



**HAL**  
open science

# Phylogenetic divergence and population genetics of the hydrothermal vent annelid genus *Hesiolyra* along the East Pacific Rise: Reappraisal using multi-locus data

Leyi Xi, Yanan Sun, Ting Xu, Zhi Wang, Man Ying Chiu, Sophie Plouviez, Didier Jollivet, Jian-wen Qiu

## ► To cite this version:

Leyi Xi, Yanan Sun, Ting Xu, Zhi Wang, Man Ying Chiu, et al.. Phylogenetic divergence and population genetics of the hydrothermal vent annelid genus *Hesiolyra* along the East Pacific Rise: Reappraisal using multi-locus data. *Diversity and Distributions*, 2023, 29 (1), pp.184-198. 10.1111/ddi.13653 . hal-04283224

**HAL Id: hal-04283224**

**<https://hal.science/hal-04283224>**

Submitted on 13 Nov 2023



**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open licence - etalab

# Phylogenetic divergence and population genetics of the hydrothermal vent annelid genus *Hesiolyra* along the East Pacific Rise: Reappraisal using multi-locus data

Leyi Xi<sup>1</sup> | Yanan Sun<sup>1,2,3</sup>  | Ting Xu<sup>2,3</sup>  | Zhi Wang<sup>4</sup> | Man Ying Chiu<sup>1</sup> | Sophie Plouviez<sup>5</sup> | Didier Jollivet<sup>6</sup> | Jian-Wen Qiu<sup>1,2</sup>

<sup>1</sup>Department of Biology, Hong Kong Baptist University, Hong Kong, China

<sup>2</sup>Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China

<sup>3</sup>Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong, China

<sup>4</sup>State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China

<sup>5</sup>Department of Biology, University of Louisiana at Lafayette, Lafayette, Louisiana, USA

<sup>6</sup>Sorbonne Université-CNRS, UMR 7144 Adaptation et Diversité en Milieu Marin, Equipe DyDiv, Station Biologique de Roscoff, Roscoff, France

## Correspondence

Yanan Sun and Jian-Wen Qiu, Department of Biology, Hong Kong Baptist University, 224 Waterloo Road, Kowloon Town, Hong Kong, China.

Email: [sunyanan8733@gmail.com](mailto:sunyanan8733@gmail.com) and [qiujiw@hkbu.edu.hk](mailto:qiujiw@hkbu.edu.hk)

## Funding information

General Research Fund of Hong Kong's Research Grants Council, Grant/Award Number: 12101021; Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Grant/Award Number: SMSEGL20Sc02

Editor: Trishna Dutta

## Abstract

**Aim:** Maintaining genetic connectivity is crucial for species that inhabit the disjunct and unstable deep-sea hydrothermal vents. We aimed to re-assess the connectivity of the annelid genus *Hesiolyra* distributed at hydrothermal vents along the East Pacific Rise (EPR). A previous study detected a major clade among five *Hesiolyra* populations spanning from 13°N to 21°S and a minor sympatric southern clade with ~1% divergence from the major clade. However, this study was based on a short locus of 366-bp *COI* gene, which might not contain sufficient informative sites.

**Location:** East Pacific Rise.

**Methods:** We sequenced 188 specimens of *Hesiolyra* from five hydrothermal vent fields along the East Pacific Rise for six mitochondrial and two nuclear loci to infer their genetic divergence, population diversity and gene flow.

**Results:** We found a minor southern clade (*Hesiolyra* aff. *bergi*) which was genetically distinct from *Hesiolyra bergi* sensu stricto for all gene markers explored except 16S. We also found shared genotypes between *H. bergi* and *H. aff. bergi* likely resulted from incomplete lineage sorting. For *H. bergi* s.s., we found a low but fixed divergence between the north and south EPR populations. We estimated the northern and southern metapopulations of *H. bergi* s.s. split ~0.45 Mya (HPD: 0.27–0.74 Mya). The northern metapopulation had a higher haplotype diversity than the southern metapopulation, indicating historical gene flow and loss of genetic diversity in the southern clade.

**Main conclusions:** We confirmed that the equatorial dispersal filter observed previously on several other vent species also applies to *H. bergi*. Our results highlight the power of the multi-locus approach in revealing the divergence and population genetic history of marine species with strong dispersal capabilities, indicating that the NEPR and the SEPR should be considered as separate biogeographic regions in environmental management and biological conservation.

## KEYWORDS

deep sea, gene flow, hydrothermal vent, Mid-ocean ridge, population genetics

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

© 2022 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Deep-sea hydrothermal vents are distributed along mid-ocean ridges and back-arc spreading centres and provide habitats that support high biomass of epibenthos that are distinct from the surrounding seabed (Van Dover et al., 2002). Animals that inhabit these reducing ecosystems such as giant siboglinid tubeworms, bathymodioline mussels and vesicomid clams often harbour symbiotic chemosynthetic bacteria that depend on simple organic chemicals, such as methane and hydrogen sulphide, as their energy source (Tunnicliffe, 1991). Because hydrothermal vent habitats are discrete and ephemeral that can be easily destroyed by volcanic eruption and tectonic events (Jollivet et al., 1999; Shank et al., 1998; Tunnicliffe et al., 1997), vent species have adapted to these environmental disturbances with characteristics such as rapid growth rates and well-developed dispersal capabilities. Hydrothermal effluent produces large quantities of mineral deposits known as seafloor massive sulphides (SMS), which contain metals like nickel, gold, cobalt and manganese, and rare earth elements, that can be used to manufacture various products such as rechargeable batteries (Miller et al., 2018). The increasing demand for these metals leads to interests in commercial mining of the deep-sea SMS, which poses immediate threats to the vent species (Levin et al., 2020; Van Dover et al., 2018). Thus, there has been substantial interests in understanding the origin, connectivity and divergence of animals living in these deep-sea environments (reviews by Levin et al., 2016; Vrijenhoek, 2010) as such knowledge is required for the understanding of their biogeography and development of conservation plans.

The East Pacific Rise (EPR) is among the few extensively explored mid-ocean ridge systems. It is fast expanding in the northern EPR (NEPR, 82–111 mm year<sup>-1</sup>) and superfast expanding in the southern EPR (SEPR, 136–151 mm year<sup>-1</sup>) (Vrijenhoek, 2010). The ridge axis of the EPR is quite linear with many vent fields distributed in similar water depths (2000–2830 m). Along-axis currents have been thought to be important in the retention and dispersal of vent invertebrate larvae, but topographical discontinuities along the EPR could disrupt the along-axis dispersal. Such topographical discontinuities can be found at the Galápagos Triple Junction – a transformed fault near the Equator where there is a deep rift valley (i.e., Hess Deep) with a strong east-flowing current (Reid, 1997), in some sections of the southern EPR where the axis walls are low or absent, and around the Easter and Juan Fernandez microplates where the ridge is highly fractured (Jang et al., 2016; Johnson et al., 2013; Won et al., 2003).

Genetic studies have shown that life-history characteristics can modulate the effects of ocean current and bottom topography on population colonization and divergence along the EPR (Hurtado et al., 2004; Plouviez et al., 2009; Vrijenhoek, 2010). For instance, the brooding amphipod *Ventiella sulfuris* has low levels of gene flow to the extent that the NEPR and Galápagos Rift (GAR) populations represent distinct species, and this species is absent from the SEPR. In contrast, the vent tubeworm *Tevnia jerichonana*, a species with planktotrophic larvae, has intermediate gene flow along the ridge,

where the southern EPR (7–17°S) acts as a zone of introgression between the northern EPR (9–13°N) and PAR (31–32°S) populations, but the sequence divergence among these metapopulations is small and therefore considered intraspecific (Zhang et al., 2015). On the other hand, the polynoid scaleworm *Branchiopolynoe symmytilida* has high levels of gene flow among populations spanning from 9°N to 32°S latitude along the EPR (Hurtado et al., 2004; Plouviez et al., 2009).

*Hesiolyra bergi* is a free-living polychaete species widely distributed on vent chimneys along the EPR (Blake, 1985; Vrijenhoek, 2010). In their comparative genetic study of EPR species, Plouviez et al. (2009) analysed five populations of *H. bergi* distributed between 13°N and 21°S latitude with a fragment of the mitochondrial cytochrome c subunit 1 (COI) gene and found an uninterrupted gene flow across the sampled vent fields. Moreover, they found a minor clade of three individuals from 17°S and 21°S with a relatively larger sequence divergence (1%) in COI from the major clade of 232 individuals. They considered this minor southern clade to be either under-sampled or to represent a nearly extinct mitochondrial lineage (Plouviez et al., 2009). Nevertheless, the COI fragment used by Plouviez et al. (2009) is considerably shorter (366 bp) than the fragment of the gene used for barcoding, limiting the number of informative polymorphic sites for inference of population genetic structure. Furthermore, studies based on a single gene may suffer from selective sweeps masking signals of population divergence (Galtier et al., 2000; Xu et al., 2021). Multi-locus data are required to detect the presence of a selective sweep and distinguish it from other factors such as population bottleneck events (e.g., Coykendall et al., 2011; Jang et al., 2016; Johnson et al., 2017; Plouviez et al., 2013; Zhang et al., 2015). We therefore re-examined the population genetic structure of *Hesiolyra* along the EPR using sequences of six mitochondrial and two nuclear genes, based on the samples collected in 1999 and 2004 (Plouviez et al., 2009) and additional samples collected in 2010 (Table S1). We aimed to answer whether the Equator also acts as a barrier to gene flow in this species, what are the potential causes of the genetic heterogeneity in the SEPR, and whether the southern minor clade really represents a distinct lineage.

## 2 | METHODS

### 2.1 | Sampling and DNA extraction

Specimens of *Hesiolyra* were collected from five hydrothermal vent fields along the East Pacific Rise during *Nautilie* submersible expeditions on board the NO L'Atalante in 1999, 2004 and 2010 (Table S1). They were individually preserved in 80% ethanol and stored at the Station biologique de Roscoff, Université Pierre & Marie Curie, Paris VI, France. We extracted genomic DNA using either the CTAB method (Doyle & Dickson, 1987) or the PureLink™ Genomic DNA Mini Kit (Invitrogen) according to the manufacturer's protocol.

