CrossMark

RESEARCH ARTICLE

Population structure and connectivity in Indo-Pacific deep-sea mussels of the *Bathymodiolus septemdierum* complex

Corinna Breusing[1] · Shannon B. Johnson[2] · Verena Tunnicliffe[3] · Robert C. Vrijenhoek[2]

Abstract Current pressures to mine polymetallic sulfide deposits pose threats to the animal communities found at deep-sea hydrothermal vents. Management plans aimed at preserving these unusual communities require knowledge of historical and contemporary forces that shaped the distributions and connectivity of associated species. As most vent research has focused on the eastern Pacific and mid-Atlantic ridge systems less is known about Indo-Pacific vents, where mineral extraction activities are imminent. Deep-sea mussels (Bivalvia: Mytilidae) of the genus *Bathymodiolus* include the morphotypic species *B. septemdierum*, *B. brevior*, *B. marisindicus*, and *B. elongatus* which are among the dominant vent taxa in western Pacific back-arc basins and the Central Indian Ridge. To assess their interpopulational relationships, we examined multilocus genotypes based on DNA sequences from four nuclear and four mitochondrial genes, and allozyme variation encoded by eleven genes. Bayesian assignment methods grouped mussels from seven widespread western Pacific localities into a single cluster, whereas the Indian Ocean mussels were clearly divergent. Thus, we designate two regional metapopulations. Notably, contemporary migration rates among all sites appeared to be low despite limited population differentiation, which highlights the necessity of obtaining realistic data on recovery times and fine-scale population structure to develop and manage conservation units effectively. Future studies using population genomic methods to address these issues in a range of species will help to inform management plans aimed at mitigating potential impacts of deep-sea mining in the Indo-Pacific region.

✉ Corinna Breusing
  cbreusing@geomar.de

[1] Evolutionary Ecology of Marine Fishes, GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany

[2] Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA

[3] University of Victoria, PO Box 1700, Victoria, BC V8W 3Y3, Canada

Introduction

Unlike most other ecosystems, deep-sea hydrothermal vent communities rely on chemosynthetic primary production, where chemoautotrophic microbes perform redox reactions to produce energy for carbon fixation. Many vent animals host sulfide- or methane-oxidizing bacteria as endo- or ectosymbionts (Dubilier et al. 2008), leading to biomass values that are among the highest on the planet (Ramirez-Llodra et al. 2010). Their dependence on chemosynthesis restricts vent species to tectonically or volcanically active seafloor regions. The patchy distribution and ephemeral nature of hydrothermal habitats increases the vulnerability of vent species to environmental disturbances and makes their dispersal among active vent fields a nontrivial challenge. Financial and technical limitations often impede thorough investigations of these remote systems, so that details about connectivity and biodiversity remain poorly understood. Despite these shortcomings, plans for mining of seafloor massive sulfide (SMS) deposits present an acute and realistic risk to these unique biological communities

🕿 Springer

(Van Dover 2011; Boschen et al. 2013). More seriously, the scarce information we have about population genetics and ecology of vent-endemic taxa comes primarily from studies in the eastern Pacific and mid-Atlantic Ocean ridge systems (Vrijenhoek 2010). However it is the less studied Indo-West-Pacific back-arc and ridge spreading centers that form current targets for SMS mining (SPC 2013). Licenses have been granted by the Fiji, Vanuatu and Tonga governments to explore SMS deposits in their territorial waters and Nautilus Minerals Inc. was granted the first mining lease for vent fields in the territorial waters of Papua New Guinea (for examples, see: http://www.nautilusminerals.com/s/Home.asp). The accelerating pace of licensing and exploration emphasizes the need for more knowledge about the genetic population structure, contemporary gene flow patterns and historical migration routes of Indo-Pacific species. Such data will aid estimates of the recovery potential of hydrothermal ecosystems subject to human impact.

Bathymodiolin mussels, hosting sulfur, methane, and/or hydrogen oxidizing bacterial symbionts (Petersen et al. 2011), are among the dominant taxa in deep-sea reducing environments worldwide. In the western (W.) Pacific, they frequently form extensive beds (Fig. 1) and dominate both biomass and community dynamics (e.g., Sen et al. 2014). More than 50 species (named and unnamed operational taxonomic units) are presently recognized (Lorion et al. 2013), where species currently placed in the genus *Bathymodiolus* are the most diverse at hydrothermal vents.

Three morphospecies with similar features were described from the W. Pacific Ocean: *B. septemdierum* Hashimoto and Okutani (August 1994); *B. brevior* Von Cosel et al. (October 1994, p. 375); and *B. elongatus* Von Cosel et al. (October 1994, p. 379). A fourth, *B.*
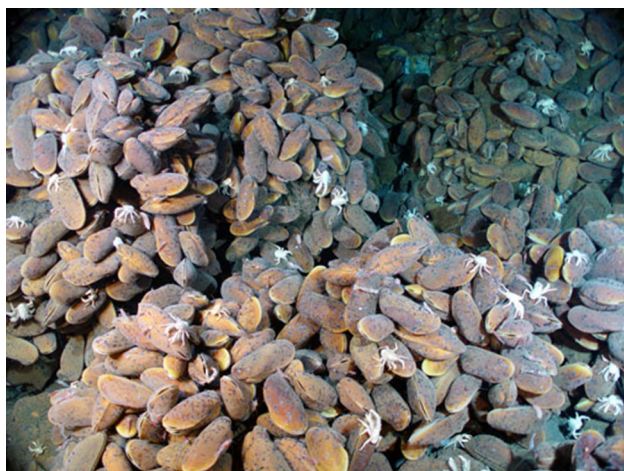

Fig. 1 A bed of *Bathymodiolus* mussels at NW Eifuku Seamount. Photo from *ROPOS* (Canadian Scientific Submersible Facility); courtesy NOAA Ocean Exploration

marisindicus Hashimoto (2001), was described from the Central Indian Ridge (CIR), although evidence from a single mitochondrial gene (*ND4* = NADH dehydrogenase subunit 4) indicated 99.2 % sequence identity with W. Pacific *B. brevior* (Van Dover et al. 2001). Subsequent studies of additional mitochondrial and nuclear genes found little evidence for differentiation among the Indian Ocean and various W. Pacific populations and raised doubts about their status as distinct species (Miyazaki et al. 2004; Jones et al. 2006; Won et al. 2008). Kyuno et al. (2009) conducted phylogeographic analyses of *B. brevior* samples from the North Fiji Basin (NFB), *B. septemdierum* from the Myojin Knoll (southern Izu-Bonin Arc), and *B. marisindicus* from the Kairei (KA) field in the Indian Ocean. Based on their analysis of mitochondrial *ND4*, Kyuno et al. proposed that: (1) essentially unimpeded gene flow exists between W. Pacific populations of *B. brevior* and *B. septemdierum* separated by about 5000 km; (2) *B. marisindicus* is "not isolated" from the W. Pacific populations, though gene flow is "relatively limited"; and (3) the ancestor of the three species might have migrated from the Southern CIR to the Izu-Bonin Arc via the Southwest Pacific.

Based on the existing molecular evidence, Thubaut et al. (2013) recognize *B. septemdierum* as the prior synonym for *B. brevior* and *B. marisindicus*. Although they keep *B. elongatus* as a separate species, they concede that this morphotype might be conspecific as well, given the low evolutionary distances to *B. septemdierum*. Von Cosel and coworkers (pers. comm., 30 October 2014) are using morphological criteria and all the available molecular data to revise the taxonomy of *Bathymodiolus* mussels.

The present study is intended to provide more robust biogeographic and genetic foundations for revision of the *B. septemdierum* complex. We examined samples from 10 vent fields distributed among five Indo-Pacific locations including ridge and back-arc spreading centers, plus volcanic arcs (Table 1; Fig. 2). Multilocus genotypes of individual mussels were based on 19 gene regions: 11 allozyme loci, four nuclear DNAs, and four mitochondrial DNAs. Genotypic assignment methods recognized two regional contemporaneously isolated metapopulations inhabiting the CIR and W. Pacific back-arc spreading centers.

Materials and methods

Samples

Mussel specimens were obtained from 10 vent habitats in the W. Pacific and Indian oceans that were visited during multiple expeditions between 1992 and 2007 (Table 1;

Table 1 *Bathymodiolus septemdierum* complex sampling localities

Locality[a]	Abbr.	Latitude	Longitude	Depth (m)	Dive No.[b]	Samples	Preserved
Central Indian Ridge[1]	CIR						
Kairei	KA	25°19.2′S	70°02.4′E	2415–2460	J1:301	30	Frozen
Edmond	ED	23°52.7′S	69°35.8′E	3290–3320	J1:296	29	Frozen
Mariana Basin/Arc[2,3]	MA						
NW Eifuku	EF	21°29.4′N	144°2.4′E	1535	R:792–793 J2:197	7	Ethanol
Mariana Trough	MT	18°12.8′N	144°42.4′E	3589	S:140–188	6	Frozen
Vanuatu	VA						
Nifonea	NF	18°8.0′S	169°31.0′E	1900	So:229	30	Ethanol
North Fiji Basin[2,4]	NFB						
White Lady	WL	16°59.5′S	173°54.9′E	1989–1992	J2:149	30	Frozen
Lau Basin[2,4]	LB						
Kilo Moana	KM	20°3.2′S	176°8.0′W	2612–2622	J2:140–141	29	Frozen
Tow Cam	TC	20°19.1′S	176°8.3′W	2714	J2:142	28	Frozen
Tui Malila	TM	21°59.4′S	176°34.1′W	1845–1900	J2:144	29	Frozen
Hine Hina	HH	22°32.3′S	176°43.0′W	1807–1819	J2:145	23	Frozen

[a] Morphotype occurrences according to ChEssBase/GBIF, Won et al. (2008), Desbruyères et al. (2006) and Van Dover et al. (2001): (1) *marisindicus*, (2) *brevior*, (3) *septemdierum*, (4) *elongatus*

[b] Submersibles/ship: J1 = *Jason I*, J2 = *Jason II*, S = *Shinkai 6500*, R = *Ropos*, So = *RV Sonne*

Fig. 2). Depths ranged from about 1500–3600 m. Upon collection with remotely operated vehicles, samples were either preserved in 70 % ethanol or dissected and stored at −80 °C.

