8510, Japan

Adaptive radiations present fascinating opportunities for studying the evolutionary process. Most cases come from isolated lakes or islands, where unoccupied ecological space is filled through novel adaptations. Here, we describe an unusual example of an adaptive radiation: symbiotic mussels that colonized island-like chemosynthetic environments such as hydrothermal vents, cold seeps and sunken organic substrates on the vast deep-sea floor. Our time-calibrated molecular phylogeny suggests that the group originated and acquired sulfur-oxidizing symbionts in the Late Cretaceous, possibly while inhabiting organic substrates and long before its major radiation in the Middle Eocene to Early Oligocene. The first appearance of intracellular and methanotrophic symbionts was detected only after this major radiation. Thus, contrary to expectations, the major radiation may have not been triggered by the evolution of novel types of symbioses. We hypothesize that environmental factors, such as increased habitat availability and/or increased dispersal capabilities, sparked the radiation. Intracellular and methanotrophic symbionts were acquired in several independent lineages and marked the onset of a second wave of diversification at vents and seeps. Changes in habitat type resulted in adaptive trends in shell lengths (related to the availability of space and energy, and physiological trade-offs) and in the successive colonization of greater water depths.

1. Introduction

Adaptive radiation is broadly defined as the rapid diversification of species and of their adaptations to the environment in response to natural selection and ecological opportunities [1]. The radiations of Darwin's finches and African Rift Lake cichlids are well known and have become popular beyond academic research, but these examples may just be the 'tip of an evolutionary iceberg', because adaptive radiations can take place on a broad range of time scales and taxonomic levels, and may even include the Cambrian explosion of life [2]. The fauna living at hydrothermal vents and cold seeps in the deep sea represents another remarkable, yet poorly understood case of radiation. These habitats are characterized by extreme physico-chemical features and a scarcity of primary photosynthetic production, and the animals living there thrive because of their symbiotic relationships with bacteria able to use sulfide and/or methane as energy sources [3]. Initially considered as archaic faunas that survived major extinction events [4], molecular age estimates and fossil records suggested that most of the modern vent and seep animals appeared during a short time interval between the Late Mesozoic and the Early Cenozoic [5,6].

Mussels of the bivalve family Mytilidae are particularly suitable as model organisms to study the roles of ecological opportunities, symbioses and other adaptations in the evolution of deep-sea chemosymbiotic faunas. Indeed, they dominate many vents and seeps, but are also common in other sulfide-rich habitats such as whale carcasses and sunken wood [7], which are considered to be evolutionary stepping stones to deep-sea vents [8]. In addition to this diversity of habitats, deep-sea mussels have a remarkable range of symbiotic types, including intracellular and extracellular symbionts, and the ability to host either or both sulfur-oxidizing and methanotrophic symbionts [7].

We investigated the evolutionary history of deep-sea symbiotic mussels by estimating phylogenetic relationships from a comprehensive dataset including all known lineages and 80% of the known species, collected from virtually all ocean basins. In an analysis of five gene fragments, including mitochondrial and nuclear DNA, calibrated with three reliable fossils, we (i) characterized speciation rates through time; (ii) reconstructed the evolution of habitat use (environmental type and depth), body size, symbiont type (sulfur and/or methane oxidizer) and the degree of the physiological integration of the symbiont with the host (intracellular versus extracellular symbioses); and (iii) evaluated the impact of these biological and ecological factors on speciation rates, and their respective roles in the evolution of deep-sea symbiotic mussels.

2. Material and methods

(a) Sampling

Fourteen species of vent and seep mussels from the western Pacific were sampled using the manned submersible Shinkai 6500 and remotely operated Hyper Dolphin vehicle (see electronic supplementary material, table S1). From experimental bone deployments in Japanese waters, we collected the three undescribed species *Idas* sp SAL4, *Idas* evolutionary significant unit (ESU) D and *Idas* ESU R [9]. Upon recovery, pieces of the gills of *Bathymodiolus aduloides* and the three undescribed species were fixed with 2.5% glutaraldehyde in filtered seawater for 24 h and preserved in filtered seawater with 10 mM sodium azide at 4°C. Remaining tissues and other specimens were fixed in 99% ethanol for DNA analysis. We also analysed alcohol-fixed foot tissues of 15 species associated with sunken organic substrates collected in the western Pacific during cruises of the Tropical Deep-sea Benthos program [9,10]. Finally, we included alcohol-fixed specimens from the Atlantic Ocean and Mediterranean Sea, including two species from organic falls, six species from hydrothermal vents, five species from cold seeps and the subtidal outgroup species *Modiolus modiolus* (Linnaeus, 1758).

(b) Molecular analyses

Template DNA from feet and gills were extracted to analyse host and symbiont genes, respectively, using the QIAamp DNA Micro Kit (Qiagen, USA). A fragment of the small subunit 16S rRNA was amplified to characterize symbionts in 10 species that had not been studied previously (see electronic supplementary material, table S1). Fragments of mitochondrial COI, NADH4 and 16S, and nuclear 28S and histone 3, were amplified from host species for phylogenetic analysis. Polymerase chain reactions (PCRs) were performed using the Ex Taq PCR kit (TaKaRa, Japan). Forward and reverse primers (0.2 μM each; see electronic supplementary material, table S2) and less than 1 μg of DNA template were added to reaction mixtures. PCR products were generated by an

initial denaturing step of 4 min at 94°C followed by 35 cycles at 94°C for 1 min, 55°C (50°C for COI and NADH4) for 1 min and 1 min at 72°C, and by a final extension at 72°C for 7 min. They were purified using Wizard SV Gels and the PCR Clean-Up System (Promega, Madison, WI). PCR products of symbiont 16S were cloned into the pCR-TOPO vectors using a TOPO TA cloning kit (Invitrogen, USA). The DNA constructs were transferred into *Escherichia coli* TOP10 cells (Invitrogen). The sequencing reaction of bacterial 16S rRNA gene clones (20–50 clones per specimen and species) and amplified eukaryotic COI, NADH4, 16S, 28S and histone 3 genes was performed using a BIGDYE TERMINATOR v. 3.1 Cycle Sequencing Kit (Applied Biosystems, USA). Sequencing was performed using an ABI PRISM v. 3730 Genetic Analyzer (Applied Biosystems). Sequences were proofread using CODONCODE ALIGNER v. 3.7.1.1 (CodonCode Corporation, www.codoncode.com). Locations of symbionts in gill tissues of *Ba. aduloides*, *Idas* nsp SAL4 and *Idas* nsp ESU R were determined using transmission electron microscopy (see electronic supplementary material, figure S1), following protocols described elsewhere [11]. Symbionts were identified by comparing their 16S rRNA sequences with GenBank databases using BLAST searches.