## 2.2 | PCR amplification and sequencing

We used two nuclear genes [Internal Transcribed Spacer 2 (*ITS2*) and alpha-methylacyl-CoA racemase (*COA*)] and six mitochondrial genes [16S ribosomal RNA (*16S*), ATP synthase membrane subunit 6 (*ATP6*), cytochrome c oxidase subunit I (*COI*), cytochrome b (*CYTB*) and NADH dehydrogenase subunits 4 (*NAD4*) and 5 (*NAD5*)]. We attempted to amplify more nuclear loci, including globin X and phosphoglucosmutase that were used to distinguish *Alvinella pompejana* populations (Jang et al., 2016; Plouviez et al., 2010), but it was unsuccessful. We amplified the fragments of *16S* and *CYTB* with primers used by Palumbi et al. (1991) and Boore and Brown (2000), respectively. For the other six genes, we designed specific primers for the genus *Hesiolyla* using Primer3Plus (Untergasser et al., 2012; Table S2). We performed polymerase chain reactions (PCR) with TaKaRa Taq DNA Polymerase Kit (Takara Bio) using programs included in Table S2. The PCR products were sequenced bidirectionally with the PCR primers at BGI (Beijing Genomics Institute) on an ABI3730XL sequencer. We visualized the sequencing results and assembled them using CodonCode Aligner v9.0.1 (<http://www.codoncode.com/>) based on the overlapping strands in the two directions. We aligned the multiple sequences with ClustalW v1.81 (Thompson et al., 1994) under the default settings (15 gap opening penalty and 6.66 gap extension penalty) and edited the sequences by eye. We resolved heterozygous positions for the *COA* fragment with PHASE implemented in DNASP v6 (Rozas et al., 2017).

## 2.3 | Genetic distance and phylogeny

We identified the three individuals with a greater divergence from all other individuals reported by Plouviez et al. (2009) by searching our *COI* sequences against the GenBank nucleotide database. We calculated pairwise distances between the minor clade of three individuals (referred to as *H. aff. bergi* hereafter) and the major clade of *H. bergi* s.s. (referred to as *Hesiolyla bergi* hereafter) for each of the eight loci using MEGA v11 (Tamura et al., 2021) based on the Kimura 2-Parameters (K2P) substitution model. As *16S* and *COI* sequences were available in the other two described species of *Hesiolyla* (*H. heteropoda* and *H. longqiensis*, Wang et al., 2020, Table S3), we also included them in the K2P distance calculation.

To determine the phylogenetic relationships among *Hesiolyla* spp., we constructed phylogenetic trees using the maximum likelihood (ML) and Bayesian Inference (BI) methods. We determined the best-fit model for ML and BI analyses using IQ-TREE v1.4.4 (Nguyen et al., 2015). We used four species of *Gyptis* as the outgroup due to their close relationship with *Hesiolyla* (Rouse et al., 2018; Wang et al., 2020). We then performed ML analysis using IQ-TREE with 1000 ultrafast bootstrap (UFBoot) replicates (Minh et al., 2013) and BI analysis using MrBayes v3.2.6 (Ronquist et al., 2012) on two separate runs of Markov chains for 50 million generations and discarded the first 25% of trees as burn-in.

## 2.4 | Statistical tests for polymorphism and divergence

We divided the specimens from the five vent fields into three groups representing the northern EPR of *H. bergi* (NEPR: 13°N and 9°N), southern EPR of *H. bergi* (SEPR: 17°S, 18°S and 21°S) and *H. aff. bergi*. For each locus, we calculated the number of haplotypes (*H*), the number of polymorphic sites (*S*), the haplotype diversity (*Hd*) and the nucleotide diversity ( $\pi$ ) for the NEPR and SEPR sites separately using DnaSP v6 (Rozas et al., 2017). To determine the genealogical relationships between haplotypes, we constructed a TCS haplotype network (Clement et al., 2002) using PopART v1.7 (Leigh & Bryant, 2015) for each gene.

We examined the genetic differentiation among the three groups for each gene by calculating the site-based fixation index ( $\Phi_{st}$ ) with 10,000 randomizations in Arlequin v3.5.2.2 (Excoffier & Lischer, 2010). We tested the isolation-by-distance (IBD) by evaluating the amount of genetic differentiation ( $\Phi_{st}$ ) between local populations (vent fields) and their geographic distances. We estimated the geographic distances among the five vent fields using the Google Maps Distance Calculator (<https://www.daftlogic.com/projects-google-maps-distance-calculator.htm>) and applied the Mantel test with 10,000 permutations in ZT software package (Bonnet & de Peer, 2002) to determine the relationship between the geographical distances and pairwise  $\Phi_{st} / (1 - \Phi_{st})$ . We also performed a hierarchical analysis of molecular variance (AMOVA) among and within the NEPR and SEPR groups using PopART v1.7 (Leigh & Bryant, 2015) to detect the distribution of genetic variations within and between populations.

To detect demographic and/or selective events, we calculated the Tajima's *D* (Tajima, 1989) and Fu's *F<sub>s</sub>* (Fu, 1997) separately for each locus at each vent field using Arlequin v3.5.2.2 (Excoffier & Lischer, 2010), with *p*-values generated using 1000 simulations.

## 2.5 | Genetic structure

We determined the genetic clusters based on a multi-locus genotype approach including all loci for each individual using STRUCTURE v2.3.4 (Pritchard et al., 2000). We used two datasets, with and without *H. aff. bergi*, to perform the genetic structure analyses. We estimated the most probable number of discrete clusters ( $K = 1-5$ ) with ten independent runs for each *K* value using an admixture model with correlated allele frequencies among populations. We performed each run with  $1.2 \times 10^6$  Markov Chain Monte Carlo (MCMC) iterations after a burn-in of the first  $2 \times 10^5$  iterations. Then, we determined the most probable value of *K* using Structure Harvester Web v0.6.94 (Earl & VonHoldt, 2012) with the  $\Delta K$  method (Evanno et al., 2005). We visualized the results using STRUCTURE PLOT (Ramasamy et al., 2014).

## 2.6 | Gene flow and demographic history

We calculated a *D*-statistic (ABBA/BABA test) to assess gene flow between *H. bergi* and *H. aff. bergi* using Dsuite (Malinsky et al., 2020),

after converting the alignments of each locus to the VCF format using the MSA2VCF tool in Jvarkit v20200206 (Lindenbaum, 2015). We tested the introgression from *H. aff. bergi* (P3) into the NEPR (P1) and SEPR (P2) groups of *H. bergi* using a guided species tree in the form of “([P1, P2],P3),O,” with *H. longqiensis* as the outgroup (O). We assessed the statistical significance by calculating Z-scores and p-values using block jackknife in Dsuite. A D-statistic that differs significantly from zero (Z-score >3, and  $p < .05$ ) indicates the presence of gene flow.

To test the effects of the Equatorial barrier on the genetic structure of *H. bergi* s.s., we performed the Isolation with Migration (IM) and Sweep\_bott analyses to the NEPR and SEPR meta-populations of *H. bergi* s.s. Because these analyses assume no recombination, we used IMgc (Woerner et al., 2007) to remove potential recombinant individuals and generate a non-recombining block (Table S2).

We conducted the IM analysis in IMA2p (Sethuraman & Hey, 2016), which estimated a splitting time parameter ( $t$ ), an effective population parameter ( $\theta$ ) and migration parameters in each direction ( $N_m$ ). We performed preliminary runs using MCMC-mode (M-model) with an inheritance scalar of 0.25 for mitochondrial markers and 1 for nuclear markers as recommended in the IMA documentation (Hey, 2009), and the Hasegawa-Kishino-Yano mutation model (HKY, Hasegawa et al., 1985). We assigned a mutation rate of  $1.1 \times 10^{-5}$  per site per year for COI (548bp) based on an evolutionary rate of 0.2% per million years that was used for other vent annelids (Chevaldonné et al., 2002). After assessing the results of preliminary runs, we set the upper boundary values of the uniform prior distribution as:  $q$  (prior population size parameter) = 30,  $m$  (migration rates) = 1 and  $t$  (divergence time) = 2, and forty geometric heating chains ( $h_a = 0.97$ ,  $h_b = 0.9$ ) for the regular M model runs. We performed seven regular M-model runs using random seeds for at least 10 million MCMC steps and discarded the first  $1.0 \times 10^5$  steps as burn-in. Then, we combined the results of the regular M-model runs in a Load-genealogies mode (L-mode) run to estimate the log maximum likelihood and credibility intervals (95% highest posterior density, HPD) for migration parameters using likelihood ratio tests (LLR) (Nielsen & Wakeley, 2001) and re-scaled parameters to time in years. Due to the absence of a confirmed mutation rate in *H. bergi*, we estimated the geometric mean of mutation rates across the eight loci using the scalar method (Hey & Nielsen, 2004) based on the mutation rate of COI and mutation rate scalar generated from the regular M-model runs. The geometric mean of mutation rates, estimated as  $1.82 \times 10^{-6}$  per site per year, was applied to the L-model run. Finally, we visualized the model parameters generated from IMA2p using the IMfig program (Hey, 2010).

We detected bottleneck and/or selective sweep events during the demographic history of the *H. bergi* s.s. populations using the program Sweep\_bott (Galtier et al., 2000). Sweep\_bott applies a coalescence-based maximum likelihood approach to estimate the time since the occurrence of a diversity-reduction event and its strength by estimating the T and S parameters, and calculates the maximum likelihood across these parameter values under three distinct population models (the null constant size model M1, the

bottleneck model M2 and the sweep model M3). Any significant departure from M1 will be considered as a diversity-reducing event [either bottleneck (M2) or selective sweep (M3)]. To discriminate between these two alternative scenarios of diversity reduction (M2 vs M3), we assumed that in M2 (bottleneck) all loci share the same time since the occurrence of the diversity-reducing event (T) and the same strength of diversity reduction (S), whereas in M3 (selective sweep) these two parameters are locus-specific. Thus, we estimated the T and S parameters and maximum likelihood under M3 model locus by locus. We performed the analyses with a dataset of six loci because the two loci 16S and ITS2 only displayed one unique segregating site in the SEPR populations. We performed the coalescent simulations with 100,000 burn-in iterations, 1000,000 second-step iterations and 20 optimization processes with a theta range from 1 to 30 in each run. Then, we compared the three models (M2 vs M1, M3 vs M1, M3 vs M2) with likelihood ratio tests (LRTs) to find the most appropriate model.