DNA isolation, PCR and sequencing

DNA extraction, PCR amplification, amplicon purification and sequencing methods mirrored those of Génio et al. (2008) and Johnson et al. (2013). Mussels were genotyped at four nuclear (*Cat* = catchin, *Col-1* = collagen type XIV, *EF1α* = elongation factor 1α, *H3* = histone 3) and four mitochondrial (*tRNA^Met* = transfer RNA for methionine, *tRNA^Val* = transfer RNA for valine, *ND4* = NADH dehydrogenase subunit 4, *COI* = cytochrome-c-oxidase subunit I) loci that were found to contain polymorphic sites in *B. septemdierum* (Table 2). These loci were chosen from a test suite of 10 genes and 17 primer pairs, as they were the only ones that consistently provided usable sequences. All obtained sequences were deposited in GenBank under accession numbers KP879256–KP881222 (Table 2).

Allozymes

Allozymes encoded by eleven gene loci (*PGDH* = 6-phosphogluconate dehydrogenase, *G6PDH* = glucose-6-phosphate dehydrogenase, *AAT1* = aspartate aminotransferase 1, *AAT2* = aspartate aminotransferase 2, *IDH2* = isocitrate dehydrogenase 2, *LAP1* = leucine aminopeptidase 1,

MPI = mannose-6-phosphate isomerase, *PEP-GL* = peptidase with glycyl-leucine, *GPI* = glucose-6-phosphate isomerase, *PGM* = phosphoglucomutase, *LDH* = lactate dehydrogenase) were examined from the subset of samples that were frozen immediately following their collection (Tables 1, 3). For each specimen, an approximately 0.2 g piece of adductor muscle was homogenized in a roughly equal volume of extraction buffer (0.01 M Tris, 2.5 mM EDTA, pH 7.0), and the homogenate was centrifuged at $12,000 \times g$ for 2 min to remove tissue debris. Screening employed cellulose-acetate gel-electrophoresis (CAGE) according to the conditions, buffers and stains outlined by Hebert and Beaton (1989).

Sequence analysis

Forward and reverse sequences for each sample and gene were quality trimmed, clipped and paired applying the De Novo Assemble tool with highest sensitivity in GEN-EIOUS v7.1.5 (http://www.geneious.com/). If required, base calls were corrected manually. Subsequently, consensus sequences were multiple aligned with the integrated MUSCLE program using 20 iterations. To check for the accuracy of the PCR amplifications, the resulting nucleotide alignment was compared against the nr database with MEGABLAST choosing an e-value of 1e-20 and a minimum similarity of 75 %. Nuclear alleles in heterozygous individuals were further resolved with PHASE v2.1.1 (Stephens et al. 2001; Stephens and Donnelly 2003). For
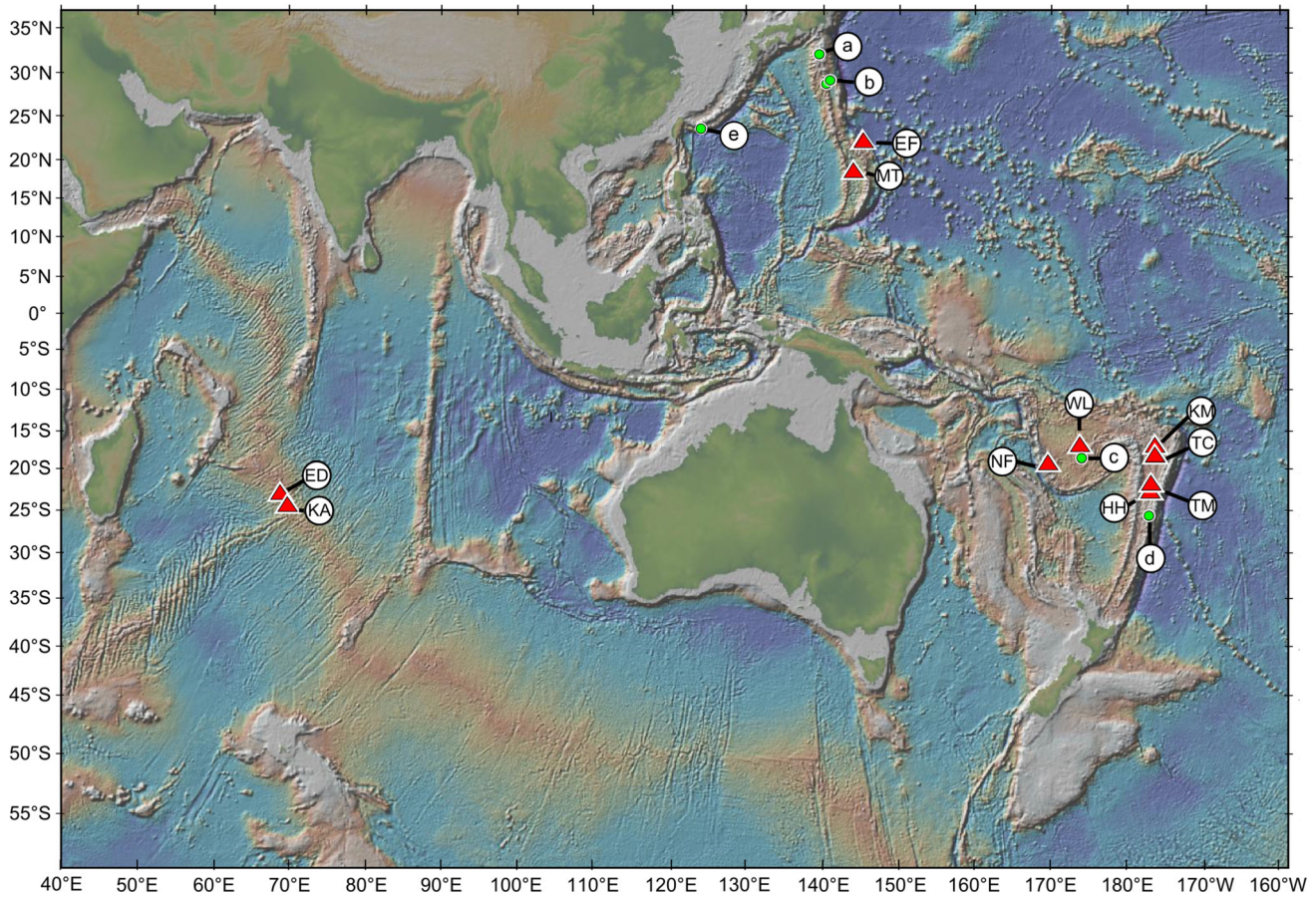
Fig. 2 Geographic distribution of the *B. septemdierum* complex in the Indo-Pacific. Sampled sites are: *KA* Kairei, *ED* Edmond, *EF* NW Eifuku, *MT* Mariana Trough, *NF* Nifonea, *WL* White Lady, *KM* Kilo Moana, *TC* Tow Cam, *TM* Tui Malila, *HH* Hine Hina. Unsampled sites with known occurrences according to ChEssBase/GBIF, Desbruyères et al. (2006), Miyazaki et al. (2010) and Tunnicliffe et al. (2009) are: *a* Mokuyo Seamount and Suiyo Seamount (Izu-Bonin Arc, morphotype *septemdierum*), *b* Myojin Knoll and Sumisu Caldera (Izu-Bonin Arc, morphotype *septemdierum*), *c* Mussel Valley (North Fiji Basin, morphotypes *brevior* and *elongatus*), *d* Monowai (Tonga Arc, morphotype *brevior*), *e* Hatoma Knoll (Okinawa Trough, morphotype *septemdierum*, only 1 specimen)

ensuring the reliability of the results, we chose 10,000 iterations of the MCMC chain, a burnin of 1000 and five different seeds for the random number generator. As $tRNA^{Met}$, $tRNA^{Val}$ and *ND4* were amplified as a continuous sequence, ARWEN v1.2 (Laslett and Canbäck 2008) was used to predict the exact tRNA gene boundaries and separate the three loci. Due to problems with PCR amplification or sequence quality, samples from EF had to be excluded from all analyses except for haplotype network generation and basic statistics.

