SCIENTIFIC REP😲RTS

OPEN

Cryptic diversity in the Japanese mantis shrimp *Oratosquilla oratoria* (Crustacea: Squillidae): Allopatric diversification, secondary contact and hybridization

Jiao Cheng & Zhong-li Sha

Mounting evidence of cryptic species in the marine realm emphasizes the necessity to thoroughly revise our current perceptions of marine biodiversity and species distributions. Here, we used mitochondrial cytochrome oxidase subunit I (mtDNA COI) and nuclear ribosomal internal transcribed spacer (nrDNA ITS) to investigate cryptic diversity and potential hybridization in the Japanese mantis shrimp *Oratosquilla oratoria* in the Northwestern (NW) Pacific. Both mitochondrial and nuclear gene genealogies revealed two cryptic species in this morphotaxon, which was further confirmed by extensive population-level analyses. One cryptic species is restricted to cold waters with a distribution range corresponding to temperate affinities, while the other dwelled warm waters influenced by the Kuroshio Current. Their divergence was postulated to be attributable to the vicariant event which resulted from the isolation of the Sea of Japan during the middle Pliocene (*c*. 3.85 Mya, 95% HPD 2.23–6.07 Mya). Allopatric speciation was maintained by limited genetic exchange due to their habitat preferences. Furthermore, the observation of recombinant nrDNA ITS sequence and intra-individual ITS polymorphism suggested recent hybridization event of the two cryptic species occurred in sympatric areas. Our study also illustrated that the Changjiang River outflow might act as an oceanic barrier to gene flow and promoted allopatric diversification in *O. oratoria* species complex.

Marine biodiversity may have always been underestimated due to the occurrence of morphologically indistinguishable cryptic or sibling species[1]. Species has been known to evolve different reproductive and physiological traits without being reflected in the external morphology, adding an extra layer of complexity to the discovery of hidden diversity[2]. The application of molecular methods to species delimitation has uncovered an overwhelming amount of unrecognized cryptic diversity in marine organisms, e.g. copepod[3, 4], barnacles[5, 6], bivalves[7, 8] and fishes[9, 10]. These patterns challenged the paradigm that an apparent homogeneity of the marine environments and pelagic larval stages in marine species limit their diversification[11]. Much progress has been achieved to resolve marine biogeographic patterns and generalize factors involved in shaping marine biodiversity[12, 13]. A growing body of studies has emphasized the importance of climate shifts, or associated sea-level changes caused by glacial cycles, interacted with species-specific ecological traits and life histories as drivers of speciation and diversification in the sea[9, 14–16].

The Northwestern (NW) Pacific is one of the world's largest subduction zones, and its shoreline and sea-basin configuration varied extensively and temporarily[17, 18]. Sea levels declined for 120–140 m during the glacial maximum, leading to the complete closure of the Sea of Japan, the semi-closure of the South China Sea and the partial or full exposure of the East China Sea and Yellow Sea[18, 19]. Meanwhile, land bridges were formed between the continent and Taiwan and between the Korean Peninsula and the main Japanese Islands, which would collectively isolate the three marginal seas[18, 20, 21]. During interglacial periods, marine transgression associated with rising sea level caused the coastline of East China Sea-Yellow Sea to move inland, resulting in the flooding of the East China Sea Shelf. The marginal seas reconnected and the islands were isolated again from the continent.