### 3 | RESULTS

#### 3.1 | New insight into the rare southern EPR lineage of *Hesiolyra*

Our analyses of the DNA sequences confirmed the result of Plouviez et al. (2009) that three specimens collected from vent fields at 17°S and 21°S of SEPR represent a lineage that differed substantially from *H. bergi*, with an even higher divergence (2%) of K2P distance with our longer 548-bp COI sequences (Table 1). The distinct southern lineage showed 1.11% to 4.18% K2P distance from *H. bergi* for CYTB, NAD4, ATP6, COA and ITS2 (Table S4). The genetic distance of COI between the EPR and Indian Ridges (IR) species ranged from 6.78% (*H. bergi* vs *H. longqiensis*) to 9.62% (*H. aff. bergi* vs *H. heteropoda*) and was lower than that between the two IR species (12.72%; Table 1).

We used a dataset of 548-bp COI sequences for the phylogenetic studies. IQ-TREE indicated that GTR+I+G was the best model for both ML and BI analyses. To reduce the complexity of the tree, we used only one sequence for each haplotype in *H. bergi*, all three available sequences of *H. aff. bergi*, and all available sequences in the other two species of *Hesiolyra* (Wang et al., 2020, Table S3). The phylogenetic analyses using the ML and BI methods generated the same tree topology, which showed that the three specimens of *H. aff. bergi* represented a well-supported monophyletic clade (UFBoot

TABLE 1 Mean inter- and intra-specific percentage K2P genetic distances for species of *Hesiolyra*, calculated based on COI sequences.

	1	2	3	4
1. <i>H. bergi</i> s.s.	0.21			
2. <i>H. aff. bergi</i>	2.07	0.87		
3. <i>H. longqiensis</i>	6.78	7.14	0.22	
4. <i>H. heteropoda</i>	9.24	9.63	12.72	1.65

>85, posterior probability >0.95, Figure 1). This clade was most closely related to a clade representing all specimens of *H. bergi* from the EPR, and the *H. aff. bergi* + *H. bergi* clade was sister to *H. longqiensis* (Figure 1). The clade representing *H. heteropoda* was sister to the other *Hesiolyra* spp. in all analyses (Figures 1 and S1).

### 3.2 | Haplotype analyses revealed gene frequency shifts across the equatorial gap

The NEPR group showed higher levels of both haplotype diversity (*Hd*) and nucleotide diversity ( $\pi$ ) than the SEPR group for all loci, except for COA (Table 2). Except for the common haplotypes of 16*S*, *H. bergi* did not share any haplotype with *H. aff. bergi* for the other seven loci (Figure 2). One common haplotype distributed throughout the sampling range was detected for both 16*S* and *ITS2*. For *ATP6*, *COI*, *CYTB*, *NAD4* and *NAD5*, we did not find any shared

haplotypes between the NEPR and SEPR populations of *H. bergi*. However, the two metapopulations of *H. bergi* shared haplotypes for the two nuclear loci, with six for COA and one for *ITS2*, respectively (Figure 2).

The 16*S* network showed the simplest topology with two main haplotypes that differed from each other by one substitution and are shared by both *H. bergi* and *H. aff. bergi*, and three surrounding singleton haplotypes that differed by one or two substitutions from these main haplotypes. For the other seven loci, *H. aff. bergi* was distinct from *H. bergi* by at least three substitutions (Figure 3). For *ATP6*, *COI*, *CYTB*, *NAD4* and *NAD5*, the networks revealed a clear geographical pattern within *H. bergi*. The NEPR and SEPR individuals of *H. bergi* were completely separated for the five loci, with the northern individuals exhibiting more haplotypes, and the southern individuals, in general, forming a star-like topology with one or two dominant haplotypes surrounded by multiple singletons. By contrast, the networks of COA and *ITS2* did not exhibit a clear geographical pattern in

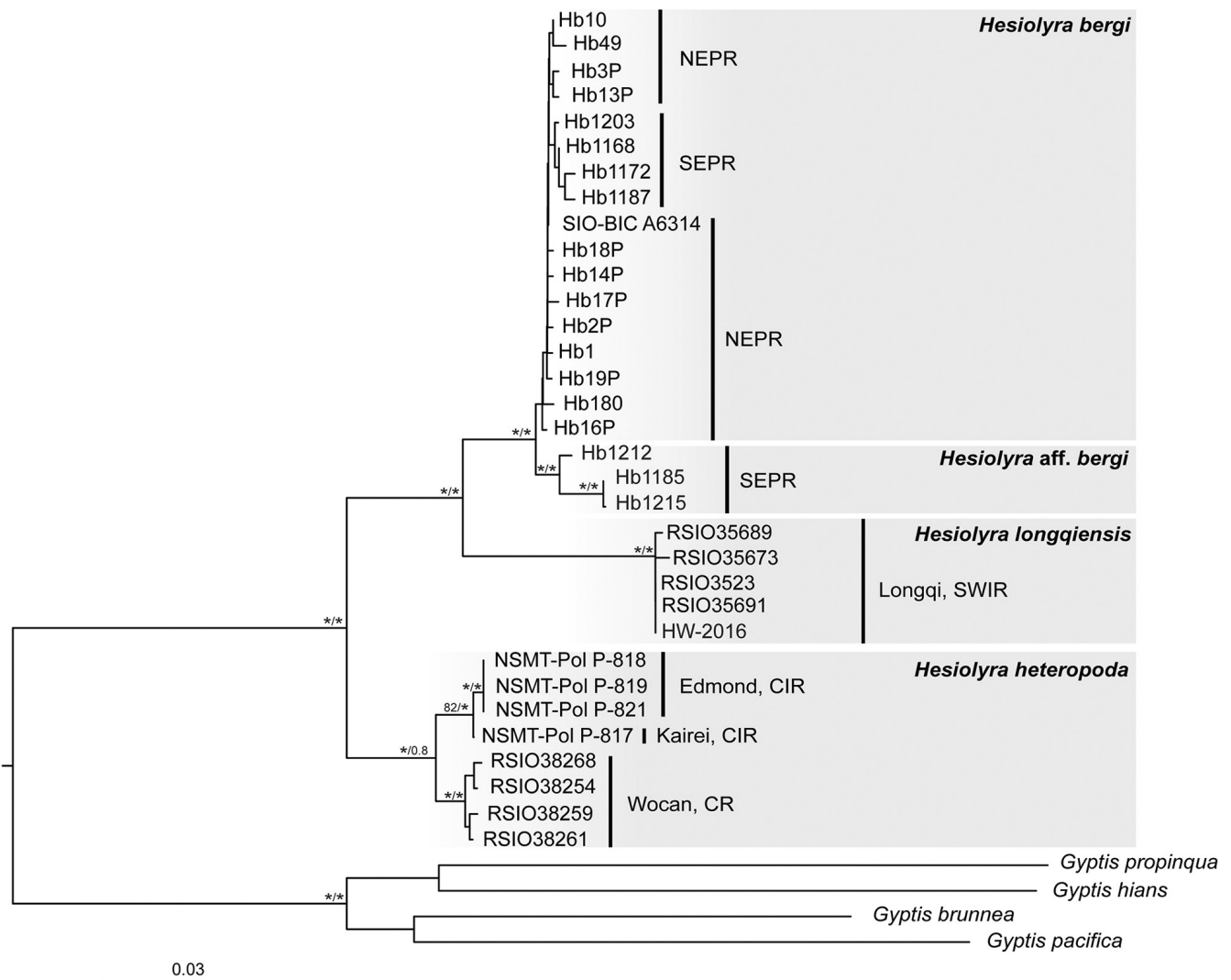


FIGURE 1 Phylogenetic tree based on the haplotype of 548-bp *COI* sequences. Tree topology is from the maximum likelihood (ML) analysis using IQ-TREE. Ultra-bootstrap values of ML analyses and posterior probability (pp) values of Bayesian inference (BI) analyses are shown at the branch node and separated by slash. Asterisks indicate bootstrap values >85 or pp values >0.95. CIR, Central Indian Ridge; CR, Carlsberg Ridge; NEPR, northern East Pacific Ridge; SEPR, southern East Pacific Ridge; SWIR, Southwest Indian Ridge.

TABLE 2 Genetic diversity of *Hesolyra bergi* populations across the East Pacific Rise.

Locus	location	N	S	H	Hd	$\pi$	Tajima's D	Fu's Fs
16S	NEPR	45	4	4	0.4636	0.0012	-0.9173	-0.6695
	SEPR	22	1	2	0.0909	0.0002	-1.5147*	-2.0784**
	Total	67	5	5	0.3794	0.0010	-1.2160	-1.3740
ATP6	NEPR	37	33	17	0.9670	0.0106	-1.1777	-11.0658*
	SEPR	42	10	8	0.6144	0.0017	-1.8956*	-10.4990***
	Total	79	50	25	0.8931	0.0102	-1.0244	-6.5141*
COI	NEPR	32	13	12	0.7057	0.0019	-2.2267***	-10.0412***
	SEPR	32	3	4	0.1815	0.0003	-1.7366*	0.5808
	Total	64	28	16	0.7499	0.0037	-1.9817*	-4.7302
CYTB	NEPR	37	22	16	0.9144	0.0065	-1.7050*	-18.4503***
	SEPR	18	4	4	0.8619	0.0032	-0.2709	-0.0919
	Total	55	26	20	0.9468	0.0102	-0.9547	-9.2711
NAD4	NEPR	32	17	12	0.9113	0.0070	-0.6805	-9.6399***
	SEPR	33	6	3	0.6250	0.0018	-1.45033	0.4945
	Total	65	33	15	0.8933	0.0088	-1.0654	-4.5727
NAD5	NEPR	36	17	15	0.8895	0.0043	-1.5728	-8.4074***
	SEPR	32	4	4	0.2946	0.0007	-1.5940	-1.8481
	Total	68	22	19	0.8125	0.0042	-1.5834	-5.1278
COA	NEPR	30	10	17	0.9356	0.0076	0.3576	-10.3951***
	SEPR	30	15	22	0.9908	0.0088	-0.4684	-26.1505***
	Total	60	19	35	0.9770	0.0105	-0.0369	-11.9475***
ITS2	NEPR	29	12	9	0.5584	0.0039	-0.4706	-0.1181
	SEPR	37	0	1	0.0541	0.0002	-1.1241	-1.4501
	Total	66	12	9	0.5720	0.0042	-1.3884	-1.7043

Abbreviations: H, number of haplotypes; Hd, haplotype diversity; N, number of sequences; S, number of segregating sites;  $\pi$ , nucleotide diversity. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

*H. bergi*, with one dominant haplotype of mixed geographical origin surrounded by one or two less frequent haplotypes and singletons (Figure 3).