Data preparation

For combined usage of DNA and allozymic data in Bayesian (STRUCTURE, BAYESCAN, BAYESASS) and general population genetic statistics (Hardy–Weinberg (HWE) and linkage (LE) equilibrium), sequence information for each locus was reduced to a number code by identifying and indexing the unique alleles. Data were subsequently transformed into appropriate input formats using CONVERT v1.31 (Glaubitz 2004), FORMATO-MATIC v0.8.1 (Manoukis 2007) or PGDSPIDER v2.0.7.2 (Lischer and Excoffier 2012). In all other analyses, truncated sequence data were utilized.

Molecular evolution and tests for selection

As a preparation for further analyses we assessed the nucleotide substitution scheme for the eight nuclear and mitochondrial genes in JMODELTEST v2.1.6 (Darriba et al. 2012; Guindon and Gascuel 2003) by evaluating 5 possible models with the Bayesian Information Criterion (BIC). Unlike commonly used likelihood ratio tests the BIC has the advantage of comparing both nested and non-nested models, determining model uncertainty and performing model averaging (Posada and Buckley 2004). For each partition we considered rate variation among sites (−g 4), a proportion of invariable sites (−i) and unequal base

Table 2 Sequence characteristics for the eight gene loci analyzed in this study

Locus	Primer	Sequence	Ref.	Length	Exon	Intron	Indels	H^a	π^b	GenBank Acc. No.
mtDNA										
COI	COIG	5'-GTATTGAATTAGCACGTCCTGGAA-3'	Génio et al. (2008)	583	1–583	N/A	N/A	0.8820	0.0062	KP879540–KP879754
	COIH	5'-ATACTATTCCAAACCGGGTAAAAAT-3'	Génio et al. (2008)							
ND4	ArgBL	5'-CAAGACCCTTGATTTCGGCTCA-3'	Bielawski and Gold (1996)	549	1–549	N/A	N/A	0.8580	0.0059	KP881001–KP881222
	NAP2H	5'-TGGAGCTTCTACGTGRGCTTT-3'	Arévalo et al. (1994)							
tRNA^Val	ArgBL	5'-CAAGACCCTTGATTTCGGCTCA-3'	Bielawski and Gold (1996)	71	N/A	N/A	2–3	0.4489	0.0071	KP881001–KP881222
	NAP2H	5'-TGGAGCTTCTACGTGRGCTTT-3'	Arévalo et al. (1994)							
tRNA^Met	ArgBL	5'-CAAGACCCTTGATTTCGGCTCA-3'	Bielawski and Gold (1996)	64	N/A	N/A	N/A	0.0269	0.0004	KP881001–KP881222
	NAP2H	5'-TGGAGCTTCTACGTGRGCTTT-3'	Arévalo et al. (1994)							
nDNA										
Cat	CatF	5'-GAGYGTCTTTCMAAGATCATCTCCA-3'	Johnson et al. (2013)	354	1–17	18–354	1–93	0.8411	0.0076	KP879256–KP879539
	Cat2R	5'-CATTTYCTGATGTTWCGCTGGAT-3'	Johnson et al. (2013)							
Col-1	Col160F	5'-GGTTCACGAYCGGAWGTTCCC-3'	Johnson et al. (2013)	445	1–50	51–445	59	0.6497	0.0027	KP879755–KP880118
	Col1R	5'-TCTCCTTCGCTATTTTTGTGG-3'	Faure et al. (2009)							
EF1α	EF1αF	5'-ACGCCTGGGTATTGGACAAAC-3'	Faure et al. (2009)	439	1–65	66–439	13	0.2650	0.0009	KP880119–KP880528
	EF1αR	5'-CCAAGAGGGGTCGTACAAATT-3'	Faure et al. (2009)							
H3	H3F	5'-ATGGCTCGTACCAAGCAGACV-3'	Colgan et al. (1998)	321	1–321	N/A	N/A	0.1863	0.0014	KP880529–KP881000
	H3R	5'-ATATCCTTRGGCATRATRGTG-3'	Colgan et al. (1998)							

Sequence lengths, exons, introns and indels are given in base pairs. Exon–intron boundaries are based on the position in the alignment

COI Cytochrome-c-oxidase subunit I, *ND4* NADH dehydrogenase subunit 4, *tRNA^Val* transfer RNA for valine, *tRNA^Met* transfer RNA for methionine, *Cat* Catchin, *Col-1* Collagen type XIV, *EF1α* Elongation factor 1α, *H3* Histone 3

[a] Haplotype/gene diversity

[b] Nucleotide diversity

Table 3 Allozyme loci and optimal buffers used for analyses

Enzyme	Locus	EC No.	Optimal buffers
Aspartate aminotransferase	*Aat-1, Aat-2*	2.6.1.2	TC 7.0
Glucose-6-phosphate dehydrogenase	*G6pdh*	1.1.1.49	TC 7.0
Glucose-6-phosphate isomerase	*Gpi*	5.3.1.9	TC 7.0
Isocitrate dehydrogenase	*Idh-2*	1.1.1.42	CA 6.2
Leucine aminopeptidase	*Lap-1*	3.4.11.1	TC 7.0
Mannose-6-phosphate isomerase	*Mpi*	5.3.1.8	TC 7.0
Lactate dehydrogenase	*Ldh*	1.1.1.27	TG 8.5
Peptidase (glycyl-leucine)	*Pep-gl*	3.4.11	TG 8.5
Phosphoglucomutase	*Pgm*	5.4.2.2	TC 7.0
6-Phosphogluconate dehydrogenase	*Pgdh*	1.1.1.44	TC 7.0

Compositions of each buffer (one liter total volume) were: TC 7.0: Trizma base (90.8 g), citric monohydrate (52.5 g), pH 7.0, and dilution factor for working buffer with deionized water (20×); TG 8.5: Trizma base (30 g), glycine (144 g), pH 8.5, and dilution factor (10×); CA 6.2: Citric acid monohydrate (42 g), N-(3-aminopropyl)-morpholine (50 ml), pH 6.2, and dilution factor (20×)

frequencies ($-f$). Nucleotide (π) and haplotype/gene (H) diversities were assessed with ARLEQUIN v3.5.1.2 (Excoffier and Lischer 2010), using the Tamura-Nei distance and a gamma correction according to the inferred substitution models. Sequence divergence was determined in MEGA6 (Tamura et al. 2013). We used a bootstrapped Maximum Composite Likelihood method (500 iterations), including both transitions and transversions and considering heterogeneous patterns among lineages. Substitution rates among sites were assumed to be uniform or modeled according to a gamma distribution, if suggested by JMO-DELTEST. Missing information was excluded following the pairwise deletion procedure. To validate the selective neutrality of our markers for usage in BAYESASS, we performed two types of selection tests. We used MEGA6 to conduct codon-based Z-tests for selection on protein-coding DNA sequences. However, as the amplified exonic fragments were short for *Cat*, *Col-1* and *EF1α*, these markers were excluded from this analysis. We used the original or modified Nei-Gojobori method with Jukes-Cantor correction and conducted 500 bootstrap replications for each test. Gaps and missing data were treated as described. Complementarily, we used BAYESCAN v2.1 (Foll and Gaggiotti 2008) to conduct F_{ST}-outlier tests for (pseudo-)selection on the allozyme and DNA markers. Default parameters were used for run length and prior odds (i.e., a neutral evolution model was considered as 10 times more likely than an adaptive one). To avoid detection of false positives, we excluded monomorphic loci from the analyses. A locus was considered as outlier, if it had a *q*-value <0.1.

Genetic structure

To identify the degree of partitioning between populations we applied the program STRUCTURE v2.3.4 (Pritchard et al. 2000). We chose the admixture model with correlated allele frequencies and a sample site prior (Falush et al. 2003; Hubisz et al. 2009) to infer the most likely number of genetic clusters K. The MCMC chain was simulated 10^6 times with a burnin period of 10^5. Posterior probabilities were calculated for K ranging from 1 to 11 with five replications for each value and corrected after the ΔK method by Evanno et al. (2005). Tests of Hardy–Weinberg and linkage equilibrium ($\alpha = 0.05$) within populations were done in ARLEQUIN and adjusted for type 1 error, using the Benjamini–Yekutieli False Discovery Rate (BY FDR) procedure (Benjamini and Yekutieli 2001). Compared to sequential Bonferroni correction, this method has the advantage of not losing test power while controlling for error rates in multiple comparisons (Narum 2006). Only two of the allozyme loci (*Mpi* and *Gpi*) were sufficiently polymorphic to warrant inclusion in the STRUCTURE analyses. The four mitochondrial markers (hereafter mtDNA) were concatenated into a single data string for each individual. We examined the consistency of the results by running analyses in four ways: (1) all polymorphic markers (*Cat*, *Col-1*, *EF1α*, *H3*, *Mpi*, *Gpi*, plus mtDNA); (2) the neutral nuclear markers alone (*Cat*, *EF1α*, *Mpi*, *Gpi*); (3) the DNA sequence markers alone (*Cat*, *Col-1*, *EF1α*, *H3*, plus mtDNA); and (4) the neutral nDNA markers alone (*Cat*, *EF1α*). STRUCTURE results were converted to vector graphics using the RUBY script "bar_plotter.rb" (http://evolution.unibas.ch/salzburger/software.htm) to improve image quality.

As a measure of population differentiation, pairwise F_{ST} (allozymes) and Φ_{ST} (mtDNA, nDNA) values were estimated in ARLEQUIN v3.5.1.2, using non-parametric permutation procedures with 10,000 replications. Haplotype distance matrices for Φ_{ST} estimations were calculated with the Tamura-Nei model. All tests were corrected following the BY FDR method, choosing 0.05 as α level.