### (c) Model selection and reconstruction of the host tree

DNA sequences from newly obtained hosts were complemented with data from GenBank and aligned using PROBALIGN v. 1.4 [12]. The best-fitting model of nucleotide substitution was selected for each gene and each partition within each gene using the corrected Akaike information criterion (AICc) in JMODELTEST v. 0.1.1 [13]. We inferred gene trees using the Bayesian method implemented in BEAST v. 1.7.2 [14] for each partitioning scheme. A Yule speciation model was used as a tree prior. We modelled rate variation among lineages using an uncorrelated lognormal relaxed clock, with the mean substitution rate fixed to 1 to get branch lengths in units of substitution per site [15]. Posterior distributions were estimated using Markov chain Monte Carlo (MCMC) sampling. Samples were drawn every 1000 steps over a total of 10 million MCMC steps. Each analysis was run four times, with mixing and convergence assessed using TRACER v. 1.5. After discarding 10% of the samples as burn-in, samples from the four runs were thinned (sampling every 4000 steps) and pooled together. The best partitioning scheme was selected for each gene by comparing marginal likelihoods using approximate Bayes factors [16] in TRACER.

We then performed a combined analysis of all five genes, using the selected partitioning scheme and substitution models. Each gene was assigned an independent relaxed clock. The protocol of the combined analysis was the same as that used for single-gene analyses, except the number of MCMC steps was increased to 20 million. The maximum-clade-credibility tree was drawn from the pooled samples. Maximum-likelihood (ML) trees were inferred using TREEFINDER [17], with the same partitioning scheme and substitution models as those used in the Bayesian analyses. Bootstrapping analyses (1000 replicates) were used to evaluate support for the ML tree.

(d) Molecular dating

Three fossils were used as calibrations in the Bayesian relaxed-clock analysis and were implemented as prior distributions for ages of nodes in the tree [18].

Mussels from a Middle Eocene (45 Myr) seep deposit in Washington State (USA) were recently assigned with some hesitation to *Vulcanidas* (as *Vulcanidas*? *goederti* Kiel & Amano, 2013). This hesitation stemmed from the missing data on anatomy and muscle attachment scars, while in all available shell characters *Vulcanidas*? *goederti* was more similar to *Vulcanidas* than to any other genus of deep-sea mussels [19]. Therefore, we modelled the divergence time of the clade including *Vulcanidas insolatus*
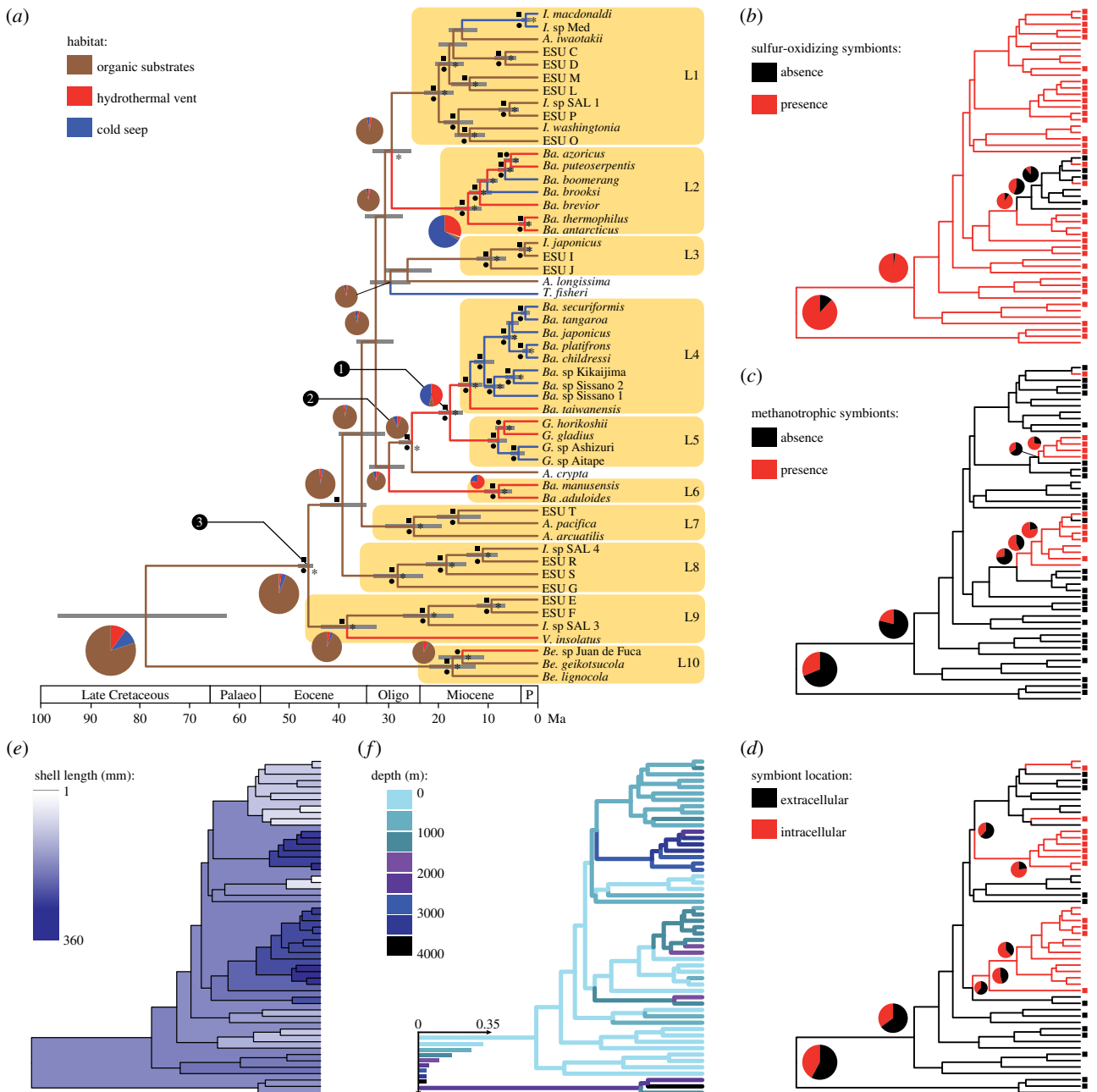
Figure 1. Maximum-clade-credibility Bayesian chronograms and estimates of ancestral character states. Pie charts indicate probabilities of each state at nodes discussed in the text (yellow rectangles). (*a*) Chronogram showing the evolution of habitat use. Grey bars are 95% HPD intervals of divergence time estimates. Black squares, circles and asterisks at nodes indicate posterior probabilities greater than or equal to 0.99, bootstrap values greater than or equal to 75% (95% for lineages highlighted in yellow), and nodes inferred in analyses of both nuclear and mitochondrial genes, respectively. The nodes that were assigned fossils to estimate divergence times of the 'childressi' group, and *Gigantidas* and *Vulcanidas* genera, are labelled in black circles by numbers 1, 2 and 3, respectively. Abbreviations on time scale: Palaeo, Palaeocene; Oligo, Oligocene; P, Pliocene to present; Ma, million years ago. (*b*) Evolution of the presence/absence of sulfur-oxidizing symbionts. (*c*) Evolution of the presence/absence of methanotrophic symbionts. (*d*) Evolution of the location of symbionts in the gill epithelium. Squares at tips of chronograms *b*, *c* and *d* indicate available data. (*e*) Evolution of shell lengths. (*f*) Evolution of depth use. For legibility, and because depth was discretized into nine states, a histogram was used instead of a pie chart to display probabilities at the root. Genus names are abbreviated as follows: *A.*, *Adipicola*; *Ba.*, *Bathymodiolus*; *Be.*, *Benthomodiolus*; *G.*, *Gigantidas*; *I.*, *Idas*; *M.*, *Modiolus*; *T.*, *Tamu*.