### 3.3 | Population differentiation

Locus-by-locus hierarchical AMOVA analyses revealed that most of the genetic variation was partitioned between the NEPR and SEPR groups (ranging from 61.7% to 78.5%,  $p < .05$ , Table 3) rather than among the individuals of each vent field (from 22.1% to 40.1%,  $p < .05$ ) for COI, CYTB, NAD4 and NAD5. By contrast, most of the genetic variance in 16S, ATP6, COA and ITS2 could be explained by differences among individuals within each vent field (71.4% to 87.3%,  $p < .05$ ), compared to the small but still significant differentiation (at least for three of them) between the two groups (18.7% to 33.1%).

The correlation between genetic and geographic distances was statistically significant for the whole dataset ( $r = 0.988$ ,  $p = .008$ , Figure S2), which did not reject the Isolation-by-distance (IBD) model in *H. bergi* across the EPR.

### 3.4 | Geographical structure and admixture

The STRUCTURE analysis based on the multi-locus dataset of *H. bergi* s.s. identified two most possible clusters divided across the equator (Figure 4a). There was genotype admixture between the NEPR and SEPR clusters in two individuals from the 13°N vent field, one individual from 17°S, one individual from 18°S and two individuals from 21°S (Figure 4a), while the STRUCTURE analysis with *H. aff. bergi* added identified three most possible clusters, represented by the NEPR and SEPR populations of *H. bergi* and *H. aff. bergi*, respectively (Figure 4b). Moreover, the analysis detected genotype admixture between the NEPR population of *H. bergi* and *H. aff. bergi* in one individual from 13°N, and genotype admixture among the three clusters in two individuals from 13°N, one individual from 17°S and two individuals from 21°S, respectively (Figure 4b).

The ABBA/BABA test did not provide statistical support for the introgression between *H. bergi* and *H. aff. bergi* ( $D = 0.18$ ,  $p > .05$ , Z-score  $< 3$ , Figure S3), indicating that incomplete lineage sorting might be a better explanation of the results, or that our nuclear genes were not sensitive enough to detect introgression.

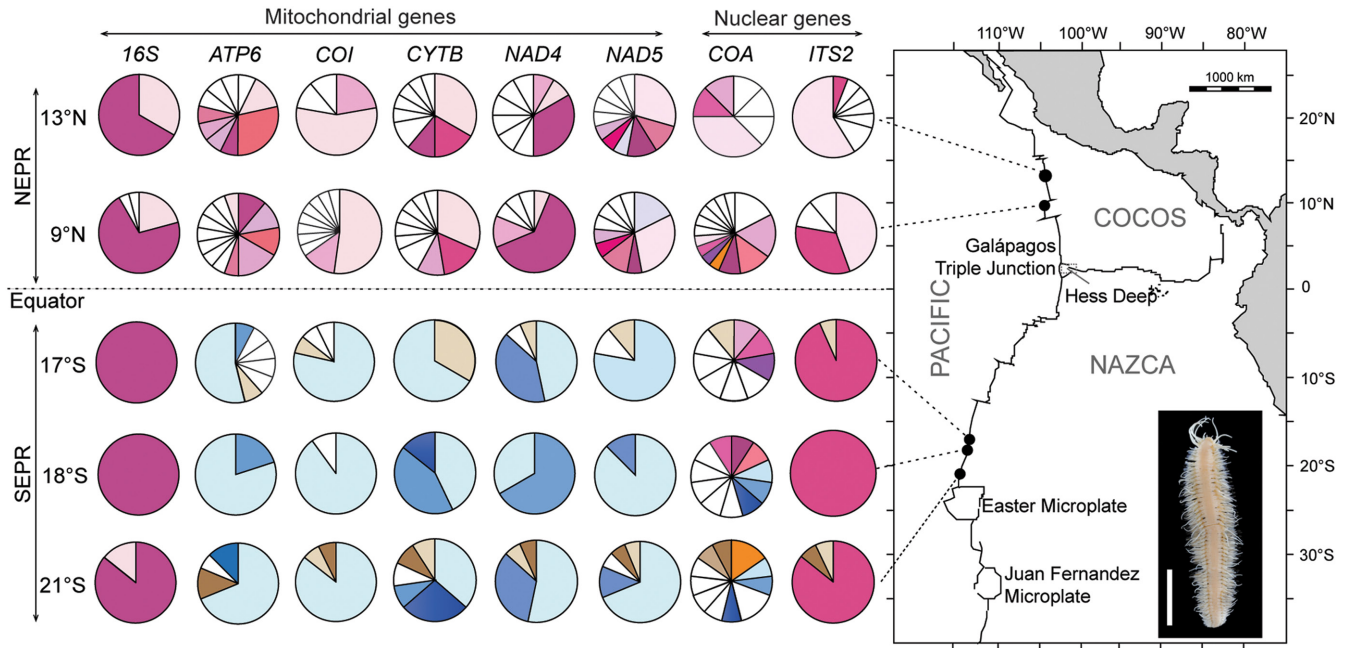


FIGURE 2 Geographic distribution of divergent haplotype of *Hesiolyra* from EPR at six mitochondrial and two nuclear loci. Shared haplotypes are coloured. Private haplotypes are put in white. Light brown and brown colours are haplotypes of *H. aff. bergi*. Scale bar: 5 mm.

### 3.5 | Historical divergence and gene flow

We estimated the effective population size, population splitting time and migration rate parameters in forward time with a two-population IM model (Figure 5 and Table S5). The estimated population size of the NEPR group was approximately nine times larger than the ancestral one ( $\theta_{\text{NEPR}} = 18.35$ ,  $\theta_{\text{SEPR}} = 1.95$ ,  $\theta_{\text{ANCESTOR}} = 2.05$ ), which indicated historical population expansion in the NEPR. The marginal posterior probability  $t_0$  was 0.536, which was converted into a divergence time of 0.45 Mya (95% HPD: 0.27–0.74 Mya). An  $N_m$  value  $>1$  indicates sufficient gene flow to maintain continuity among populations (Yang et al., 2022). The multi-locus estimates of migration revealed a weak and rather equal bidirectional gene flow between the NEPR and SEPR meta-populations:  $2N_m = 0.484$  from NEPR to SEPR, and  $2N_m = 0.475$  from SEPR to NEPR (Figure 5).

### 3.6 | Demographic dynamics

Significant negative values of Tajima's  $D$  and Fu's  $F_s$  indicate a recent population expansion or a selective sweep at the locus. Tajima's  $D$  values were negative except for  $COA$  in the SEPR group, but the values were statistically significant only for  $COI$  (Table 2). All Fu's  $F_s$  values were negative, with statistically significant differences from zero for  $ATP6$  and  $COI$  in both the NEPR and SEPR groups, and for  $CYTB$  and  $NAD5$  in the NEPR group.

Results of the Sweep\_bott analyses are summarized in Table 4. For both NEPR and SEPR groups, the bottleneck model (M2) showed a significantly better fit than the constant-size model (M1) (M2/

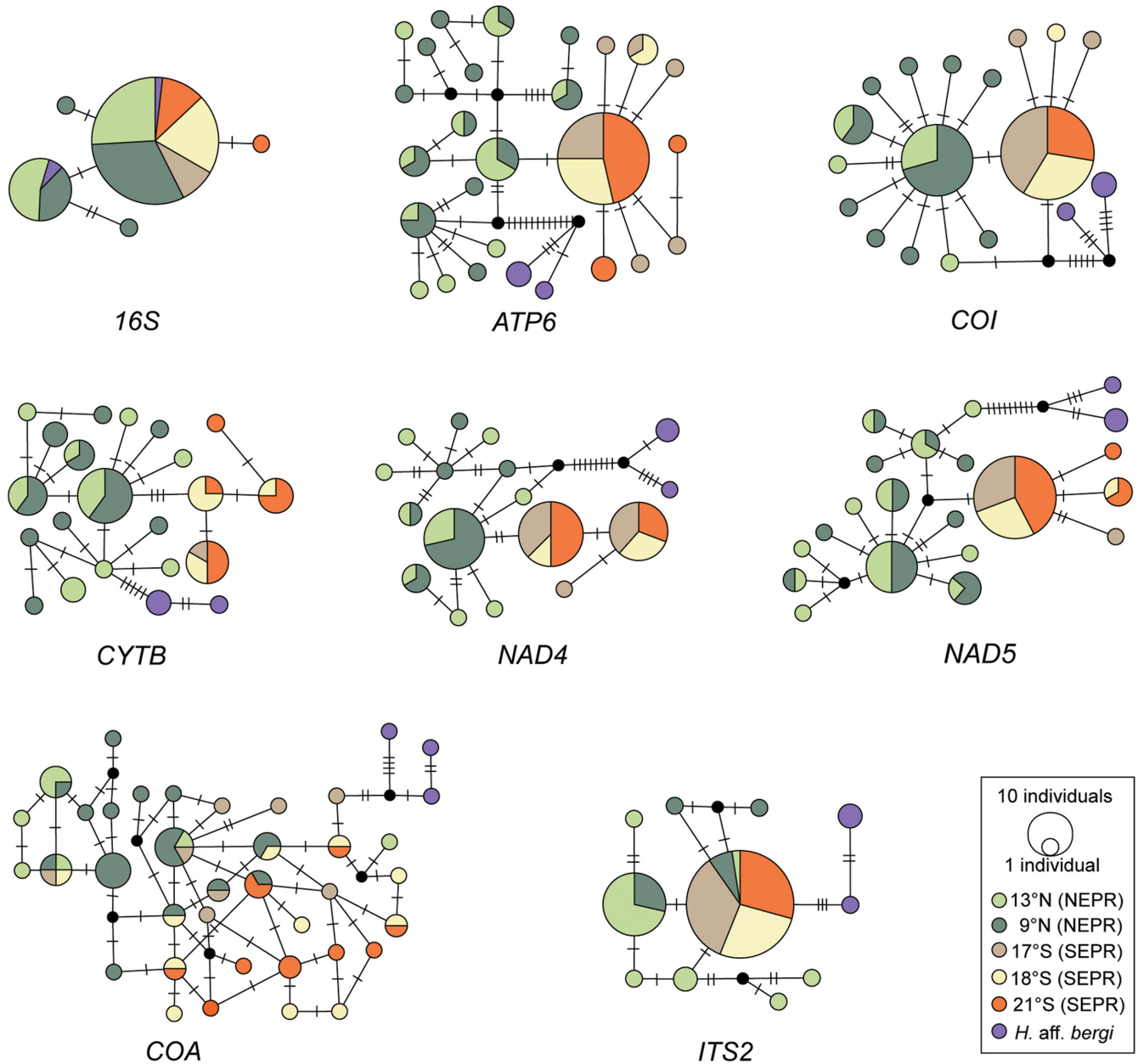
M1 LRTs 12.56 and 8.34, respectively,  $p < .05$ ). However, LRTs for selective-sweep model (M3) fit the dataset significantly better than the other two models in the NEPR group (LRTs M3/M1 = 51.72, M3/M2 = 39.16,  $p < .05$ ), suggesting that the sweep model was appropriate for the NEPR group, whereas the bottleneck model was appropriate for the SEPR group. In addition, under the bottleneck model, the NEPR group had a higher  $T$  value than the SEPR group ( $T = 1.51$  in NEPR vs  $T = 1.19$  in SEPR), indicating an older bottleneck event in the NEPR group.