Hierarchical structuring of population differentiation was tested with the Analysis of Molecular Variance (AMOVA) framework in ARLEQUIN v3.5.1.2, grouping demes according to regional affinities (W. Pacific versus Indian Ocean). Owing to missing genotypes in our datasets, covariance components and Φ-statistics were assessed for each locus separately and combined into a global weighted average. Statistical significance of the Φ estimates was evaluated by 10000 permutations.

Contemporary migration

Potential recent migration events based on non-linked, neutral polymorphic nDNA and allozyme markers (*Cat*, *EF1α*, *Mpi*, *Gpi*) were analyzed with BAYESASS v3.0—a Bayesian assignment software that can accurately detect immigrants over the last two generations if population pairwise F_{ST}s are greater than 0.05 and loci are in LE (Wilson and Rannala 2003; Faubet et al. 2007). Since genetic differentiation between individual populations was low and initial test runs with unpooled data led to inconsistent results and low ESS values, we merged samples belonging to the same arc or mid-ocean ridge area (Table S1). In addition, we performed separate calculations for the LB samples to estimate connectivity within the W. Pacific back-arc spreading centers (Table S2). Runs were iterated 10^8 times, using a sampling interval of 2000 and discarding the first 10^7 simulations. Consistency of results was evaluated by repeating calculations 10 times with different random number seeds, while convergence was checked with TRACER v1.6 (http://beast.bio.ed.ac.uk/Tracer). To achieve the recommended acceptance rates (20–60 %) for changes in mixing parameters, we set the step length for migration rates and inbreeding coefficients to 1.0 and the step length for allele frequencies to 0.6 (LB data set) or 1.0 (pooled data set). For obtaining the most reliable result, we calculated the Bayesian deviance for each run based on the R script by Meirmans (2014) and selected the output with the lowest value. The significance of the migration rate estimates for all pairs of sites was evaluated by assessing overlaps of the 95 % confidence intervals with those simulated for uninformative data.

Phylogenetic relationships

We used the MJ method in NETWORK v4.6.1.2 (www.fluxus-engineering.com; Bandelt et al. 1999; Forster et al. 1996; Saillard et al. 2000) to reconstruct phylogenies for mtDNA and nDNA sequence data. Weights of characters and mutation types were left at their defaults. All calculations were repeated three times with the epsilon parameter fixed at values of 0, 10 and 20 to validate the

consistency of the results. Finally, redundant median vectors were removed from all networks using the MP option (Polzin and Daneshmand 2003). In cases where large gaps were present in the alignments, network calculations produced unreasonable results in terms of mutation steps or homoplasies. We therefore either excluded problematic sequences (2 *Col-1* sequences) or marked indels in the network graphs (*EF1α*, *Cat*). As we did not observe any apparent sub-structuring at the population level, we pooled samples according to arc and mid-ocean ridge area to simplify the visual output.

Results

DNA sequence variation and selection regimes

DNA fragments of eight genes were examined in mussel samples from 10 locations (Table 1). Haplotype and nucleotide diversities for mitochondrial *COI* and *ND4* were among the highest observed in this study (Table 2). Although the *tRNA*^{Val} haplotypes were less variable, nucleotide polymorphism was larger than in all other mtDNA genes. Overall, *tRNA*^{Met} exhibited the lowest haplotype and nucleotide variation. Even so, the level of nucleotide polymorphism was similar to that of some nuclear genes (*EF1α*, *H3*). Phylogenetic networks generated for *COI* and *ND4* were topologically similar (Fig. 3). The most frequent haplotypes in the W. Pacific samples were entirely absent from the Indian Ocean samples. The W. Pacific and Indian Ocean regions were almost completely differentiated for *tRNA*^{Val} haplotypes, but the most common *tRNA*^{Met} haplotype dominated both regions.

In general, the protein-coding nuclear genes were less polymorphic than the protein-coding mitochondrial genes (Table 2; Fig. 3). Only *Cat* showed a haplotype diversity (*H*) that was comparable to that of *COI* and *ND4*, and its nucleotide diversity (π) was greatest of all other genes. *H3* and *EF1α* had the lowest *H* and π diversities, respectively. The dominant *H3*, *EF1α* and *Col-1* haplotypes were shared between the W. Pacific and Indian Ocean regions (Fig. 3). In contrast, the Indian Ocean samples were mostly unique for *Cat* haplotypes. A few rare and potentially private polymorphisms were observed for all genes. Except for *Cat*, the overall and net between-region distances were typically smaller for the nuclear than for the mitochondrial genes (Table S3). Overall and net between-region divergences for *Cat* were in the upper range of the mitochondrial values. Sequence divergence within the W. Pacific was usually close to zero and did not exceed 0.00018 for all genes (Table S4).

Various tests for selection differed among the loci. Codon-based Z-tests suggested purifying selection for all
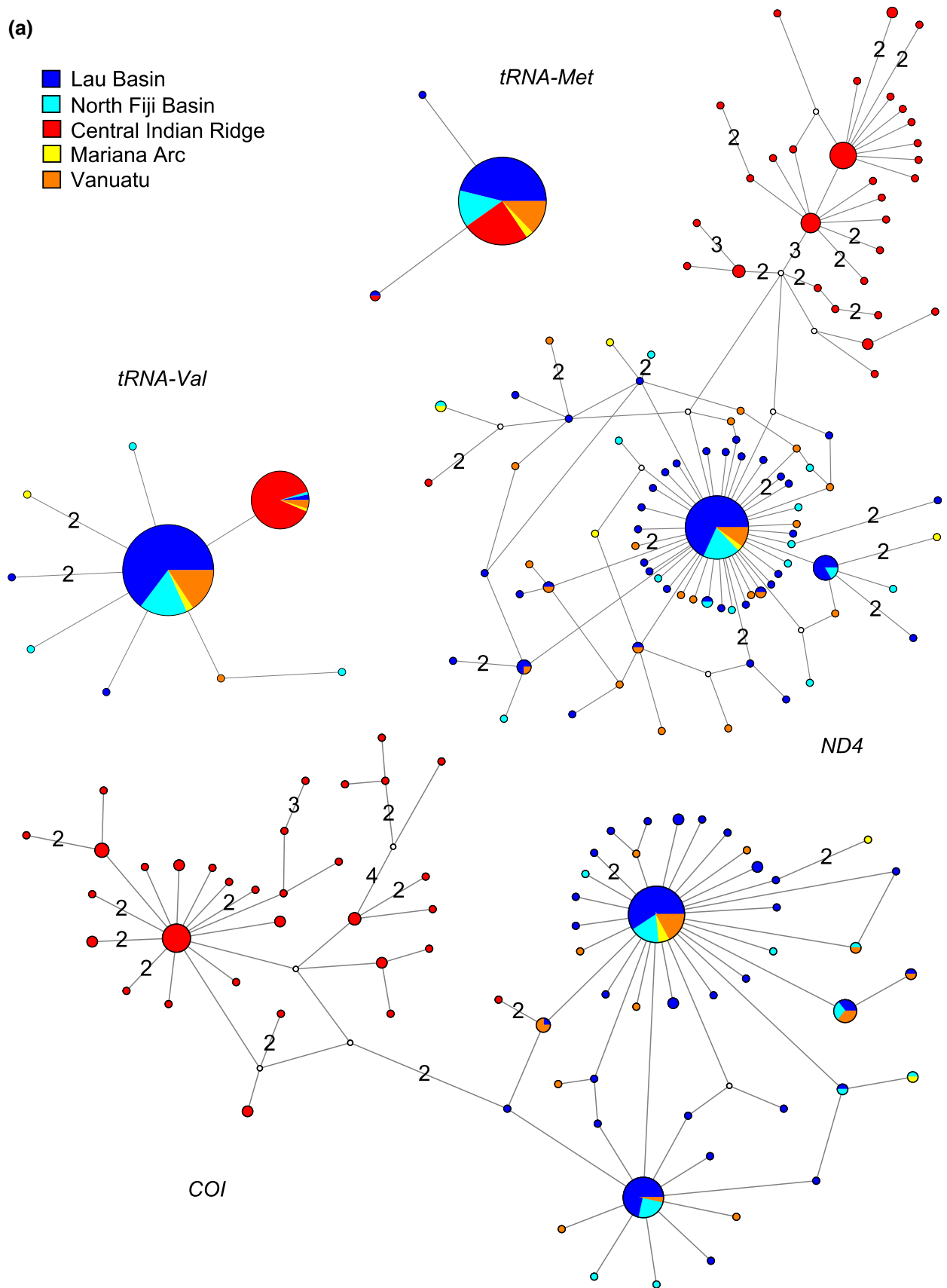
(a)