Cosel & Marshall, 2010 from other species (figure 1*a*, lineage L9) using an exponential prior with a hard minimum bound of 45 Myr and a mean of 1.2 Myr. This resulted in a soft maximum matching the beginning of the Middle Eocene (48.6 Myr).

Two fossil species from seep deposits on the eastern North Island of New Zealand were recently assigned to extant clades based on muscle attachment scars. The Middle Miocene (Lillburnian; 12.7–15.1 Myr) *Gigantidas coseli* Saether *et al*. 2010 was used to calibrate the divergence of *Gigantidas* (figure 1*a*, lineage L5) from other lineages. An exponential prior with a minimum bound of 15.1 Myr and a mean of 1.3 Myr was used for this node, resulting in a soft

maximum matching the beginning of the Clifdenian stage (15.9 Myr). The Early Miocene (Waitakian; 21.7–25.2 Myr) *Bathymodiolus heretaunga* Saether *et al*. 2010 was assigned to the 'childressi' clade (figure 1*a*, lineage L4 + L5) and the divergence of this clade was assigned an exponential prior with a minimum bound of 25.2 Myr and a mean of 0.7 Myr. This yielded a soft maximum matching the beginning of the Late Oligocene Duntroonian stage (27.3 Myr).

Each of the three fossil calibrations was used in a separate analysis, and then we combined them in a single analysis. For each analysis, four runs of 20 million MCMC steps, 10% burn-in

and a 4000-step subsampling of individual runs were used. The COI substitution rate and the divergence between *Bathymodiolius thermophilus* Kenk & Wilson, 1985 and *Bathymodiolius antarcticus* Johnson, Won, Harvey & Vrijenhoek 2013, which is thought to have occurred at 5 Ma [20], were used to further evaluate the consistency of results obtained from our three fossil calibrations. Molecular dating was also performed using non-parametric rate smoothing in R8s v. 1.8 [21,22]. Uncertainties in the fossil calibrations were modelled using age constraints that corresponded to the intervals used in our Bayesian analyses. This analysis was run 10 times from different starting points.

(e) Character evolution and diversification analyses

Data on habitat use, shell length, depth, presence/absence of sulfur-oxidizing and methanotrophic symbionts, and symbiont location on the gill epithelium were compiled from published sources and added to our new data (see electronic supplementary material, table S1). Because depth and depth range vary greatly among deep-sea mussels, depth was discretized using intervals of 500 m and multiple states were used to account for large depth ranges. Transitions between character states were estimated using transition matrices for categorical data [23], and a Brownian diffusion model was used for natural logarithms of shell lengths [24]. These components were added to the analysis that included all five genes and fossil calibrations, without the outgroup. As in the Bayesian dating analysis, we used four runs of 20 million MCMC steps, 10% burn-in and a 4000-step subsampling of individual runs.

To assess heterogeneity of diversification rates through time, a likelihood analysis of speciation and extinction rates [25] was conducted using the R package LASER. This analysis was based on the chronogram estimated by non-parametric rate smoothing of the ML tree (free of any speciation model). Distribution of AICc differences among a set of speciation models was obtained by simulating 5000 trees under the null hypothesis of constant diversification. The robustness of the results to sampling bias was evaluated by running simulations under the assumptions that 0, 25 and 35% of species were missing.

A similar simulation approach was carried out using the R package TESS to test whether including extinction events in the birth–death speciation model improved the likelihood score [26]. Simulations were conditioned on the age of the root and the number of species in input trees. To assess the impact of methanotrophic symbionts and intracellular symbionts on diversification rates, character-dependent Bisse [27] and Yule speciation models were fitted to the ML tree using the R package *diversitree* [28]. The significance of likelihood differences between models was tested using likelihood ratio tests. These analyses were conducted from the ML tree after removing missing data and from the patterns inferred from Bayesian analyses, assuming sampling biases of 0, 25 and 35%.

Unsupported nodes were collapsed and missing data were removed before further analyses. Correlations among categorical variables were tested (10 iterations, 1000 simulations) using Pagel's method [29] as implemented in MESQUITE v. 2.75. Influences of habitat use, presence/absence of methanotrophic symbionts, and presence/absence of intracellular symbionts on natural logarithms of shell lengths and depth were tested using phylogenetic analyses of variance [30] with the R package *geiger* (5000 simulations) [31]. Correlation between natural logarithm of shell lengths and depth was tested using the phylogenetic generalized least-squares method [32] with the R package *caper*.

3. Results and discussion

The species tree inferred from the concatenated dataset (figure 1*a*) revealed support for 10 clades, which were consistent with those observed in previous studies [9,33,34]. Our Bayesian relaxed-clock analysis, calibrated using fossils at three nodes, yielded an estimated mean age of all deep-sea symbiotic mussels of 85 Myr (95% HPD interval: 69–102). Although this result is consistent with their divergence from other Mytilidae in the Late Mesozoic or Early Cenozoic [5,33], there is a lack of mussels at the few methane seep deposits and organic falls of this age [35–39]. This discrepancy may be due to higher levels of homoplasy along deep branches and/or to the lack of age calibrations at deep nodes in the tree. Alternatively, earliest species might be absent from the fossil record [5].

Most of the 10 extant clades started diverging from each other about 45 Ma (figure 1*a*), and the main vent and seep clades (L2, L4 and L5) diversified within the last 30 Myr, consistent with previous analyses using biogeographic calibration points [40]. The robustness of these date estimates was supported by the consistency of results obtained using each of the fossil calibrations separately (see electronic supplementary material, figure S2). Differences in mean estimates among fossil calibrations were at most 0.38% for the COI substitution rate, 0.61 Myr for the divergence between the sister species *Ba. thermophilus* and *Ba. antarcticus*, 4.05 Myr for the divergence of the genus *Gigantidas* Cosel & Marshall, 2003 (lineage L5), and 5.67 Myr for the divergence of the '*childressi*' clade (lineages L4+L5). In the analysis including the three fossil calibrations, the COI substitution rate estimated across the entire tree was 1.62×10^{-2} substitutions per site per Myr (95% HPD interval: 1.22×10^{-2}–2.09×10^{-2}), a value higher than that estimated in vent-endemic annelid taxa, but close to those found in other invertebrates from non-chemosynthetic shallow water environments [41,42]. Our estimate for the divergence between *Ba. thermophilus* and *Ba. antarcticus* was 2.77 Myr (95% HPD interval: 1.7–4.0), a result similar to those obtained from population genetic studies [20,43].