## 4 | DISCUSSION

### 4.1 | Presence of a southern EPR lineage in *Hesiolyra*

In this study, we used multi-locus data to unveil the genetic structure of *Hesiolyra* along the EPR from 13°N and 21°S. We found that three individuals from 17°S and 21°S were ~2% divergent from other individuals in the mitochondrial  $COI$  gene. The three individuals, in this study referred to as *H. aff. bergi*, were also substantially different from *Hesiolyra bergi* s.s. in other mitochondrial ( $CYTB$ ,  $NAD4$ ,  $NAD5$ ,  $ATP6$ ) and nuclear ( $COA$  and  $ITS2$ ) markers. In a large-scale study that involved amplification of the  $COI$  fragment from 1876 polychaete specimens representing 333 provisional species, Carr et al. (2011) found that the average intraspecific K2P distance was 0.38%, and in most species (85%) this distance was  $<1\%$ . Plouviez et al. (2009) suggested that the three divergent individuals represented either an under-sampled clade or a nearly extinct mitochondrial lineage. While we cannot rule out the first possibility, it appears that, since





**FIGURE 3** TCS haplotype networks based on six mitochondrial and two nuclear loci of *Hesiolyra* from EPR. Each circle represents a different haplotype and sizes of circles correspond to the frequency of the haplotypes. Each dash on each link indicates one substitution. Black dots indicate unknown or missing haplotypes.

the divergence affected multiple mitochondrial and nuclear genes, and *Hesiolyra* specimens have been collected from vent fields as far as 32°S (Robert Vrijenhoek, personal communication), these individuals may represent a distinct lineage that is more common in the northeast Pacific Antarctic Ridge (PAR). While this hypothesis can be tested when samples from PAR become available for investigation in the future, it is consistent with the suggestion that the formation of the Easter Microplate (2.5–5.3 Mya) was responsible for the divergence of vent species, such as crabs *Bythyograea laubieri* and *B. vrijenhoeki* (Guinot & Hurtado, 2003), mussels *Bathymodiolus thermophilus* and *B. antarcticus* (Johnson et al., 2013), and partial isolation of polychaete *Riftia pachyptila* (Coykendall et al., 2011; Hurtado

et al., 2004), *Alvinella pompejana* (Jang et al., 2016) and *Tevnia jerichonana* (Hurtado et al., 2004; Zhang et al., 2015). The presence of the three divergent *Hesiolyra* individuals in the SEPR also supports the idea that the Easter Microplate, which links the PAR and SEPR, allows secondary contact between the two previously isolated species around this region (Johnson et al., 2013; Plouviez et al., 2013; Zhang et al., 2015).

The pattern of admixture between *H. bergi* and *H. aff. bergi* was likely due to the incomplete lineage sorting. Incomplete lineage sorting has been considered to be associated with recently separated populations (Degnan & Rosenberg, 2009), and accounts for the genotype distribution pattern of *Bathymodiolus septemdiemum* – a

**TABLE 3** Analysis of molecular variance (AMOVA) for populations of *Hesiolyra bergi* from five vent fields assigned to NEPR and SEPR groups, respectively, based on six mitochondrial gene fragments and two nuclear gene fragments.

Gene	df	SS	VC	%
Mitochondrial genes				
16S				
Between groups	1	1.938	0.065	19.057
Among populations within groups	3	0.117	-0.022	-6.332
Within populations	62	18.482	0.298	87.276***
ATP6				
Between groups	1	109.409	3.139	33.165***
Among populations within groups	3	2.559	-0.435	-4.591
Within populations	77	439.403	6.760	71.427***
COI				
Between groups	1	34.312	1.059	68.300***
Among populations within groups	3	1.363	-0.004	-0.233
Within populations	59	29.200	0.495	31.934***
CYTB				
Between groups	1	261.671	11.132	78.528***
Among populations within groups	3	6.848	-0.092	-0.650
Within populations	54	169.345	3.136	22.123***
NAD4				
Between groups	1	206.453	7.130	70.7133***
Among populations within groups	3	1.086	-0.260	-2.577
Within populations	53	170.271	3.213	31.864***
NAD5				
Between groups	1	96.613	2.957	61.723***
Among populations within groups	3	2.546	-0.087	-1.8144
Within populations	60	115.241	1.921	40.092***
Nuclear genes				
COA				
Between groups	1	7.986	0.268	26.176***
Among populations within groups	3	4.325	0.116	11.287
Within populations	41	26.255	0.640	62.537
ITS2				
Between groups	1	2.963	0.090	18.693*
Among populations within groups	3	1.397	0.007	1.3856
Within populations	54	20.894	0.387	79.922***

Abbreviations: df, degrees of freedom; SS, sum of squares; VC, variance components; %, percentage of variation.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

complex of deep-sea mussels widely distributed in Indo-Pacific vent fields (Breusing et al., 2015).

## 4.2 | The 9°N–17°S barrier to gene flow

Previous studies have identified several transform faults that offset the ridge axis between 9°N–17°S along the EPR. These faults initiated 1–2 Mya (Francheteau et al., 1990; Kureth & Rea, 1981),

later than the formation of the Easter Microplate. In particular, the Hess Deep, a 6000m depression at the Galápagos triple junction near the Equator, has been considered to impede the dispersal of the negatively buoyant larvae of *A. pompejana* (Hurtado et al., 2004). Furthermore, a strong eastward deep-sea current near the Equator (Reid, 1997) generates northern and southern gyres that may transport pelagic larvae like those of crabs off the ridge axis (Hurtado et al., 2004). Consistent with these geological and oceanographic features, many of the previously studied vent

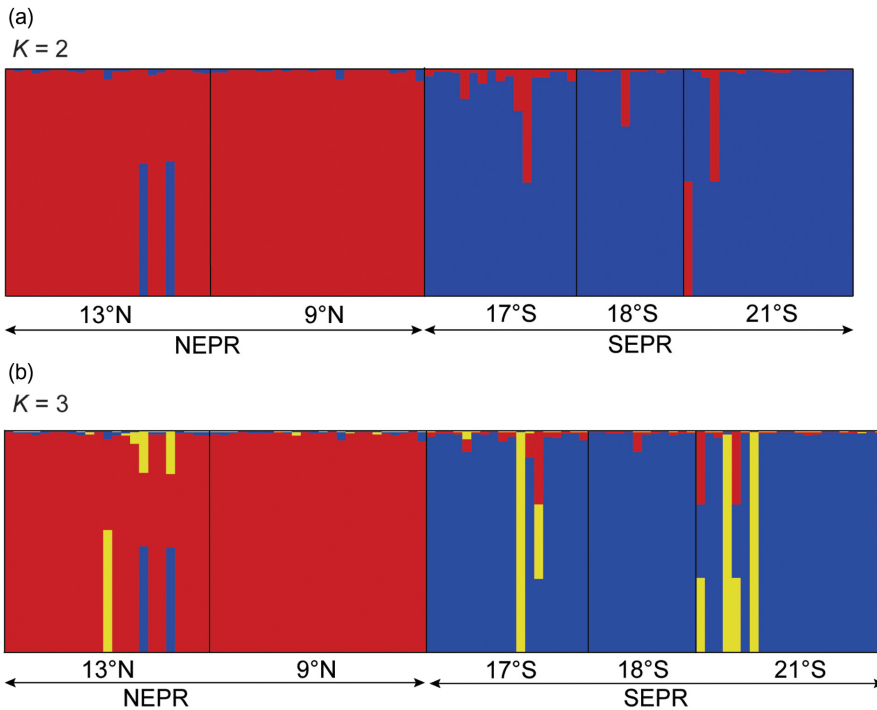


FIGURE 4 (a) STRUCTURE plot at  $K = 2$  based on eight loci of *Hesiolyra bergi* s.s. (b) STRUCTURE plot at  $K = 3$  based on eight loci of *H. bergi* s.s. and *H. aff. bergi*. Each individual is represented by a vertical bar, different colours indicate different genetic clusters.