◄**Fig. 3** Phylogenetic networks for **a** mitochondrial and **b** nuclear DNA. Pie sizes for each network are proportional to the frequency of the respective gene variant (n = 1 for the *smallest circle*). *White* circles represent unknown/missing haplotypes. Numbers on branches show the number of mutations between alleles. If no numbers are indicated, only one mutation step occurred. Branch lengths are not to scale to improve visualization. *indel

investigated sequences (i.e., *COI, ND4*, and *H3*; Table S3). F_{ST}-outlier analyses were conducted with all the nuclear genes (Table S5; Fig. S1). Only *Col-1* deviated from neutral expectations, possibly indicating balancing or negative selection.
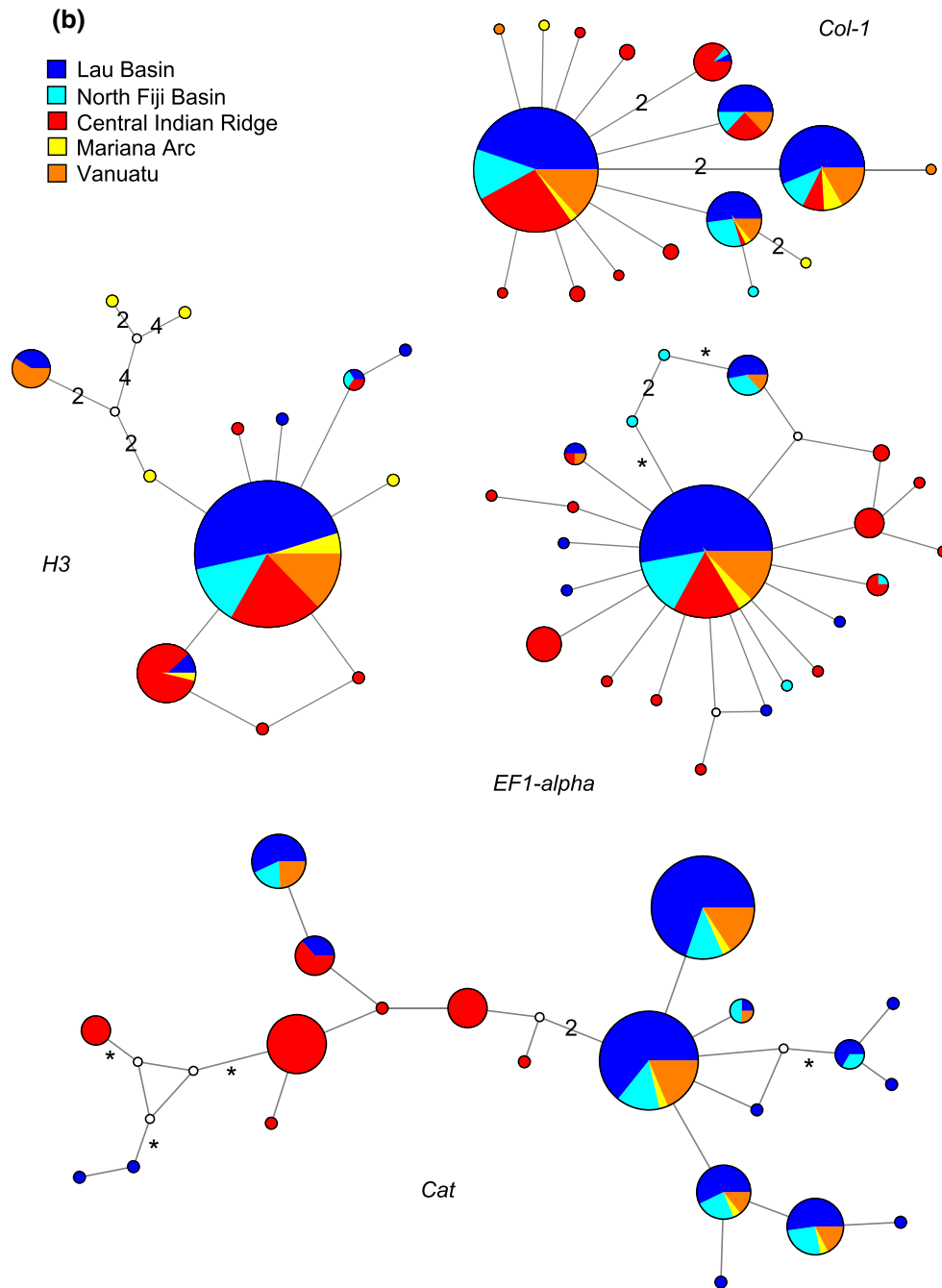

Fig. 3 continued

Allozyme variation

Allozyme diversity was scored in the 8 frozen samples (Table 1). The number of alleles was lower than for the DNA markers, ranging from 1 to 5. Because only non-synonymous charge-based polymorphisms are revealed with CAGE methods, actual variation might be substantially greater. Two loci were polymorphic in all eight samples (*Mpi* and *Gpi*), three had rare alleles ($q < 0.05$) that occurred in two or more samples (*G6pdh, Lap-1*, and *Pep-gl*), four had alleles that were "private" to a single sample (*Pgdh, Aat-1, Aat-2*; and *Idh-2*), and two were monomorphic in all samples (*Pgm* and *Ldh*). Despite some minor differences in allele frequencies, all Indian Ocean and W. Pacific samples were usually dominated by the same most common variant, regardless of the locus analyzed (Table S6). F_{ST}-outlier analyses were non-significant for all loci tested.

Hardy–Weinberg and linkage equilibrium

Tests for HWE indicated 20 significant departures from random mating expectations out of 62 possible comparisons. After BY FDR correction ($\alpha = 0.01061$) the number of significant tests decreased to 10, but deviations were not uniform across populations or loci (Table S7). In the case of LE a total of 170 performable tests resulted in 26 significant patterns of genetic linkage, which were reduced to 13 significant associations following BY FDR correction ($\alpha = 0.00875$). However, the pairing of genes was dissimilar among populations (Table S8). This lack of consistency suggests that our genetic data were free of statistical biases and that the observed departures from HWE and LE most likely resulted from scoring problems, allelic dropouts in heterozygous individuals (Piyamongkol et al. 2003) or limitations of the likelihood ratio test for samples with unknown gametic phase (Excoffier and Slatkin 1998).

Population structure

Regardless of the marker types, all STRUCTURE simulations produced qualitatively similar population clusters that distinguished samples from the Indian Ocean versus W. Pacific regions (Fig. 4, two examples). After BY FDR correction Φ_{ST} indices based on the mitochondrial loci were significant for all pairwise comparisons of the Indian and W. Pacific samples (Table 4). In contrast, within-region comparisons were not significant except for contrasts involving TM and KM ($\Phi_{ST} = 0.0824$), MT ($\Phi_{ST} = 0.2150$) and NF ($\Phi_{ST} = 0.1077$). Populations were less divergent for the nuclear genes, although all contrasts between Indian Ocean and W. Pacific samples were significant except for the KA-

HH and ED-HH pairs (Table 4). All within-region contrasts were not significant. Pairwise F_{ST}s based on the allozyme loci *Mpi* and *Gpi* were not significant with the exception of the ED-HH ($F_{ST} = 0.1548$), KM-HH ($F_{ST} = 0.1262$) and TC-HH ($F_{ST} = 0.0724$) comparisons (Table 5).

AMOVAs were conducted separately for mtDNA, nDNA and allozymes (Table 6). Most of the mtDNA variance (66.84 %) occurred between regions, a small but statistically significant amount (0.49 %) occurred among samples within regions, and 32.68 % resided within samples. Less of the nDNA variance (22.40 %) occurred between regions, an insignificant amount (0.32 %) occurred among samples within regions, and most (77.28 %) resided within samples. In contrast, nearly all of the allozyme variation (97.35 %) resided within samples, no significant variation (−0.28 %) occurred between regions, and little existed among samples within regions (2.93 %).

Migration

BAYESASS outputs implied a strong self-seeding rate (≥0.6828; fraction of individuals migrating from source vent per generation) (Table S9; Table S10). On a broader scale, limited immigration was only seen from LB to NFB (0.2330) and to VA (0.2709). By contrast, no significant immigration was detected for the LB localities. As populations were poorly differentiated within the W. Pacific, we cannot exclude that some migration events have been overlooked or that a fraction of immigrants were wrongly assigned due to limitations of the BAYESASS algorithm (Faubet et al. 2007; Meirmans 2014). Additionally, the sample size for LB was higher than for the other demes, so that detection of contemporary emigration events from this site might have been facilitated. Nevertheless, all 10 of the BAYESASS replicates produced the same outcomes and varied only slightly in the Bayesian deviance (pooled data set: 4884.783–4888.187; LB data set: 1927.533–1928.588). We experienced no problems with convergence and we achieved high effective sample sizes (ESS > 200) for all parameters. This indicates that most of the estimates should be reasonably accurate.