Institute of Oceanology, Chinese Academy of Sciences, Qingdao, 266071, China. Correspondence and requests for materials should be addressed to Z.-l.S. (email: shazl@qdio.ac.cn)
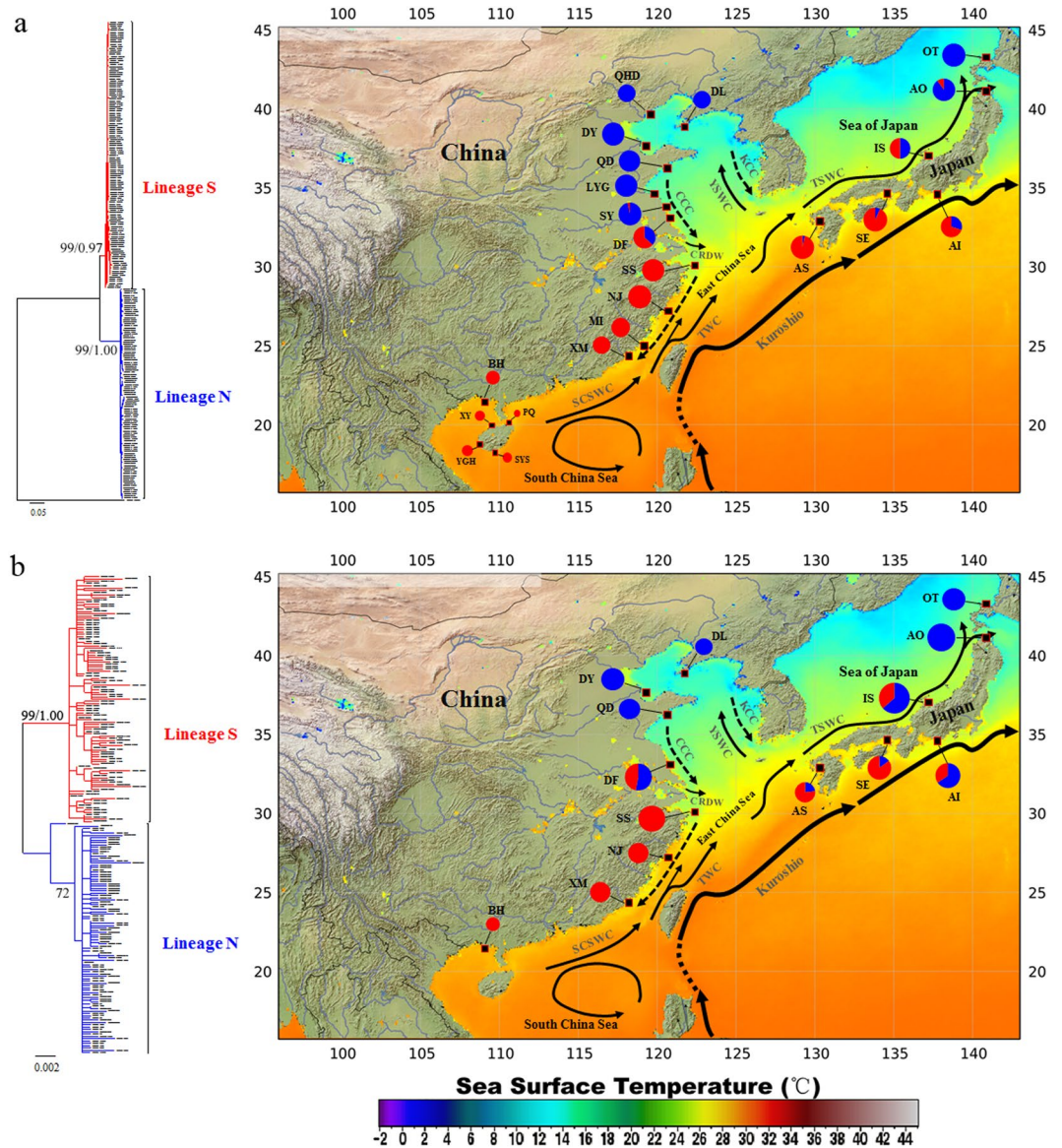
Figure 1. Sample locations and relative frequency of two genetic lineages inferred from mtDNA COI (**a**) and nrDNA ITS (**b**). For each dataset, the corresponding maximum-likelihood (ML) and Bayesian inference (BI) trees are also presented. Numbers above branches represent maximum likelihood bootstrap (left) and Bayesian posterior probabilities (right) for the two major lineages. The map also presented Sea Surface Temperature (SST) in the NW Pacific in winter (2002–2011), which was modified from ref. 78. SCSWC, South China Sea Warm Current; TWC, Taiwan Warm Current; CCC, China Coastal Current; CRDW, Changjiang River Diluted Water; YSWC, Yellow Sea Warm Current; TSWC, Tsushima Warm Current.

These paleogeographical dynamics in the NW Pacific should have enhanced vicariance through glacial geographic isolations and range extension via postglacial dispersal in the marine organisms[9, 22]. This hypothesis has been confirmed by recent phylogeographic investigations in the NW Pacific[16, 23–25]. However, the effects of these evolutionary events in shaping speciation and diversification of marine organisms in this region have not been explored extensively.