The inference that the divergences among and within clades occurred within the last 45 Myr is remarkable because the ecological niche of deep-sea mussels—being epifaunal and taking up sulfide from the water column—was unoccupied since their inferred origin in the Late Cretaceous [44–45]. Even assuming the lower (younger) bound of the 95% HPD interval as the time of origin of the deep-sea symbiotic mussels still implies that it took 25 Myr from the origin to the diversification event seen in our tree (figure 1*a*). This long basal branch led other authors to propose the hypothesis that the genus *Benthomodiolus* (L10), which is a sister group to the remaining nine clades, may be a relic of lineages that became extinct during the global anoxia/dysoxia event associated with the Palaeocene/Eocene thermal maximum (PETM) around 57 Ma [33,46]. We found some support for that hypothesis, since the likelihood of our ML and Bayesian maximum-clade-credibility trees under a birth–death speciation model significantly improved ($p = 0.002$ and $p = 0.024$, respectively) when including an extinction event (5% survival rate) at 57 Ma. However, our simulations also show that a scenario without extinction during this time produced very similar trees, inducing a lack of statistical power in the case of Bayesian trees (figure 2). Overall, our analyses also suggested that chemosymbiotic mussels were not very diverse during the PETM and that it induced, if anything, a simple slowdown of the early diversification rather than a dramatic extinction event.

The slow initial evolution of deep-sea mussels contrasts with the timing of branching events that occurred from 45 Ma. A likelihood analysis of speciation and extinction
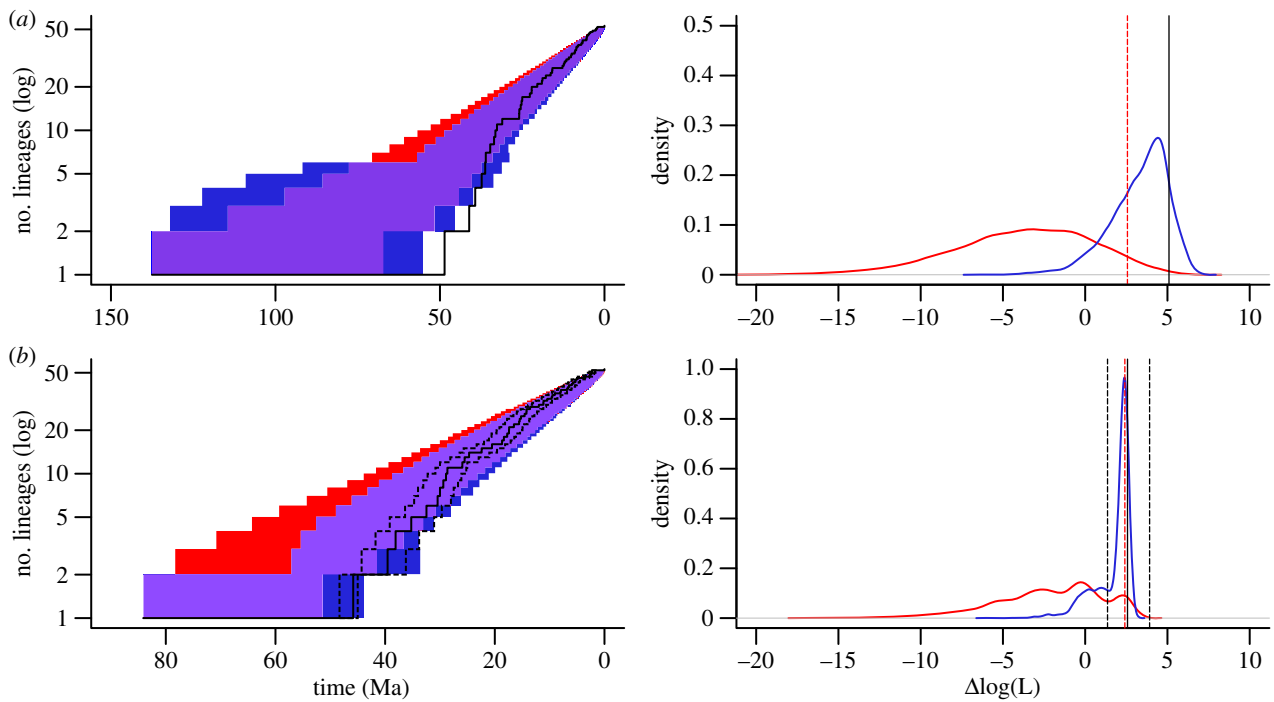
Figure 2. Comparison between extinction-free birth–death models (H0) and models including a mass extinction at 57 Ma (H1). Analyses were performed on chronograms inferred using (*a*) non-parametric rate smoothing and (*b*) Bayesian phylogenetic analysis. Lineages-through-time plots (left) obtained from these trees are represented by solid black lines, with additional dashed black lines for the 95% HPD interval estimated from the entire distribution of sampled Bayesian trees. Red- and blue-shaded areas correspond to 95% CIs obtained by simulation under H0 and H1, respectively, and purple areas show the overlap in confidence intervals between models. Distributions of likelihood differences between both models fitted to datasets simulated under H0 (red) and H1 (blue) are given on the right. Dashed red lines represent H0's 5% rejection levels. Likelihood differences calculated from real trees are given by solid vertical black lines, with an additional 95% HPD interval for Bayesian trees (vertical black dashed lines). Ma, million years ago; Δlog(L), difference between natural logarithms of likelihood.

rates allowed us to identify a moderate but distinct increase in speciation rates, with about 0.17 speciation events per million years from 41 to 32 Ma (Middle Eocene to earliest Oligocene; figure 3). From 32 Ma to the present day, the estimated speciation rate was three times lower. The corresponding Yule model with three distinct speciation rates provided a significantly better fit ($p = 0.007$) to the tree than the best constant-rate model (i.e. pure birth), even in simulations assuming 25% ($p = 0.011$) and 35% ($p = 0.009$) sampling bias (table 1). Moreover, parts of the gene trees and species tree that fall within that interval are consistently characterized by poorly resolved nodes and short internode distances (figure 1*a*). It is worth noting that a complementary analysis of speciation rates from 41 Myr to the present suggests that a slow decrease of speciation rates, following a Weibull distribution ($\beta = 1.42$), might also be an alternative to the simple Yule model ($\chi^2 = 8.33$, $p = 0.004$). This indicates that mussels have continued to diversify after their initial radiation, as corroborated by the existence of several young species and species complexes [20,46].