Discussion

Analyses of genetic variation within the *Bathymodiolus septemdierum* complex confirmed earlier findings of low differentiation among the morphospecies and populations (Kyuno et al. 2009). Nonetheless, assignment procedures partitioned the multilocus genotypes into two distinct regional metapopulations: (1) all sampled W. Pacific localities including the nominal species *B. septemdierum*, *B. brevior* and *B. elongatus*; versus (2) the Indian Ocean samples comprising the

Fig. 4 STRUCTURE analysis. *Bar* plots showing the clustering of individuals based on K = 2 and **a** all polymorphic markers (*Cat*, *Col-1*, *EF1α*, *H3*, *Mpi*, *Gpi*, concatenated mtDNA) and **b** only neutral polymorphic markers (*Cat*, *EF1α*, *Mpi*, *Gpi*). Each *vertical line* represents one mussel sampled at the respective location, where numbers on the *left* indicate the genetic content an individual inherits from each cluster. *KA* Kairei, *ED* Edmond, *KM* Kilo Moana, *TC* Tow Cam, *TM* Tui Malila, *HH* Hine Hina, *WL* White Lady, *MT* Mariana Trough, *NF* Nifonea
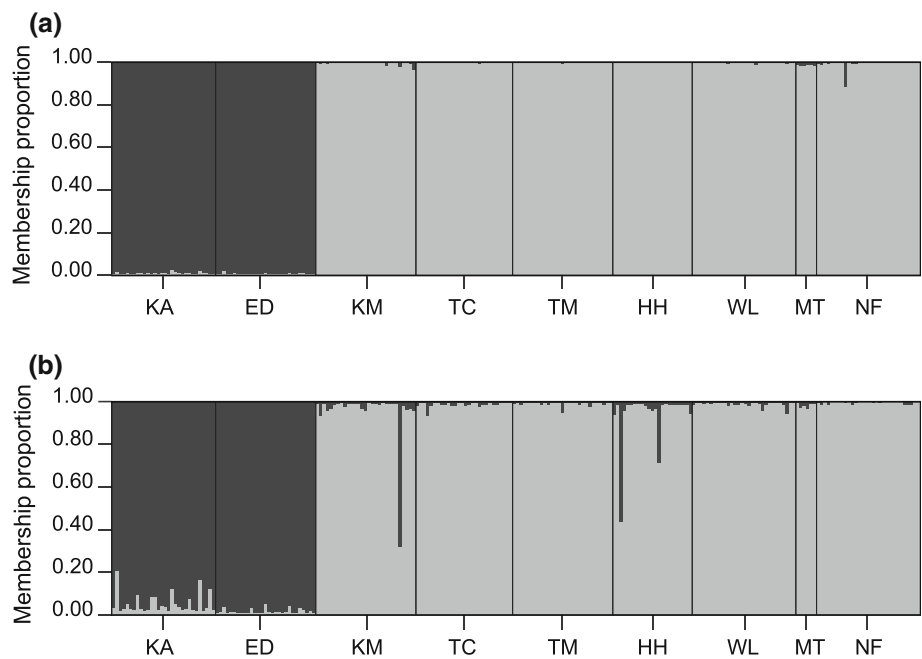

Table 4 Pairwise $\Phi_{ST}s$ for mtDNA (above diagonal) and nDNA (below diagonal)

	KA	ED	KM	TC	TM	HH	WL	MT	NF
KA	*	0.01502	**0.65473**	**0.66734**	**0.66641**	**0.64439**	**0.65503**	**0.59349**	**0.63770**
ED	−0.02380	*	**0.57333**	**0.58706**	**0.58383**	**0.54831**	**0.57297**	**0.48429**	**0.55500**
KM	**0.25756**	**0.21985**	*	0.00091	**0.08238**	−0.02430	−0.00013	0.04091	−0.00388
TC	**0.31339**	**0.28062**	−0.02909	*	0.03609	−0.05480	−0.01956	0.07785	0.00586
TM	**0.30667**	**0.24914**	0.00048	−0.04595	*	0.02564	0.02748	**0.21500**	**0.10767**
HH	0.04646	−0.10351	0.04270	−0.08689	0.10600	*	0.00089	0.10494	0.01792
WL	**0.24698**	**0.20816**	−0.01156	0.01611	−0.00217	−0.01906	*	0.06220	0.01787
MT	**0.30366**	**0.30933**	0.00785	0.03818	0.02990	−0.10352	0.03231	*	0.01465
NF	**0.17413**	**0.13559**	0.03839	0.05039	0.04078	−0.03982	0.02506	0.04588	*

Bold values indicate significant differences after BY FDR correction

KA Kairei, *ED* Edmond, *KM* Kilo Moana, *TC* Tow Cam, *TM* Tui Malila, *HH* Hine Hina, *WL* White Lady, *MT* Mariana Trough, *NF* Nifonea

Table 5 Pairwise $F_{ST}s$ for the two most polymorphic allozyme loci *Mpi* and *Gpi*

	KA	ED	KM	TC	TM	HH	WL	MT
KA	*							
ED	0.03161	*						
KM	0.01263	−0.00092	*					
TC	−0.00928	0.00887	0.00633	*				
TM	−0.01326	0.04416	0.01900	−0.00693	*			
HH	0.04469	**0.15475**	**0.12619**	**0.07243**	0.05062	*		
WL	−0.00209	0.06301	0.04272	0.01261	0.00415	0.00286	*	
MT	0.00812	0.12836	0.07762	0.03842	0.00742	−0.03710	−0.01861	*

Bold values indicate significant differences after BY FDR correction

KA Kairei, *ED* Edmond, *KM* Kilo Moana, *TC* Tow Cam, *TM* Tui Malila, *HH* Hine Hina, *WL* White Lady, *MT* Mariana Trough, *NF* Nifonea

Table 6 AMOVA results for the hierarchical grouping of 9 sampled sites into ocean region

	mtDNA		nDNA		Allozymes (*Mpi*, *Gpi*)	
	% variation	Φ-statistics	% variation	Φ-statistics	% variation	Φ-statistics
Among regions	66.83515	0.66835***	22.40400	0.22404***	−0.28305	−0.00283
Among sites within region	0.48682	0.01468***	0.31669	0.00408	2.92871	0.02920**
Within sites	32.67803	0.67322***	77.27931	0.22721***	97.35434	0.02646**

W. Pacific region MT, NF, WL, KM, TC, TM, and HH, *Indian Ocean region* KA and ED

** $p < 0.01$, *** $p < 0.001$

nominal species *B. marisindicus.* Despite overall low levels of sequence divergence, *Cat* alleles and mitochondrial haplotypes for *ND4*, *COI* and *tRNA*^{Val} were almost completely distinct between the two regions. On the other hand, gene variants at the other nDNA and allozymic loci as well as *tRNA*^{Met} were broadly shared between the Indian and W. Pacific samples. These discordances in mitochondrial and nuclear differentiation were reflected by the fixation indices based on AMOVA Φ-statistics (Φ_{CT} (mtDNA) = 0.6684; Φ_{CT} (nDNA) = 0.2240; F_{CT} (allozymes) = –0.0028). Such differences among genetic marker systems are fairly typical, resulting in part from the inverse relationship between effective population size (2–4 times smaller for mtDNA) and rates of lineage sorting (reviewed in Toews and Brelsford 2012), from synonymous substitutions affecting DNA sequences versus the inability to see them with allozyme electrophoresis, and from discrepancies in selective constraints experienced by these markers.

For example, the *H3* exon showed evidence of purifying selection according to the codon-based Z-test, but the F_{ST}-outlier test failed to detect selection, although the populations were almost invariant. Perhaps the same few alleles provided a universal advantage, which is in concordance with the essential function of *H3* in DNA packaging and the slower mutation rate (= lower allelic diversity) of conserved genes (Brown et al. 1979). In contrast, differentiation of the allozymes and *EF1α* did not deviate from neutral expectations. The broad Indo-Pacific distributions of the most frequent alleles might have resulted from incomplete lineage sorting—a phenomenon that is especially pronounced in recently separated populations (Pamilo and Nei 1988; Degnan and Rosenberg 2009). F_{ST}-outlier tests suggested possibly balancing selection for locus *Col-1*, which might point to genetic hitchhiking scenarios (Barton 2000). Balancing selection at a closely linked locus could eliminate differences between populations while maintaining *Col-1* diversity. In contrast, *Cat* polymorphisms appeared to be selectively neutral and almost fixed between the Indian Ocean and W. Pacific samples, maybe due to a founder event resulting from colonization of W. Pacific vents by Indian Ocean emigrants (Desbruyères et al. 2006; Kyuno et al. 2009). Variation at

the two protein-coding mitochondrial markers, *ND4* and *COI*, appeared to be shaped by negative selective forces, as expected for a rapidly evolving clonal genome that is susceptible to an accumulation of slightly deleterious mutations (i.e., Muller's ratchet; reviewed in Stewart and Larsson 2014). Although purifying selection appears to be contradictory to the observed levels of polymorphism, mtDNA variation might be maintained due to increased mutation rates (Brown et al. 1979), co-evolution with the nuclear genome (Dowling et al. 2007) or doubly uniparental inheritance, which is known to occur in some mytilid species (Zouros et al. 1994; Skibinski et al. 1994; Hoeh et al. 1996; Zouros 2013). Interestingly, the mitochondrial tRNAs showed contrasting results concerning the Indo-Pacific divergence. While variation at both loci is likely to be influenced by close linkage to *ND4*, changes in the *tRNA*^{Met} gene might be constrained by its essential role in translation initiation.