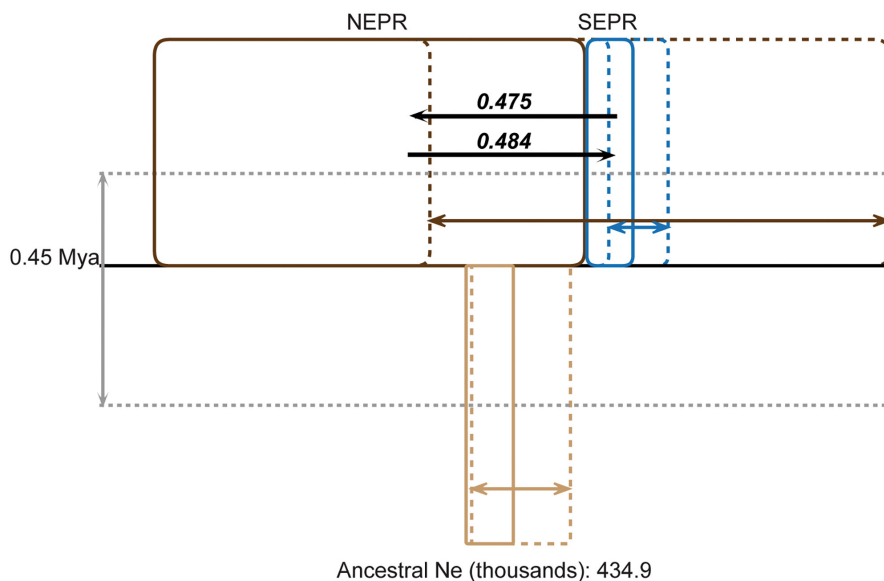


FIGURE 5 Diverging processes of *Hesiolyra bergi* s.s. from EPR estimated with IMA2p. The width of the black boxes represents the effective population size of each population and each ancestral population; the width of the dashed boxes represents the 95% highest posterior density (HPD) interval. Horizontal lines represent divergence time between populations. The black arrows represent the migration rate ( $2N_m$ ) between populations forward in time.

species, such as the tubeworms *R. pachyptila* and *T. jerichonana*, the mussel *B. thermophilus* and the gastropods *Lepetodrilus elevatus* and *L. ovalis*, exhibit different levels of divergence between the northern and southern EPR (Johnson et al., 2013; Matabos et al., 2011; Matabos & Jollivet, 2019; Plouviez et al., 2013; Vrijenhoek, 2010; Zhang et al., 2015).

In this study, we found a clear difference in haplotype compositions between the NEPR and SEPR populations of *H. bergi* in five of the eight studied genes. This result, although not surprising given that these genes are linked in the mitochondrial genome, is another empirical evidence that the 9°N–17°S region represents a barrier to gene flow. Nevertheless, the nuclear *COA* and *ITS2* genes showed lower levels of differentiation, and the mitochondrial *16S* gene

showed no detectable differentiation across this region of the ridge axis (Figure 2). An explanation of the inter-loci differences is that natural selection acts against the homogenizing effect of gene flow (Nosil et al., 2009). Further studies with more nuclear loci may allow us to determine whether such differential permeability is related to local adaptation or constraints of some particular functional genes.

Despite their clear differentiation, the NEPR and SEPR populations of *H. bergi* exhibited only a low level of divergence, which is similar to several other studied vent species in this region (Coykendall et al., 2011; Johnson et al., 2013; Plouviez et al., 2013; Vrijenhoek, 2010). This low level of divergence reflects the recent population splitting time (0.45 Mya, 95% HPD: 0.32–0.87 Mya) estimated in our study (Figure 5, Table S5). Plouviez et al. (2009)

TABLE 4 Maximum likelihood estimation (MLE) and likelihood ratio tests (LRT) estimated using Sweep\_bott for coalescent models of bottleneck, selective sweeps and constant population size.

Model	Multilocus		Single locus							COA
	log (L)		ATP6	COI	CYTB	NAD4	NAD5			
NEPR	M1 (6p)	-530.80	-184.05	-51.91	-111.78	-76.43	-91.70	-14.93		
	M2 (8p)	-524.52 (1.51, 0.12)								
	M3 (18p)	-504.94	-176.40 (0.38, 0.01)	-36.53 (0.25, 0.16)	-113.76 (1.31, 0.38)	-74.57 (2.36, 1.28)	-90.82 (2.77, 1.87)	-12.86 (0.92, 0.02)		
LRT (df)	M2/M1	12.56 (2) **								
	M3/M1	51.72 (12) ***	15.3 <sup>a</sup> ***	30.78 ***	3.96 <sup>a</sup>	3.72	1.76	4.14 <sup>a</sup>		
	M3/M2	39.16 (10) ***								
SEPR	M1 (6p)	-133.39	-27.42	-7.26	-18.35	-27.23	-20.10	-33.03		
	M2 (8p)	-129.22 (1.19, 0.33)								
	M3 (18p)	-127.18	-25.26 (1.72, 0.58)	-7.45 (1.48, 1.4)	-17.34 (0.45, 0.1)	-27.76 (0.92, 1.19)	-19.44 (0.92, 1.43)	-29.93 (1.78, 1.26)		
LRT (df)	M2/M1	8.34 (2) *								
	M3/M1	12.42 (12)	4.32 <sup>a</sup>	0.38	2.02	1.06	1.32	6.2 <sup>a</sup>		
	M3/M2	4.08 (10)								

Note: Values in brackets (T, S) correspond to MLE of the time since and the strength of either the bottleneck or the sweep in units of 2N generations, respectively. Abbreviations: df, degrees of freedom within parentheses; Log(L), logarithm of the maximum likelihood; M1, Null constant-size model; M2, Bottleneck model; M3, Selective-sweep model. <sup>a</sup>df = 1 for M3-M1 single-locus LRT. \*p < .05; \*\*p < .01; \*\*\*p < .001.

estimated that the divergence time between the NEPR and SEPR populations ranged from 0.7 to 11.6 Mya in different vent species and suggested that a vicariance event that occurred at ~1.3 Mya in this region was responsible for the simultaneous divergence of these species. Nevertheless, younger divergence times were estimated for *Branchiopolynoe symmytilida* (0.7 Mya; Plouviez et al., 2009) and *A. pompejana* (0.79 Mya, 95% HPD: 0.07–6.67 Mya) (Jang et al., 2016) across the equatorial barrier. Given the very low number of fixed mutations between the NEPR and SEPR populations of *H. bergi*, our estimation of a more recent splitting date appears reasonable. It is possible that the Equatorial barrier has trapped some introgressed and more divergent alleles from PAR, which could have been used for the multi-species divergence time estimate by Plouviez et al. (2009), leading to overestimation of the divergence time. Alternatively, the studied loci could have slower mutation rates in *Hesiolyra* and *Branchiopolynoe* than in other taxa. A physical barrier to dispersal would then take longer for *Hesiolyra* and *Branchiopolynoe* to translate into genetic divergence, leading to an underestimation of the time of the vicariance event in these species.

#### 4.3 | Mechanisms of historical dispersal and demographic changes

Our data from all gene markers exhibited isolation-by-distance (IBD) like signals. The IBD pattern has been observed in a few vent species from the EPR, such as the amphipod *Ventiella sulfuris*, the mussel *B. thermophilus* (France et al., 1992; Won et al., 2003) and the tubeworms *R. pachyptila* and *T. jerichonana* (Black et al., 1994; Zhang et al., 2015), which was considered to be the result of stepping-stone dispersals (Audzijonyte & Vrijenhoek, 2010). However, most of the positive correlations in vent species might be due to the great geographic gap between the sampled southern and northern populations. Significant barriers can also produce a pattern of genetic differentiation similar to that of IBD (Audzijonyte & Vrijenhoek, 2010; Plouviez et al., 2009, 2010, 2013).

Previous studies have not produced consistent results about the direction of gene flow crossing the 9°N–17°S barrier for vent species inhabiting EPR. For instance, among the seven species studied by Plouviez et al. (2009), four (*B. thermophilus*, *B. symmytilida*, *L. ovalis* and *L. elevatus*) exhibit a high asymmetric gene flow from SEPR to NEPR, two (*A. pompejana* and *E. vitrea*) show nearly no gene flow, and one (*H. bergi*) shows a high asymmetric gene flow from NEPR to SEPR. However, gene flow across the 9°N–17°S barrier is detected in *A. pompejana* in multi-locus studies (Hurtado et al., 2004; Jang et al., 2016; Plouviez et al., 2010). In this study, we detected weak ( $2N_m < 1$ ) but statistically significant bidirectional gene flow across the 9°N–17°S barrier in *H. bergi* s.s. Life-history traits likely determine the dispersal abilities of these vent species. The lack of information on the reproduction and larval development of *Hesiolyra* spp. (Wang et al., 2020) prevents us from inferring their dispersal pattern based on characters like egg buoyancy and larval

development mode (Kinlan et al., 2005), but the genetic data indicate that the larvae of *H. bergi* were able to disperse across the Equatorial barrier. The superfast spreading rate in the SEPR could have caused a high frequency of extinction-recolonization events, thus contributing to the observed demographic instability and decreased genetic diversity in the SEPR (Vrijenhoek, 2010). In the SEPR, Plouviez et al. (2009) detected a bottleneck that occurred at about 0.5 Ma and suspected it to be the result of a large catastrophic/eruptive event that destroyed most of the vent fauna. The bottleneck event was also detected in our study. The effective population size of *H. bergi* has increased in the NEPR compared to their hypothetical ancestral population, but become smaller in the SEPR (Figure 5), supporting the hypothesis of a loss in genetic diversity due to a recent bottleneck in the SEPR populations. The trend of decreasing population size in the SEPR was also observed in *A. pompejana* (Jang et al., 2016; Plouviez et al., 2010). In contrast with most vent species in this region (Plouviez et al., 2009, 2010), our Tajima's *D* and Fu's *F<sub>s</sub>* of *H. bergi* results did not fully support the presence of recent demographic expansions following the bottleneck in the SEPR, suggesting that in this species the recovery of historic effective population size after a bottleneck event is slow.