To test if and how symbiont type and location have influenced the diversification of mussels, we estimated the ancestral states of these characters. Our results indicate that sulfur-oxidizing symbionts are an early acquisition going back to the stem of the group, almost 30 Myr before the main burst of diversification (figure 1*b*). In contrast, the ability to host intracellular symbionts was detected in our tree from the Late Oligocene onward, and methanotrophic symbionts only from the Middle Miocene onward (figures 1*c,d* and 3). This suggests that the evolution of the various symbiotic relationships did not trigger the Middle Eocene–Early Oligocene burst of diversification, although the association with sulfur-

oxidizing symbionts was probably a prerequisite. A potential environmental trigger for this burst might have been the appearance of whales, because their carcasses, which produce large amounts of sulfide, are thought to be dispersal stepping stones [47]. This hypothesis was originally based on the congruence of molecular age estimates for the chemosymbiotic bivalve family Vesicomyidae and the rise of whales, but it has been challenged because Eocene and Oligocene whale falls lacked associated vesicomyid fossils [48]. However, these early whale falls were abundantly colonized by mussels, and thus the 'whale stepping stone' hypothesis might apply to mussels. Alternatively, the geological record shows a rapid increase in the abundance of seep carbonates in the Late Eocene [49], indicating a worldwide increase in methane seepage. Analogous to the rationale of the 'whale stepping stone' hypothesis, this increase in habitat availability and in potential dispersal stepping stones could have triggered the Late Eocene radiation. A third possible trigger is a drop in deep-water temperature beginning in the Middle to Late Eocene, associated with the initial glaciation of Antarctica [50]. Low ambient water temperatures could have decreased the metabolic rate of mussel larvae, thus increasing their longevity and enhancing their dispersal capability. This would have been an advantage for mussels living in patchy deep-sea habitats such as vents, seeps and organic substrates.

Even if the acquisition of intracellular and methanotrophic symbionts was not involved in the Middle Eocene to Early Oligocene radiation, symbionts played a role in the subsequent evolution of the group. Pagel's tests showed that both the presence of methanotrophic symbionts and the intracellular location of symbionts are correlated with the occurrence at vents and seeps (table 2). This supports the
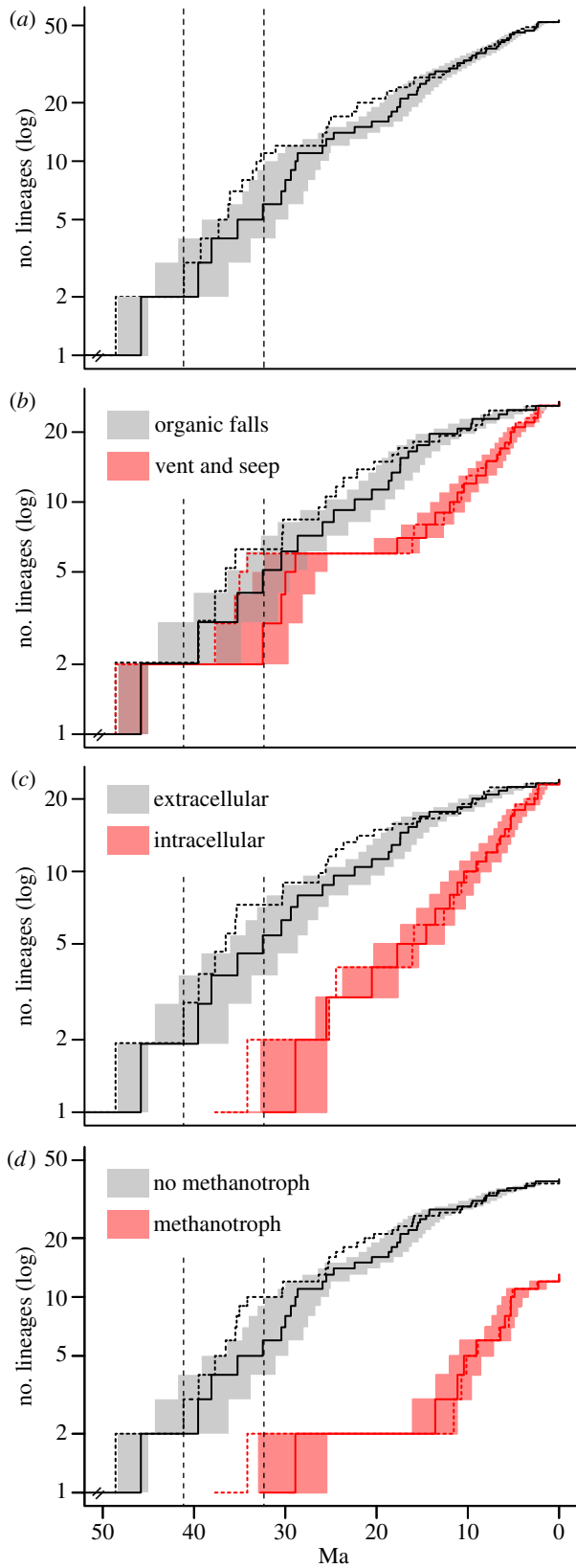
Figure 3. (a) Global and (b–d) character-dependent lineages-through-time plots (log-scaled) estimated from the ML tree smoothed using non-parametric rate smoothing (dashed lines) and from the Bayesian chronogram (solid lines). Shaded areas represent 95% HPD intervals estimated from the entire distribution of sampled Bayesian trees. Vent and seep lineages from figure 1a were pooled together in the habitat-dependent plot. Colour-coded groups in plots based on the location and the presence/absence of methanotrophic symbionts match the distributions of character states among taxa in figure 1d and 1c, respectively. Vertical dashed lines represent shifts of the net diversification rate estimated in the likelihood analysis of speciation and extinction rates.

Table 1. Results of the likelihood analysis of speciation and extinction rates (LASER). Characteristics of each model are abbreviated as follows: RC, rate constant; RV, rate variable; L, likelihood; $r1$, first diversification rate; $r2$, second diversification rate; $r3$, third diversification rate; a, extinction fraction of the birth−death model (ratio extinction/speciation); xp, x-parameter from the exponential variant of the density-dependent speciation rate (DDX) model; k, k-parameter from the logistic variant of the density-dependent speciation rate (DDL) model; $s1$, first break in diversification rate (million years); $s2$, second break in diversification rate (million years); dAIC, the difference in AIC scores between the model and the overall best-fit model.

model	parameters	type	L	r1	r2	a	k	xp	s1	s2	r3	AIC	dAIC
pure birth	r1	RC	−54.49	0.05	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	111.0	12.4
birth−death	r1, a	RC	−54.44	0.04	n.a.	0.02	n.a.	n.a.	n.a.	n.a.	n.a.	112.9	14.3
DDL	r1, k	RV	−54.45	0.05	n.a.	n.a.	429	n.a.	n.a.	n.a.	n.a.	112.9	14.3
DDX	r1, x	RV	−54.03	0.03	n.a.	n.a.	n.a.	−0.13	n.a.	n.a.	n.a.	112.1	13.5
Yule 2 rates	r1, r2, ts	RV	−50.30	0.01	0.05	n.a.	n.a.	n.a.	39.6	n.a.	n.a.	106.6	8.0
Yule 3 rates	r1, r2, r3, s1, s2	RV	−44.29	0.01	0.17	n.a.	n.a.	n.a.	41.2	33.8	0.05	98.6	0

Table 2. Tests of correlations between characters. Model used, number of species for which data were available, criteria estimated, degrees of freedom (d.f.) and p-values are given for each pair of characters.