Mussels from vents and seeps show a variety of distribution patterns from highly restricted to the very broad range shown by *B. septemdierum* in this study. The known occurrences are indicated on Fig. 2: from Monowai Volcano in the southern Tonga Arc to Mokuyo Seamount in the Izu-Bonin Arc, a distance of 7500 km; from Monowai to KA the range is 9600 km. These ranges are the largest known among vent animals. Such a broad distribution has some precedent in the study of Olu et al. (2010) who recognize pan-Atlantic populations of two seep mussel species. *Bathymodiolus* is known to have species with long-lived larvae that may spend over a year in a teleplanic mode (Arellano and Young 2009), which, in the case of *B. septemdierum*, may be able to maintain an extended metapopulation in the Pacific through larval exchange. Furthermore, Arellano et al. (2014) show that these larvae perform ontogenetic vertical migrations to the ocean surface where faster currents would transport them over longer distances. While larval movement is necessary to connect populations, it is not sufficient. Habitat must be suitable and available. The wide depth range recorded for this species (1100 m to nearly 3500 m) suggests broad adaptation. But there are many vent sites in the Pacific where other species of *Bathymodiolus* may exclude *B.*

septemdierum, for example, *B. manusensis* in the Manus Basin (MB) and *B. platifrons* in the Okinawa Trough.

Community similarities between the Indian Ocean and W. Pacific vents (Van Dover et al. 2001) indicate historical connectivity either north of New Guinea (Hessler and Lonsdale 1991) or from the South East Indian Ridge via the Macquarie Ridge Complex and Kermadec Arc to the W. Pacific. Alternatively, gene flow between Indian and Pacific systems is likely active via the Indonesian Flow-through and supported by venting on subsea volcanoes in the Molucca and Banda Seas (e.g., McConachy et al. 2004). While these connections would generally support the low genetic divergence found in our study, sea level drops during past glacial periods might have temporarily isolated Indian and Pacific populations (Barber et al. 2000), which could have contributed to the differences seen in mitochondrial and *Cat* variants. Despite possibly long planktonic larval periods, the results from our BAYESASS analysis suggested a high degree of isolation between all investigated vent regions and sites, at least over the last two generations. Given the shallow population structure in the *B. septemdierum* complex, these patterns imply that contemporary gene flow exists, but that it occurs over a longer time frame in an indirect way. This hypothesis would agree with a recent modeling study of coral larval drift pathways (Wood et al. 2014), which shows that present connectivity in the Indo-Pacific area can be achieved in a stepping stone fashion. Consequently, it is plausible that the opportunistic timing of metamorphosis, the availability of sufficient intermediate habitats and an effective establishment at new settlement sites rather than long-distance dispersal *per se* are the driving forces for the broad biogeography of this species.

Studies of other Indo-Pacific vent taxa suggest that contemporary gene flow is limited in the W. Pacific back-arcs for some species and that migration routes are often asymmetric. For example, hidden genetic diversity and species crypticism was detected in gastropods of the genus *Alviniconcha*, where either two or three species were found to co-occur in NFB (*A. kojimai*, *A. boucheti*), MB (*A. kojimai*, *A. boucheti*) and LB (*A. kojimai*, *A. boucheti*, *A. strummeri*) (Johnson et al. 2014). Populations of the snail *Ifremeria nautilei* were undifferentiated within and between NFB and LB, but genetically distinct from MB sites based on mitochondrial and microsatellite markers (Thaler et al. 2011). As with our findings, the authors observed weak asymmetric movement from LB to NFB, in accordance with prevailing ocean currents. We also identified limited gene flow from LB to VA, suggesting that LB might function as a source population for *Bathymodiolus* and *Ifremeria* metapopulations in other W. Pacific back-arc and arc volcano settings. This hypothesis may be relevant to potential impacts from mining of seafloor massive

sulfides in the South Pacific. Nonetheless, contrasting outcomes of gene flow studies in other co-distributed species highlight the need for comprehensive genetic and life history investigations of vent taxa and their likely dispersal routes (reviewed in Vrijenhoek 2010). Vent metacommunity dynamics will influence behavior of both reserves and mined areas after anthropogenic disturbances. Making useful recommendations for sustainable mining plans would almost certainly be illusory, if data are based on a limited number of species. Thus, even if our results suggest that sites in the W. Pacific back-arc basins have high gene flow on a longer time scale, it will be necessary to determine settlement success of foreign larvae, characterize source, sink and intermediate stepping stone sites and measure actual recovery rates of already perturbed vent habitats to assess the impacts of mining activities and design management units for biodiversity conservation (Boschen et al. 2013).

A drawback of our study stems from sampling only one locality per back-arc basin in most cases; so, true levels of within-basin differentiation are uncertain. Consequently, sampling of multiple populations at a finer scale will be required to detect cryptic population structure—a phenomenon that is becoming increasingly apparent in hydrothermal systems (e.g., Kojima et al. 2001; Vrijenhoek 2009; Shank and Halanych 2007; Johnson et al. 2014; Puetz 2014). Multilocus genotyping approaches tend to be more informative than morphometric or unigenic methods, but they are unable to elucidate the whole picture of genetic subdivision. Next-generation technologies like RAD sequencing provide a promising and cost-effective solution to this obstacle by uncovering restriction site associated SNPs across the entire genome without the need for completely sequencing it (Davey et al. 2011; Wang et al. 2012; Reitzel et al. 2013). Future studies could exploit this method to reveal genomic differentiation, unambiguously assigning natal origins of recruiting larvae, and identifying genes involved in speciation (e.g., Puebla et al. 2014), thereby helping to understand the molecular mechanisms that underlie biodiversity dynamics at hydrothermal vents and improving attempts to mitigate impacts of deep-sea mining in these unique ecosystems.

Conclusions

The present analysis identified recently separated regional metapopulations: mussels on the Central Indian Ridge (comprising the morphotypic species *B. marisindicus*) and mussels in the W. Pacific back-arc spreading centers (comprising the morphotypic species *B. septemdierum*, *B. elongatus* and *B. brevior*). Discordant patterns of gene flow inferred from the genetic markers probably reflect the

different evolutionary histories of the investigated loci. Importantly, the shared genetic variation between populations seemed to be a result of past and infrequent or indirect contemporary migration events (stepping stone connectivity), as virtually no evidence for immigration was detected over the last two generations.

Our results have significant ramifications for biodiversity conservation in the Indo-Pacific region, where deep-sea mining projects are currently ongoing and exploration is rapidly increasing. Additional studies using highly informative markers and targeting a variety of taxa with different life-history characteristics are necessary to assess the degree of present gene flow on a genome-wide, multi-species level. Detailed knowledge about repopulation rates, intermediate settlement sites and other mechanisms influencing connectivity will further be crucial for designing efficient management plans for the protection of hydrothermal vent ecosystems.

Conflict of interest The authors declare that they have no conflict of interest.

References

Arellano SM, Young CM (2009) Spawning, development, and the duration of larval life in a deep-sea cold-seep mussel. Biol Bull 216:149–162

Arellano SM, Van Gaest AL, Johnson SB, Vrijenhoek RC, Young CM (2014) Larvae from deep-sea methane seeps disperse in surface waters. Proc Biol Sci. doi:10.1098/rspb.2013.3276

Arévalo E, Davis SK, Sites JW (1994) Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in Central Mexico. Syst Biol 43:387–418

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Barber PH, Palumbi SR, Erdmann MV, Moosa MK (2000) Biogeography: a marine Wallace's line? Nature 406:692–693

Barton NH (2000) Genetic hitchhiking. Philos Trans R Soc Lond B Biol Sci 355:1553–1562

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188

Bielawski JP, Gold JR (1996) Unequal synonymous substitution rates within and between two protein-coding mitochondrial genes. Mol Biol Evol 13:889–892

Boschen RE, Rowden AA, Clark MR, Gardner JPA (2013) Mining of deep-sea seafloor massive sulfides: a review of the deposits, their benthic communities, impacts from mining, regulatory frameworks and management strategies. Ocean Coast Manage 84:54–67

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76:1967–1971

Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J, Cassis G, Gray MR (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Aust J Zool 46:419–437

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499–510

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol Evol 24:332–340

Desbruyères D, Hashimoto J, Fabri M-C (2006) Composition and biogeography of hydrothermal vent communities in western Pacific back-arc basins. In: Christie DM, Fisher CR, Lee S-M, Givens S (eds) Back-arc spreading systems: geological, biological, chemical, and physical interactions. American Geophysical Union, Washington, DC, pp 215–234

Dowling DK, Friberg U, Hailer F, Arnqvist G (2007) Intergenomic epistasis for fitness: within-population interactions between cytoplasmic and nuclear genes in *Drosophila melanogaster*. Genetics 175:235–244

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

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

Excoffier L, Slatkin M (1998) Incorporating genotypes of relatives into a test of linkage disequilibrium. Am J Hum Genet 62:171–180

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. Mol Ecol 16:1149–1166

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

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977–993

Forster P, Harding R, Torroni A, Bandelt HJ (1996) Origin and evolution of native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945

Génio L, Johnson SB, Vrijenhoek RC, Cunha MR, Tyler PA, Kiel S, Little CTS (2008) New record of "*Bathymodiolus*" *mauritanicus* Cosel 2002 from the Gulf of Cadiz (NE Atlantic) mud volcanoes. J Shellfish Res 27:53–61

Glaubitz JC (2004) CONVERT: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Mol Ecol Notes 4:309–310

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Hashimoto J (2001) A new species of *Bathymodiolus* (Bivalvia: Mytilidae) from hydrothermal vent communities in the Indian Ocean. Venus Jpn J Malacol 60:141–149

Hashimoto J, Okutani T (1994) Four new mytilid mussels associated with deep-sea chemosynthetic communities around Japan. Venus Jpn J Malacol 53:61–83