#### 4.4 | Conservation implications

Deep-sea mining on seafloor massive sulphides has been predicted to cause habitat loss, reduced biodiversity, and declined resilience in ecosystem functions (Levin et al., 2020). An effective environmental management plan needs to be established under the framework of the International Seabed Authority before any exploratory activities can be conducted (Van Dover, 2011). Stakeholders need information on species composition, distribution and population connectivity of vent fauna to understand their vulnerability to the removal of fauna at a particular site or several sites, as well as their likelihood of recovery after the mining impacts. In particular, studies on population connectivity can determine whether a site is a suitable set-aside (i.e., no-mining site) within a network of vents or vent fields by providing information on source-sink dynamics and effective population sizes, as well as connectivity models (i.e., panmixia or isolation-by-distance) (Boschen et al., 2016). In this context, our study has implications on conservation planning of vent communities. For instance, quite a few previous studies using single mitochondrial gene loci revealed no genetic divergence for vent and seep populations distributed across distances of thousands of kilometres (e.g., multiple taxa reviewed by Vrijenhoek, 2010; *Bathymodiulus* mussels by Miyazaki et al., 2013; pectinodontid limpets by Chen et al., 2019), but more recent studies re-examining these populations using genome-wide SNP markers (Xu et al., 2018, 2021) successfully detected subtle genetic differentiation indicative of restricted gene flow between populations and separated demographic histories. Our study, using multi-locus mitochondrial and nuclear genes, detected the recently diverged NEPR and SEPR metapopulations of *H. bergi*, and their different effective

population sizes, supporting a widespread catastrophic habitat disruption (vent extinction/eruptive event, Plouviez et al., 2009) as a force that has created the bottleneck in the SEPR populations. Together with other studies indicating the Equator region as a dispersal filter for several vent species distributed along the EPR, our results suggest that the NEPR and SEPR should be considered as different biogeographic regions. The small migration rates detected in our study imply that quick recolonization between the NEPR and SEPR could not be assumed after disturbance. Further studies should consider the influence of congeneric species in calculation of population genetic parameters, and the use of such results to guide conservation planning.

## 5 | CONCLUSIONS

Our multi-locus data confirmed *H. aff. bergi* as a distinct lineage in SEPR. Of the *H. bergi* populations widely distributed along EPR, a variable barrier to gene flow is present between 9°N–17°S, with the NEPR and SEPR metapopulations being clearly separated for most of the gene markers examined. The genetic diversity in NEPR populations was in general higher than in SEPR populations even if this species was highly abundant in the South. This pattern can be explained by the hypothesis of reduced genetic diversity due to a historical bottleneck (0.5 Mya, Plouviez et al., 2009) in this section of superfast-spreading ridges, and frequent local population extinction (and a more recent bottleneck) due to the re-generation of new vent sites. The low gene flow across the Equatorial barrier since the split of the NEPR and SEPR populations of *H. bergi* 0.27–0.74 Mya implies that quick recolonization could not be assumed after mining disturbance. Thus, the NEPR and SEPR should be targeted as two biogeographic regions in biodiversity conservation. Furthermore, each individual vent site might serve as the only stepping stone during the long-distance dispersal of vent species and be critical in maintaining species distributions. Even removing one vent can significantly impede the gene flow. Further exploration for potentially undiscovered vent fields across the Equator is warranted to refine the results of population gene flow and divergence. Such fine-scale population structure is necessary to develop conservation units and guide the effective management of hydrothermal vent communities.

## ACKNOWLEDGEMENTS

We are grateful to all the chief scientists (F. Lallier, D. Jollivet and N. Le Bris) and “Nautile” and “L’Atalante” crews for their technical support and effort during the cruises HOPE1999, BIOSPEEDO2004 and MESCAL2010. This work was supported by the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (SMSEGL20Sc02) and General Research Fund of Hong Kong’s Research Grants Council (12101021).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

All DNA data obtained in this study have been deposited in GenBank under the accession numbers shown in Table S3.

## ORCID

Yanan Sun  <https://orcid.org/0000-0003-3215-3434>

Ting Xu  <https://orcid.org/0000-0002-5911-6097>

## REFERENCES

- Audzijonyte, A., & Vrijenhoek, R. C. (2010). When gaps really are gaps: Statistical phylogeography of hydrothermal vent invertebrates. *Evolution*, 64, 2369–2384.
- Black, M. B., Lutz, R. A., & Vrijenhoek, R. C. (1994). Gene flow among vestimentiferan tube worm (*Riftia pachyptila*) populations from hydrothermal vents of the eastern Pacific. *Marine Biology*, 120, 33–39.
- Blake, J. A. (1985). Polychaeta from the vicinity of deep-sea geothermal vents in the eastern Pacific. I: Euprosinidae, Phyllostodidae, Hesionidae, Nereididae, Glyceridae, Dorvilleidae, Orbiniidae and Maldanidae. *Bulletin of the Biological Society of Washington*, 6, 67–101.
- Bonnet, E., & de Peer, Y. V. (2002). zt: A software tool for simple and partial Mantel tests. *Journal of Statistical Software*, 7(10), 1–12.
- Boore, J. L., & Brown, W. M. (2000). Mitochondrial genomes of *Galathealinum*, *Helobdella*, and *Platynereis*: Sequence and gene arrangement comparisons indicate that Pogonophora is not a phylum and Annelida and Arthropoda are not sister taxa. *Molecular Biology and Evolution*, 17, 87–106.
- Boschen, R. E., Rowden, A. A., Clark, M. R., Pallentin, A., & Gardner, J. P. A. (2016). Seafloor massive sulfide deposits support unique megafaunal assemblages: Implications for seabed mining and conservation. *Marine Environmental Research*, 115, 78–88.
- Breusing, C., Johnson, S. B., Tunnicliffe, V., & Vrijenhoek, R. C. (2015). Population structure and connectivity in Indo-Pacific deep-sea mussels of the *Bathymodiolus septemdiemum* complex. *Conservation Genetics*, 16, 1415–1430.
- Carr, C. M., Hardy, S. M., Brown, T. M., Macdonald, T. A., & Hebert, P. D. N. (2011). A tri-oceanic perspective: DNA barcoding reveals geographics structure and cryptic diversity in Canadian polychaetes. *PLoS One*, 6(7), e22232.
- Chen, C., Watanabe, H. K., Nagai, Y., Toyofuku, T., Xu, T., Sun, J., Qiu, J.-W., & Sasaki, T. (2019). Complex factors shape phenotypic variation in deep-sea limpets. *Biology Letters*, 15, 20190504.
- Chevaldonné, P., Jollivet, D., Desbruyères, D., Lutz, R. A., & Vrijenhoek, R. C. (2002). Sister-species of eastern Pacific hydrothermal vent worms (Ampharetidae, Alvinellidae, Vestimentifera) provide new mitochondrial COI clock calibration. *Cahiers de Biologie Marine*, 43, 367–370.
- Clement, M., Snell, Q., Walke, P., Posada, D., & Crandall, K. (2002). TCS: Estimating gene genealogies. In *Proceedings 16th International Parallel and Distributed Processing Symposium* (p. 7). IEEE. <https://doi.org/10.1109/ipdps.2002.1016585>.
- Coykendall, D. K., Johnson, S. B., Karl, S. A., Lutz, R. A., & Vrijenhoek, R. C. (2011). Genetic diversity and demographic instability in *Riftia pachyptila* tubeworms from eastern Pacific hydrothermal vents. *BMC Evolutionary Biology*, 11(1), 1–12.
- Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescence. *Trends in Ecology & Evolution*, 24, 332–340.
- Doyle, J. J., & Dickson, E. E. (1987). Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, 36(4), 715–722.
- Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361.

- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567.
- France, S. C., Hessler, R. R., & Vrijenhoek, R. C. (1992). Genetic differentiation between spatially-disjunct populations of the deep-sea, hydrothermal vent-endemic amphipod *Ventrella sulfuris*. *Marine Biology*, 114, 551–559.
- Francheteau, J., Armijo, R., Cheminee, J. L., Hekinian, R., Lonsdale, P., & Blum, N. (1990). 1 Ma East Pacific Rise oceanic-crust and uppermost mantle exposed by rifting in Hess Deep (equatorial Pacific-ocean). *Earth and Planetary Science Letters*, 101, 281–295.
- Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147(2), 915–925.
- Galtier, N., Depaulis, F., & Barton, N. H. (2000). Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics*, 155, 981–987.
- Guinot, D., & Hurtado, L. A. (2003). Two new species of hydrothermal vent crabs of the genus *Bythograea* from the southern East Pacific Rise and from the Galapagos Rift (Crustacea Decapoda Brachyura Bythograeidae). *Comptes Rendus Biologies*, 326(4), 423–439.
- Hasegawa, M., Kishino, H., & Yano, T. A. (1985). Dating of the human-ape splitting by molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2), 160–174.
- Hey, J. (2009). *IMA documentation*. [https://bio.cst.temple.edu/~tuf29449/pdf/Using\\_IMa\\_12\\_17\\_09.pdf](https://bio.cst.temple.edu/~tuf29449/pdf/Using_IMa_12_17_09.pdf)
- Hey, J. (2010). The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. *Molecular Biology and Evolution*, 27, 921–933.
- Hey, J., & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, 167(2), 747–760.
- Hurtado, L. A., Lutz, R. A., & Vrijenhoek, R. C. (2004). Distinct patterns of genetic differentiation among annelids of eastern Pacific hydrothermal vents. *Molecular Ecology*, 13, 2603–2615.
- Jang, S. J., Park, E., Lee, W.-K., Johnson, S. B., Vrijenhoek, R. C., & Won, Y.-J. (2016). Population subdivision of hydrothermal vent polychaete *Alvinella pompejana* across equatorial and Easter Microplate boundaries. *BMC Evolutionary Biology*, 16, 235.
- Johnson, S. B., Krylova, E. M., Audzijonyte, A., Sahling, H., & Vrijenhoek, R. C. (2017). Phylogeny and origins of chemosynthetic vesicomyid clams. *Systematics and Biodiversity*, 15(4), 1–16.
- Johnson, S. B., Won, Y.-J., Harvey, J. B., & Vrijenhoek, R. C. (2013). A hybrid zone between *Bathymodiolus* mussel lineages from eastern Pacific hydrothermal vents. *BMC Evolutionary Biology*, 13, 21.
- Jollivet, D., Chevalloné, P., & Planque, B. (1999). Hydrothermal-vent alvinellid polychaete dispersal in the eastern Pacific. II. A metapopulation model based on habitat shifts. *Evolution*, 53(4), 1128–1142.
- Kinlan, B. P., Gaines, S. D., & Lester, S. E. (2005). Propagule dispersal and the scales of marine community process. *Diversity and Distributions*, 11, 139–148.
- Kureth, C. L., & Rea, D. K. (1981). Large-scale oblique features in an active transform fault, the Wilkes fracture zone near 9° S on the East Pacific Rise. *Marine Geophysical Research*, 5, 119–137.
- Leigh, J. W., & Bryant, D. (2015). popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116.
- Levin, L. A., Amon, D. J., & Lily, H. (2020). Challenges to the sustainability of deep-seabed mining. *Nature Sustainability*, 3(10), 784–794.
- Levin, L. A., Baco, A. R., Bowden, D. A., Colaco, A., Cordes, E. E., Cunha, M. R., Demopoulos, A. W. J., Gobin, J., Grupe, B., Le, J., Metaxas, A., Netburn, A., Rouse, G., Thurber, A., Tunnicliffe, V., Van Dover, C. L., Vanreusel, A., & Watling, L. (2016). Hydrothermal vents and methane seeps: Rethinking the sphere of influence. *Frontiers in Marine Science*, 3, 72.
- Lindenbaum, P. (2015). Jvarkit: Java-based utilities for Bioinformatics. figshare. <https://doi.org/10.6084/m9.figshare.1425030>
- Malingsky, M., Matschiner, M., & Svardal, H. (2020). Dsuite – Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, 21, 584–595.
- Matabos, M., & Jollivet, D. (2019). Revisiting the *Lepetodrilus elevatus* species complex (Vetigastropoda: Lepetodrilidae), using samples from the Galápagos and Guaymas hydrothermal vent systems. *Journal of Maluscan Studies*, 85, 154–165.
- Matabos, M., Plouviez, S., Hourdez, S., Desbruyères, D., Legendre, P., Warén, A., Jollivet, D., & Thiébaud, E. (2011). Faunal changes and geographic crypticism indicate the occurrence of a biogeographic transition zone along the southern East Pacific Rise. *Journal of Biogeography*, 38, 575–594.
- Miller, K. A., Thompson, K. F., Johnston, P., & Santillo, D. (2018). An overview of seabed mining including the current state of development, environmental impacts, and knowledge gaps. *Frontiers in Marine Science*, 4, 418.
- Minh, B. Q., Nguyen, M. A. T., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 30, 1188–1195.
- Miyazaki, J., Beppu, S., Kaio, S., Dobashi, A., Kawato, M., Fujiwara, Y., & Hirayama, H. (2013). Dispersal ability and environmental adaptability of deep-sea mussels *Bathymodiolus* (Mytilidae: Bathymodiolinae). *Open Journal of Marine Science*, 3, 31–39.
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274.
- Nielsen, R., & Wakeley, J. (2001). Distinguishing migration from isolation: A Markov chain Monte Carlo approach. *Genetics*, 158, 885–896.
- Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18, 375–402.
- Palumbi, S., Martin, A., Romano, S., McMillan, W. O., Stice, L., & Grabowski, G. (1991). *The simple fool's guide to PCR, Ver. 2.0*. University of Hawaii.
- Plouviez, S., Faure, B., Le Guen, D., Lallier, F. H., Bierne, N., & Jollivet, D. (2013). A new barrier to dispersal trapped old genetic clines that escaped the Easter Microplate tension zone of the Pacific vent mussels. *PLoS One*, 8(12), e81555.
- Plouviez, S., Le Guen, D., Lecompte, O., Lallier, F. H., & Jollivet, D. (2010). Determining gene flow and the influence of selection across the equatorial barrier of the East Pacific Rise in the tube-dwelling polychaete *Alvinella pompejana*. *BMC Evolutionary Biology*, 10, 220.
- Plouviez, S., Shank, T. M., Faure, B., Daguin-Thiébaud, C., Viard, F., Lallier, F., & Jollivet, D. (2009). Comparative phylogeography among hydrothermal vent species along the East Pacific Rise reveals vicariant processes and population expansion in the South. *Molecular Ecology*, 18, 3903–3917.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- Ramasamy, R. K., Ramasamy, S., Bindroo, B. B., & Naik, V. G. (2014). STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus*, 3(1), 431.
- Reid, J. L. (1997). On the total geostrophic circulation of the South Pacific Ocean: Flow patterns, tracers and transports. *Progress in Oceanography*, 39(4), 263–351.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3), 539–542.

- Rouse, G. W., Cavajal, J. I., & Pleijel, F. (2018). Phylogeny of Hesioniidae (Aciculata, Annelida), with four new species from deep-sea eastern Pacific methane seeps, and resolution of the affinity of *Hesiolyla*. *Invertebrate Systematics*, 32(5), 1050–1068.
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34(12), 3299–3302.
- Sethuraman, A., & Hey, J. (2016). IMA2p – Parallel MCMC and inference of ancient demography under the Isolation with migration (IM) model. *Molecular Ecology Resources*, 16, 206–215.
- Shank, T. M., Fornari, D. J., Von Damm, K. L., Lilley, M. D., Haymon, R. M., & Lutz, R. A. (1998). Temporal and spatial patterns of biological community development at nascent deep-sea hydrothermal vents (9°50'N, East Pacific Rise). *Deep-Sea Research Part II: Topical Studies in Oceanography*, 45, 465–515.
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585–595.
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022–3027.
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673–4680.
- Tunnicliffe, V. (1991). The biology of hydrothermal vents: Ecology and evolution. *Oceanography and Marine Biology: An Annual Review*, 29, 319–407.
- Tunnicliffe, V., Embley, R. W., Holden, J. F., Butterfield, D. A., Massoth, G. J., & Juniper, S. K. (1997). Biological colonization of new hydrothermal vents following an eruption on Juan de Fuca Ridge. *Deep Sea Research Part I: Oceanographic Research Papers*, 44, 1627–1644.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3–new capabilities and interfaces. *Nucleic Acids Research*, 40(15), e115.
- Van Dover, C. L. (2011). Mining seafloor massive sulphides and biodiversity: What is at risk? *ICES Journal of Marine Science*, 68(2), 341–348.
- Van Dover, C. L., Arnaud-Haond, S., Gianni, M., Helmreich, S., Huber, J. A., Jaeckel, A. L., Metaxas, A., Pendleton, L. H., Peterseni, S., Ramirez-Llodraj, E., Steinbergk, P. E., Tunnicliffe, V., & Yamamoto, H. (2018). Scientific rationale and international obligations for protection of active hydrothermal vent ecosystems from deep-sea mining. *Marine Policy*, 90, 20–28.
- Van Dover, C. L., German, C. R., Speer, K. G., Parson, L. M., & Vrijenhoek, R. C. (2002). Evolution and biogeography of deep-sea vent and seep invertebrates. *Science*, 295(5558), 1253–1257.
- Vrijenhoek, R. C. (2010). Genetic diversity and connectivity of deep-sea hydrothermal vent metapopulations. *Molecular Ecology*, 19(20), 4391–4411.
- Wang, Z., Xu, T., Zhang, Y., Zhou, Y., Liu, Z., Chen, C., Watanabe, H. K., & Qiu, J.-W. (2020). Molecular phylogenetic and morphological analyses of the 'monospecific' *Hesiolyla* (Annelida: Hesioniidae) reveal two new species. *Deep-Sea Research Part I: Oceanographic Research Papers*, 166, 103401.
- Woerner, A. E., Cox, M. P., & Hammer, M. F. (2007). Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, 23(14), 1851–1853.
- Won, Y. J., Young, C. R., Lutz, R. A., & Vrijenhoek, R. C. (2003). Dispersal barriers and isolation among deep-sea mussel populations (Mytilidae: *Bathymodiolus*) from eastern Pacific hydrothermal vents. *Molecular Ecology*, 12, 169–184.
- Xu, T., Sun, J., Watanabe, H. K., Chen, C., Nakamura, M., Ji, R., Feng, D., Lv, J., Wang, S., Bao, Z., Qian, P.-Y., & Qiu, J.-W. (2018). Population genetic structure of the deep-sea mussel *Bathymodiolus platifrons* (Bivalvia: Mytilidae) in the Northwest Pacific. *Evolutionary Applications*, 11, 1915–1930.
- Xu, T., Wang, Y., Sun, J., Chen, C., Waanabe, H. K., Chen, J., Qian, P.-Y., & Qiu, J.-W. (2021). Hidden historical habitat-linked population divergence and contemporary gene flow of a deep-sea patellogastropod limpet. *Molecular Biology and Evolution*, 38(12), 5640–5654.
- Yang, Q.-Q., Ip, J. C.-H., Zhao, X.-X., Li, J.-N., Jin, Y.-J., Yu, X.-P., & Qiu, J.-W. (2022). Molecular analyses revealed three morphologically similar species of non-native apple snails and their patterns of distribution in freshwater wetlands of Hong Kong. *Diversity and Distributions*, 28, 97–111.
- Zhang, H. B., Johnson, S. B., Flores, V. R., & Vrijenhoek, R. C. (2015). Intergradation between discrete lineages of *Tevnia jerichonana*, a deep-sea hydrothermal vent tubeworm. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 121, 53–61.

#### BIOSKETCHES

**Leyi Xi**, PhD student at Hong Kong Baptist University, studies marine biology.

**Yanan Sun**, Postdoctoral fellow at the Hong Kong University of Science and Technology, studies the biodiversity, biogeography and evolution of marine invertebrates, especially annelids.

**Jian-Wen Qiu**, Professor at the Hong Kong Baptist University, studies diversity, ecology and evolution of invertebrates.

Author contributions: Y.S., J.-W.Q. and D.J. and S.P. designed the study; D.J. and S.P. collected the samples; L.X. and Y.S. conducted laboratory work; L.X., Y.S. and D.J. conducted data analyses; Y.S., J.-W.Q. and L.X. wrote the first draft. All authors read, revised and approved the manuscript for submission.

#### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Xi, L., Sun, Y., Xu, T., Wang, Z., Chiu, M. Y., Plouviez, S., Jollivet, D., & Qiu, J.-W. (2022).

Phylogenetic divergence and population genetics of the hydrothermal vent annelid genus *Hesiolyla* along the East Pacific Rise: Reappraisal using multi-locus data. *Diversity and Distributions*, 00, 1–15. <https://doi.org/10.1111/ddi.13653>