model	variables	species	criterion	d.f.	p-value
Pagel's test	methanotroph ∼ symbiont location	24	$\Delta\log(L) = 4.916$		0.001
	methanotroph ∼ habitat	36	$\Delta\log(L) = 3.900$		0.024
	symbiont location ∼ habitat	24	$\Delta\log(L) = 4.331$		0.006
pGLS	log(size) ∼ depth	48	intercept: $t = 4.460$		0
			slope: $t = 0.471$		0.640
			$F = 0.222$	46	0.802
pAnova	log(size) ∼ habitat	48	$F = 61.328$	1	0
			phylogenetic p-value		0
	log(size) ∼ symbiont location	24	$F = 9.571$	1	0.005
			phylogenetic p-value		0.053
	log(size) ∼ methanotroph	36	$F = 3.341$	1	0.057
			phylogenetic p-value		0.143
	log(depth) ∼ symbiont location	24	$F = 6.994$	1	0.015
			phylogenetic p-value		0.097
	log(depth) ∼ methanotroph	36	$F = 6.153$	1	0.018
			phylogenetic p-value		0.067

adaptive significance of these two character states and the hypothesis that symbionts 'increase the metabolic capabilities, and therefore the number of ecological niches' colonized by mussels [51, p. 475]. We further explored this hypothesis by comparing simple Yule speciation models with models in which the diversification rates depend on character states. Even when assuming high levels of sampling bias, the results suggest that diversification rates are higher in lineages hosting intracellular symbionts and methanotrophic symbionts (table 3). The acquisition of these two characters is correlated with a second wave of mussel diversification at vents and seeps (figure 3), and is probably the key to the success of vent and seep clades (lineages L2, L4–L6 in figure 1a). These include nearly as many species as from organic falls, but have higher speciation rates because they started diversifying only since the Miocene. Methanotrophic symbionts were only acquired in species hosting intracellular symbionts (figure 1c,d; electronic supplementary material, table S1) and both characters are strongly correlated (table 2). This suggests that the ability to host intracellular symbionts was a prerequisite for the acquisition of methanotrophic symbionts.

Our estimation of ancestral habitat types indicates that vents and seeps were colonized from organic falls independently in six lineages, consistent with the hypothesis that these mussels took 'wooden steps to deep-sea vents' [8]. These habitat transitions were unidirectional, suggesting that the adaptation to vents and seeps is an irreversible specialization. This might be partly explained by the evolution of symbioses toward more integrated relationships (i.e. intracellular symbioses) and the concomitant reduction of the digestive tract in several species [46,52]. Specialization might also have resulted from the evolution of other anatomical and physiological traits.

Indeed, we found a significant correlation between habitat type and shell size (table 2). We also inferred a trend towards larger shells in vent/seep species from the Early to Middle Miocene onwards and towards smaller shells in lineages living on organic falls (figure 1e). This pattern is remarkably consistent with the fossil record, which shows that seep-inhabiting mussels did not exceed 50 mm in length until the Early Miocene [53–55], and then rapidly increased in size from 100 mm in the Middle Miocene to more than 300 mm today. Compared with the typically small and ephemeral organic substrates on the seafloor, vents and seeps provide large amounts of energy and habitable space that might explain the increased size of species living there [54]. By contrast, species living on sparse organic debris would benefit from becoming smaller to reduce competition for space and/or for early allocation of resources to reproduction. The latter hypothesis is corroborated by studies suggesting that mussels from organic substrates reach sexual maturity quickly and at much smaller size [56–58] than their vent and seep relatives [59].

Earlier studies indicated a trend of successive adaptation to greater water depth among mussels [33]. This trend is also seen in our data for post-Eocene taxa (figure 1f); however, it coincides with habitat transitions from organic substrates to vents and seeps, and the latter habitats typically occur in deeper water than organic falls (Mann–Whitney: $p = 0.001$). Thus, the colonization of greater depth might just be a consequence of the habitat transitions. Alternatively, there may be a systematic sampling bias because wood and bones are much more difficult to locate on the vast deep-sea floor than vents and seeps, which can be detected using temperature anomalies or seismically reflective sediment layers [60,61].

In conclusion, the evolutionary history of chemosymbiotic deep-sea mussels shows all the characteristics of an adaptive radiation: a burst of diversification into a suite of new habitats due to an ecological opportunity, along with physiological and morphological adaptations to these habitats. Although these mussels rely on their symbionts for nutrition, it is remarkable that the initial acquisition of sulfur-oxidizing symbionts did not trigger a major radiation. This adds to the growing

Table 3. Impact of methanotrophic symbionts and symbiont location on diversification rates. Character-dependent Bisse (H1: diversification rate λ_0 under state 0 and λ_1 under state 1) and simple Yule (H0: global diversification rate λ_{Yule}) speciation models were fitted to sampled data and patterns inferred during the Bayesian analysis. Levels of sampling bias considered are expressed as ratios and reflect estimated proportions of extant species with state 0 and 1 included in the phylogeny. Likelihood ratios (LRs) calculated between Bisse and Yule models and associated p-values (p) are given.

dataset	methanotrophic symbionts					symbiont location				
	λ_0^{absent}	$\lambda_1^{present}$	λ_{Yule}	LR	p-value	$\lambda_0^{extracellular}$	$\lambda_1^{intracellular}$	λ_{Yule}	LR	p-value
sampled data	0.010	0.088	0.037	12.938	0	0.004	0.055	0.028	10.276	0.001
sampled data, bias ~ 0.75 : 1	0.012	0.099	0.041	12.517	0	0.003	0.064	0.031	9.891	0.002
sampled data, bias ~ 0.65 : 1	0.013	0.105	0.043	12.346	0	0.003	0.068	0.033	9.749	0.002
sampled data, bias ~ 1 : 0.75	0.039	0.010	0.091	14.312	0	0.007	0.057	0.031	11.562	0.001
sampled data, bias ~ 1 : 0.65	0.039	0.010	0.093	14.880	0	0.007	0.060	0.032	12.303	0
inferred pattern	0.022	0.099	0.043	14.711	0	0.021	0.084	0.043	8.146	0.004
inferred pattern, bias ~ 0.75 : 1	0.027	0.107	0.049	11.871	0.001	0.038	0.084	0.047	6.948	0.008
inferred pattern, bias ~ 0.65 : 1	0.030	0.112	0.052	10.562	0.001	0.041	0.084	0.050	5.928	0.015
inferred pattern, bias ~ 1 : 0.75	0.022	0.105	0.045	17.79	0	0.021	0.092	0.046	12.515	0
inferred pattern, bias ~ 1 : 0.65	0.022	0.107	0.045	19.315	0	0.033	0.110	0.047	16.696	0

evidence that an early burst might be rare in adaptive radiations [62]. The link of the burst of diversification in the Middle Eocene to Early Oligocene to ecological opportunities such as increased habitat availability or dispersal capability raises the question of whether other chemosymbiotic deep-sea taxa were similarly affected. Interestingly, a compilation and reassessment of molecular age estimates for 14 vent/seep taxa showed that the inferred origins of six of these (Alvinellidae, Lepetodrilidae, *Alviniconcha/Ifremeria*, *Provanna*, Bresiliidae and Bythograeidae) fall within the time interval of the mussels' burst of diversification [41]. For the mussels, we consider the acquisition of sulfur-oxidizing symbionts as a prerequisite for their adaptation to, and successful radiation within, chemosynthetic environments. By contrast, the subsequent acquisition of methanotrophic symbionts allowed the colonization of new niches within the vent and seep environment, and resulted in a second wave of diversification.