Hebert PDN, Beaton MJ (1989) Methodologies for allozyme analysis using cellulose acetate electrophoresis. Helena Laboratories, Beaumont

Hessler RR, Lonsdale PF (1991) Biogeography of Mariana Trough hydrothermal vent communities. Deep Sea Res A 38:185–199

Hoeh WR, Stewart DT, Sutherland BW, Zouros E (1996) Cytochrome c oxidase sequence comparisons suggest an unusually high rate of mitochondrial DNA evolution in *Mytilus* (Mollusca: Bivalvia). Mol Biol Evol 13:418–421

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Johnson SB, Won YJ, Harvey JBJ, Vrijenhoek RC (2013) A hybrid zone between *Bathymodiolus* mussel lineages from eastern Pacific hydrothermal vents. BMC Evol Biol 13:21

Johnson SB, Warén A, Tunnicliffe V, Dover CV, Wheat CG, Schultz TF, Vrijenhoek RC (2014) Molecular taxonomy and naming of five cryptic species of *Alviniconcha* snails (Gastropoda: Abyssochrysoidea) from hydrothermal vents. Syst Biodivers 2014:1–18

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

Kojima S, Segawa R, Fijiwara Y, Fujikura K, Ohta S, Hashimoto J (2001) Phylogeny of hydrothermal-vent-endemic gastropods *Alviniconcha* spp. from the western Pacific revealed by mitochondrial DNA sequences. Biol Bull 200:298–304

Kyuno A, Shintaku M, Fujita Y, Matsumoto H, Utsumi M, Watanabe H, Fujiwara Y, Miyazaki J-I (2009) Dispersal and differentiation of deep-sea mussels of the genus *Bathymodiolus* (Mytilidae Bathymodiolinae). J Mar Biol. doi:10.1155/2009/625672625672:1-15

Laslett D, Canbäck B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24:172–175

Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28:298–299

Lorion J, Kiel S, Faure B, Kawato M, Ho SY, Marshall B, Tsuchida S, Miyazaki J, Fujiwara Y (2013) Adaptive radiation of chemosymbiotic deep-sea mussels. Proc Roy Soc B 280:20131243

Manoukis NC (2007) FORMATOMATIC: a program for converting diploid allelic data between common formats for population genetic analysis. Mol Ecol Notes 7:592–593

McConachy TF, Permana H, Binns RA, Zulkarnain I, Parr JM, Yeats CJ, Hananto ND, Priadi B, Burhanuddin S, Utomo EP (2004) Recent investigations of submarine hydrothermal activity in Indonesia. PACRIM 2004 Congr AusIMM Publ Ser 5:161–172

Meirmans PG (2014) Nonconvergence in Bayesian estimation of migration rates. Mol Ecol Resour 14:726–733

Miyazaki J-I, Shintaku M, Kyuno A, Fujiwara Y, Hashimoto J, Iwasaki H (2004) Phylogenetic relationships of deep-sea mussels of the genus *Bathymodiolus* (Bivalvia : Mytilidae). Mar Biol 144:527–535

Miyazaki J-I, Martins LdO, Fujita Y, Matsumoto H, Fujiwara Y (2010) Evolutionary process of deep-sea *Bathymodiolus* mussels. PLoS One 5:e10363

Narum SR (2006) Beyond Bonferroni: less conservative analyses for conservation genetics. Conserv Genet 7:783–787

Olu K, Cordes EE, Fisher CR, Brooks JM, Sibuet M, Desbruyères D (2010) Biogeography and potential exchanges among the Atlantic Equatorial Belt cold-seep faunas. PLoS One 5:e11967

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. Mol Biol Evol 5:568–583

Petersen JM, Zielinski FU, Pape T, Seifert R, Moraru C, Amann R, Hourdez S, Girguis PR, Wankel SD, Barbe V, Pelletier E, Fink D, Borowski C, Bach W, Dubilier N (2011) Hydrogen is an energy source for hydrothermal vent symbioses. Nature 476:176–180

Piyamongkol W, Bermúdez MG, Harper JC, Wells D (2003) Detailed investigation of factors influencing amplification efficiency and allele drop-out in single cell PCR: implications for preimplantation genetic diagnosis. Mol Hum Reprod 9:411–420

Polzin T, Daneshmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett 31:12–20

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information Criterion and Bayesian approaches over likelihood ratio tests. Syst Biol 53:793–808

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Puebla O, Bermingham E, McMillan WO (2014) Genomic atolls of differentiation in coral reef fishes (*Hypoplectrus* spp., Serranidae). Mol Ecol 23:5291–5303

Puetz L (2014) Connectivity within a metapopulation of the ecologically significant species, *Ridgeia piscesae* Jones (Annelida, Siboglinidae), from the Endeavour Hydrothermal Vents Marine Protected Area on Juan de Fuca Ridge. Master Thesis, University of Victoria

Ramirez-Llodra E, Brandt A, Danovaro R, De Mol B, Escobar E, German CR, Levin LA, Arbizu PM, Menot L, Buhl-Mortensen P, Narayanaswamy BE, Smith CR, Tittensor DP, Tyler PA, Vanreusel A, Vecchione M (2010) Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. Biogeosciences 7:2851–2899

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. Mol Ecol 22:2953–2970

Saillard J, Forster P, Lynnerup N, Bandelt HJ, Norby S (2000) mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. Am J Hum Genet 67:718–726

Sen A, Podowski EL, Becker EL, Shearer EA, Gartman A, Yücel M, Hourdez S, Luther GW III, Fisher CR (2014) Community succession in hydrothermal vent habitats of the Eastern Lau Spreading Center and Valu Fa Ridge, Tonga. Limnol Oceanogr 59:1510–1528

Shank TM, Halanych KM (2007) Toward a mechanistic understanding of larval dispersal: insights from genomic fingerprinting of the deep-sea hydrothermal vent tubeworm *Riftia pachyptila*. Mar Ecol 28:25–35

Skibinski DO, Gallagher C, Beynon CM (1994) Mitochondrial DNA inheritance. Nature 368:817–818

SPC (2013) In: Baker E, Beaudoin Y (eds) Deep Sea Minerals: Sea-Floor Massive Sulphides, a physical, biological, environmental, and technical review, vol 1A. Secretariat of the Pacific Community, Noumea

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162–1169

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Stewart JB, Larsson NG (2014) Keeping mtDNA in shape between generations. PLoS Genet 10:e1004670

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 30:2725–2729

Thaler AD, Zelnio K, Saleu W, Schultz TF, Carlsson J, Cunningham C, Vrijenhoek RC, Van Dover CL (2011) The spatial scale of genetic subdivision in populations of *Ifremeria nautilei*, a hydrothermal-vent gastropod from the southwest Pacific. BMC Evol Biol 11:372

Thubaut J, Puillandre N, Faure B, Cruaud C, Samadi S (2013) The contrasted evolutionary fates of deep-sea chemosynthetic mussels (Bivalvia, Bathymodiolinae). Ecol Evol 3:4748–4766

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol 21:3907–3930

Tunnicliffe V, Davies KTA, Butterfield DA, Embley RW, Rose JM, Chadwick WW Jr (2009) Survival of mussels in extremely acidic waters on a submarine volcano. Nat Geosci 2:344–348

Van Dover CL (2011) Tighten regulations on deep-sea mining. Nature 470:31–33

Van Dover CL, Humphris SE, Fornari D, Cavanaugh CM, Collier R, Goffredi SK, Hashimoto J, Lilley MD, Reysenbach AL, Shank TM, Von Damm KL, Banta A, Gallant RM, Gotz D, Green D, Hall J, Harmer TL, Hurtado LA, Johnson P, McKiness ZP, Meredith C, Olson E, Pan IL, Turnipseed M, Won Y, Young CR 3rd, Vrijenhoek RC (2001) Biogeography and ecological setting of Indian Ocean hydrothermal vents. Science 294:818–823

Von Cosel R, Métivier B, Hashimoto J (1994) Three new species of *Bathymodiolus* (Bivalvia: Mytilidae) from hydrothermal vents in the Lau Basin and the North Fiji Basin, western Pacific, and the Snake Pit area, Mid-Atlantic Ridge. Veliger 37:374–392

Vrijenhoek RC (2009) Cryptic species, phenotypic plasticity, and complex life histories: assessing deep-sea faunal diversity with molecular markers. Deep Sea Res II 56:1713–1723

Vrijenhoek RC (2010) Genetic diversity and connectivity of deep-sea hydrothermal vent metapopulations. Mol Ecol 19:4391–4411

Wang S, Meyer E, McKay JK, Matz MV (2012) 2b-RAD: a simple and flexible method for genome-wide genotyping. Nat Methods 9:808–810

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191

Won YJ, Jones WJ, Vrijenhoek RC (2008) Absence of cospeciation between deep-sea mytilids and their thiotrophic endosymbionts. J Shellfish Res 27:129–138

Wood S, Paris CB, Ridgwell A, Hendy EJ (2014) Modelling dispersal and connectivity of broadcast spawning corals at the global scale. Global Ecol Biogeogr 23:1–11

Zouros E (2013) Biparental inheritance through uniparental transmission: the Doubly Uniparental Inheritance (DUI) of mitochondrial DNA. Evol Biol 40:1–31

Zouros E, Ball AO, Saavedra C, Freeman KR (1994) Mitochondrial DNA inheritance. Nature 368:818