References

1. Simpson GG. 1953 *The major features of evolution*. New York, NY: Columbia University Press.

2. Glor RE. 2010 Phylogenetic insights on adaptive radiation. *Annu. Rev. Ecol. Evol. Syst.* **41**, 251–270. (doi:10.1146/annurev.ecolsys.39.110707.173447)

3. Dubilier N, Bergin C, Lott C. 2008 Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat. Rev. Microbiol.* **6**, 725–740. (doi:10.1038/nrmicro1992)

4. Newman WA. 1985 The abyssal hydrothermal vent invertebrate fauna. A glimpse of antiquity? *Bull. Biol. Soc. Washington* **6**, 231–242.

5. Little CTS, Vrijenhoek RC. 2003 Are hydrothermal vent animals living fossils? *Trends Ecol. Evol.* **18**, 582–588. (doi:10.1016/j.tree.2003.08.009)

6. Kiel S, Little CTS. 2006 Cold-seep mollusks are older than the general marine mollusk fauna. *Science* **313**, 1429–1431. (doi:10.1126/science.1126286)

7. Duperron S. 2010 The diversity of deep-sea mussels and their bacterial symbioses. In *The vent and seep biota* (ed. S Kiel), pp. 137–167. New York, NY: Springer.

8. Distel DL, Baco AR, Chuang E, Morrill W, Cavanaugh C, Smith CR. 2000 Marine ecology: do mussels take wooden steps to deep-sea vents? *Nature* **403**, 725–726. (doi:10.1038/35001667)

9. Lorion J, Buge B, Cruaud C, Samadi S. 2010 New insights into diversity and evolution of deep-sea Mytilidae (Mollusca: Bivalvia). *Mol. Phylogenet. Evol.* **57**, 71–83. (doi:10.1016/j.ympev.2010.05.027)

10. Bouchet P, Heros V, Lozouet P, Maestrati P. 2008 A quarter-century of deep-sea malacological exploration in the South and West Pacific: where do we stand? How far to go? In *Mémoires du Muséum national d'histoire naturelle* (ed. Editions du Muséum.), pp. 9–40. Paris, France: Editions du Muséum.

11. Fujiwara Y, Kawato M, Noda C, Kinoshita G, Yamanaka T, Fujita Y, Uematsu K, Miyazaki J-I. 2010 Extracellular and mixotrophic symbiosis in the whale-fall mussel *Adipicola pacifica*: a trend in evolution from extra- to intracellular symbiosis. *PLoS ONE* **5**, e11808. (doi:10.1371/journal.pone.0011808)

12. Roshan U, Livesay DR. 2006 Probalign: multiple sequence alignment using partition function posterior probabilities. *Bioinformatics* **22**, 2715–2721. (doi:10.1093/bioinformatics/btl472)

13. Posada D. 2008 jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256. (doi:10.1093/molbev/msn083)

14. Drummond AJ, Rambaut A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)

15. Drummond AJ, Ho SY-W, Phillips MJ, Rambaut A. 2006 Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, 699–710. (doi:10.1371/journal.pbio.0040088)

16. Suchard MA, Weiss RE, Sinsheimer JS. 2001 Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* **18**, 1001–1013. (doi:10.1093/oxfordjournals.molbev.a003872)

17. Jobb G, von Haeseler A, Strimmer K. 2004 TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* **4**, 18. (doi:10.1186/1471-2148-4-18)

18. Ho SY-W, Phillips MJ. 2009 Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* **58**, 367–380. (doi:10.1093/sysbio/syp035)

19. Kiel S, Amano K. 2013 The earliest bathymodiolin mussels: an evaluation of Eocene and Oligocene taxa from deep-sea methane seep deposits in Western Washington State, USA. *J. Paleontol.* **87**, 589–602. (doi:10.1666/12-135)

20. Won Y, Young CR, Lutz RA, Vrijenhoek RC. 2003 Dispersal barriers and isolation among deep-sea mussel populations (Mytilidae: *Bathymodiolus*) from eastern Pacific hydrothermal vents. *Mol. Ecol.* **12**, 169–184. (doi:10.1046/j.1365-294X.2003.01726.x)

21. Sanderson MJ. 1997 A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1231. (doi:10.1093/oxfordjournals.molbev.a025731)

22. Sanderson MJ. 2003 r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301–302. (doi:10.1093/bioinformatics/19.2.301)

23. Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009 Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**, e1000520. (doi:10.1371/journal.pcbi.1000520)

24. Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010 Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885. (doi:10.1093/molbev/msq067)

25. Rabosky DL. 2006 LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol. Bioinformatics* **2**, 247–250.

26. Höhna S. 2013 Fast simulation of reconstructed phylogenies under global, time-dependent birth–death processes. *Bioinformatics* **29**, 1367–1374. (doi:10.1093/bioinformatics/btt153)

27. Maddison WP, Midford PE, Otto SP. 2007 Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710. (doi:10.1080/10635150701607033)

28. FitzJohn RG. 2012 Diversitree: comparative phylogenetic analyses of diversification in R. *Meth. Ecol. Evol.* **3**, 1084–1092. (doi:10.1111/j.2041-210X.2012.00234.x)

29. Pagel M. 1994 Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. B* **255**, 37–45. (doi:10.1098/rspb.1994.0006)

30. Garland T, Dickerman AW, Janis CM, Jones JA. 1993 Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* **42**, 265–292. (doi:10.2307/2992464)

31. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008 GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131. (doi:10.1093/bioinformatics/btm538)

32. Grafen A. 1989 The phylogenetic regression. *Phil. Trans. R. Soc. Lond. B* **326**, 119–157. (doi:10.1098/rstb.1989.0106)

33. Jones WJ, Won YJ, Maas PAY, Smith PJ, Lutz RA, Vrijenhoek RC. 2006 Evolution of habitat use by deep-sea mussels. *Mar. Biol.* **148**, 841–851. (doi:10.1007/s00227-005-0115-1)

34. Kyuno A, Shintaku M, Fujita Y, Matsumoto H, Utsumi M, Watatabe H, Fujiwara Y, Miyazaki J-I. 2009 Dispersal and differentiation of deep-sea mussels of the genus *Bathymodiolus* (Mytilidae, Bathymodiolinae). *Mar. Biol.* **2009**, 625672. (doi:10.1155/2009/625672)

35. Hikida Y, Suzuki S, Togo Y, Ijiri A. 2003 An exceptionally well-preserved seep community from the Cretaceous Yezo forearc basin in Hokkaido, northern Japan. *Paleontol. Res.* **7**, 329–342. (doi:10.2517/prpsj.7.329)

36. Kaim A, Kobayashi Y, Echizenya H, Jenkins RG, Tanabe K. 2008 Chemosynthesis-based associations on Cretaceous plesiosaurid carcasses. *Acta Palaeontol. Pol.* **53**, 97–104. (doi:10.4202/app.2008.0106)

37. Nobuhara T, Onda D, Kikuchi N, Kondo Y, Matsubara K, Amano K, Jenkins RG, Hikida Y, Majima R. 2008 Lithofacies and fossil assemblages of the Upper Cretaceous Sada Limestone, Shimanto City, Kochi Prefecture, Shikoku, Japan. *Palaeontol. Soc. Jpn* **84**, 47–60.

38. Kiel S, Amano K, Hikida Y, Jenkins RG. 2009 Wood-fall associations from Late Cretaceous deep-water sediments of Hokkaido, Japan. *Lethaia* **42**, 74–82. (doi:10.1111/j.1502-3931.2008.00105.x)

39. Kiel S, Birgel D, Campbell KA, Crampton JS, Schiøler P, Peckmann J. In press. Cretaceous methane-seep deposits from New Zealand and their fauna. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*

40. Miyazaki J-I, Martins LD, Fujita Y, Matsumoto H, Fujiwara Y. 2010 Evolutionary process of deep-sea *Bathymodiolus* mussels. *PLoS ONE* **5**, e10363. (doi:10.1371/journal.pone.0010363)

41. Vrijenhoek RC. 2013 On the instability and evolutionary age of deep-sea chemosynthetic communities. *Deep-Sea Res. II* **92**, 189–200. (doi:10.1016/j.dsr2.2012.12.004)

42. Lessios HA. 2008 The Great American Schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annu. Rev. Ecol. Evol. Syst.* **39**, 63–91. (doi:10.1146/annurev.ecolsys.38.091206.095815)

43. Faure B, Jollivet D, Tanguy A, Bonhomme F, Bierne N. 2009 Speciation in the deep sea: multi-locus analysis of divergence and gene flow between two hybridizing species of hydrothermal vent mussels. *PLoS ONE* **4**, e6485. (doi:10.1371/journal.pone.0006485)

44. Kiel S. 2010 The fossil record of vent and seep mollusks. In *The vent and seep biota* (ed. S Kiel), pp. 255–278. New York, NY: Springer.

45. Jenkins RG, Kaim A, Little CTS, Iba Y, Tanabe K, Campbell KA. 2013 Worldwide distribution of modiomorphid bivalve genus *Caspiconcha* in late Mesozoic hydrocarbon seeps. *Acta Palaeontol. Pol.* **58**, 357–382.

46. Olu-Le Roy K, von Cosel R, Hourdez S, Carney SL, Jollivet D. 2007 Amphi-Atlantic cold-seep *Bathymodiolus* species complexes across the equatorial belt. *Deep-Sea Res. I* **54**, 1890–1911. (doi:10.1016/j.dsr.2007.07.004)

47. Baco AR, Smith CR, Peek AS, Roderick GK, Vrijenhoek RC. 1999 The phylogenetic relationships of whale-fall vesicomyid clams based on mitochondrial COI DNA sequences. *Mar. Ecol. Progr. Ser.* **182**, 137–147. (doi:10.3354/meps182137)

48. Kiel S, Goedert JL. 2006 Deep-sea food bonanzas: Early Cenozoic whale-fall communities resemble wood-fall rather than seep communities. *Proc. R. Soc. B* **273**, 2625–2631. (doi:10.1098/rspb.2006.3620)

49. Kiel S. 2009 Global hydrocarbon seep-carbonate precipitation correlates with deep-water temperatures and eustatic sea-level fluctuations since the Late Jurassic. *Terra Nova* **21**, 279–284. (doi:10.1111/j.1365-3121.2009.00882.x)

50. Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001 Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693. (doi:10.1126/science.1059412)

51. Cavanaugh CM, McKiness ZP, Newton ILG, Stewart FJ. 2006 Marine chemosynthetic symbioses. In *The Procaryotes: a handbook on the biology of bacteria: symbiotic associations, biotechnology, applied microbiology* (eds M Dworkin, S Falkow, E Rosenberg, K Schleifer, E Stackebrandt), 3rd edn, pp. 475–507. New York, NY: Springer.

52. Von Cosel R. 2002 A new species of bathymodioline mussel (Mollusca, Bivalvia, Mytilidae) from Mauritania (West Africa), with comments on the genus *Bathymodiolus* Kenk & Wilson 1985. *Zoosystema* **24**, 259–271.

53. Kiel S, Campbell KA, Gaillard C. 2010 New and little known mollusks from ancient chemosynthetic environments. *Zootaxa* **2390**, 26–48.

54. Saether KP, Little CTS, Campbell KA, Marshall BA, Collins M, Alfaro AC. 2010 New fossil mussels (Bivalvia: Mytilidae) from Miocene hydrocarbon seep deposits, North Island, New Zealand, with general remarks on vent and seep mussels. *Zootaxa* **2577**, 1–45.

55. Amano K, Jenkins RG. 2011 New fossil *Bathymodiolus* (s. l.) (Mytilidae, Bivalvia) from Oligocene seep-carbonates in eastern Hokkaido, Japan — with remarks on the evolution of *Bathymodiolus* (s. l.). *Nautilus* **125**, 29–35.

56. Tyler PA, Marsh L, Baco-Taylor A, Smith CR. 2009 Protandric hermaphroditism in the whale-fall bivalve mollusc *Idas washingtonia*. *Deep-Sea Res. Part II* **56**, 1689–1699. (doi:10.1016/j.dsr2.2009.05.014)

57. Ockelmann KW, Dinesen GE. 2011 Life on wood—the carnivorous deep-sea mussel *Idas argenteus* (Bathymodiolinae, Mytilidae, Bivalvia). *Mar. Biol. Res.* **7**, 71–84. (doi:10.1080/17451001003714504)

58. Gaudron SM, Demoyencourt E, Duperron S. 2012 Reproductive traits of the cold-seep symbiotic mussel *Idas modiolaeformis*: gametogenesis and larval biology. *Biol. Bull.* **222**, 6–16.

59. Berg CJJ. 1985 Reproductive strategies of mollusks from abyssal hydrothermal vent communities. *Bull. Biol. Soc. Wash.* **6**, 185–187.

60. Hornbach MJ, Holbrook WS, Gorman AR, Hackwith KL, Lizarralde D, Pecher I. 2003 Direct seismic detection of methane hydrate on the Blake Ridge. *Geophysics* **68**, 92–100. (doi:10.1190/1.1543196)

61. German CR, Yoerger DR, Jakuba M, Shank TM, Langmuir CH, Nakamura K-I. 2008 Hydrothermal exploration with the Autonomous Benthic Explorer. *Deep-Sea Res. Part I* **55**, 203–219. (doi:10.1016/j.dsr.2007.11.004)

62. Harmon LJ et al. 2010 Early bursts of body size and shape evolution are rare in comparative data. *Evolution* **64**, 2385–2396. (doi:10.1111/j.1558-5646.2010.01025.x)