Additionally, intricate hydrological systems represented by oceanic currents and freshwater outflow in the NW Pacific may have influenced current population genetic patterns and biogeographic histories of marine species[16, 26–30]. For instance, the Kuroshio Current transports warm and high salinity water northeastward from the Lozon Straits, passes both the east and west coasts of Taiwan and reaches the steep east-west continental shelf break of the East China Sea. It then splits into two branches, the major one flowing northeast towards the southern coast of Kyushu, and the minor one (the Tsushima Warm Current, TSWC) flowing northward into the Sea of Japan[31, 32]. The cold and low-saline China Coastal Current (CCC) flows southwards along the China coast from the Bohai Sea to the Taiwan Strait[33]. The Korean Coastal Current (KCC) connects the East China Sea and the Sea of Japan (Fig. 1). These currents can not only affect dispersal of marine larvae and ultimately population

genetic connectivity, but also generate a heterogeneous landscape boundary for marine organisms. Furthermore, the Changjiang River outflow, also known as Changjiang Diluted Water (CDW), has a significant effect on the salinity, nutrient concentrations and planktonic community of the East China Sea[34, 35]. It has been proposed as a physical barrier for genetic arrangements of coastal species inhabiting the region (e.g. *Trachypenaeus curvirostris*[36]). Thus, the NW Pacific appears to offer an ideal model to disentangle the interaction of contemporary and historical factors driving differentiation and speciation of marine species. One species or a group of closely related species inhabiting marginal seas in the NW Pacific can serve as a model to test these hypotheses.

The Japanese mantis shrimp *Oratosquilla oratoria* (De Haan, 1844) is distributed from Peter the Great Bay, Russia through Japan, Korea and China coastal waters[37, 38] and exhibits habitat preferences from the shore down to coral reefs and level substrates[39]. Its distribution range spans a wide geographical area throughout tropical, subtropical and temperate coastal areas and covers contrasting environmental features, providing an ideal system to investigate local adaptation and adaptive radiation. *O. oratoria* is subjected to intense and unregulated inshore fisheries in the NW Pacific because of its abundance and wide consumer acceptance, especially in Asian countries[40, 41]. Excessive exploitation has led to severe damage of the natural resources of *O. oratoria*[42]. Thus, an applicable strategy for conservation and management of this over-exploited species is urgent, and the precise knowledge about its taxonomy and population structure is the prerequisite. Previous studies have revealed a contrasting divergence pattern of *O. oratoria* in China Sea. Zhang *et al.*[43, 44] found a significant north-south population structure contributed by long-term isolation of the Taiwan Strait during glacial periods. Conversely, Du *et al.*[45] detected a genetic break between the East China Sea and Yellow Sea which is in accordance with the Changjiang River discharge based on mitochondrial evidence. Nevertheless, these afore-mentioned studies were constrained by small sampling sizes, large geographical gaps in experimental design and inadequate molecular data, which become an impediment to precisely evaluate population genetic structure and cryptic speciation of *O. oratoria*.

Given all uncertainties mentioned above, this study aimed at elucidating a detailed genetic pattern in *O. oratoria* with extensive sampling and revealing the existence of cryptic species. If a species complex did exist, we further attempted to (i) address the driving forces of cryptic speciation; (ii) investigate whether hybridization occurred among cryptic species; (iii) determine the factors influencing current genetic architecture and biogeographic distribution of each cryptic species. To this end, we sequenced mitochondrial cytochrome oxidase subunit I (mtDNA COI) and nuclear ribosomal internal transcribed spacer (nrDNA ITS) sequences from among 22 *O. oratoria* populations across its distribution range. These results can enhance our understanding of how paleoclimate changes and environmental heterogeneity contributed to speciation and diversification of coastal species in the NW Pacific.

Results

Phylogeographical structure. *mtDNA COI.* Aligned 658 bp of mtDNA COI gene comprised 144 polymorphic sites which defined 236 haplotypes in 498 specimens. A total of 142 transitions and 16 transversions were scored and no indels were found. All mtDNA COI haplotypes were deposited in GenBank under Accession nos. KY197015–KY197250.

MtDNA COI-based ML and Bayesian trees yielded identical topology and supported two highly-diverged monophyletic lineages in *O. oratoria* (hereafter referred as northern and southern lineages, labelled N and S, Fig. 1a), suggesting two candidate cryptic species. The minimum spanning tree (MST) constructed from mtDNA COI also retrieved two distinct haplogroups. A total of 33 bp of substitutions were required to connect them. For each lineage, common haplotypes were found in the center that produced multiple star-like polytomies with relatively shallow genetic divergences (Fig. 2a). The haplotype network of each lineage indicated no clustering of haplotypes corresponding to localities. The pairwise uncorrected *p*-distances between the two lineages was 6%, whereas the mean genetic distance between haplotypes within each lineage was 0.7% and 0.8%, respectively. Apparently, this satisfied the "4 × rule"[46] with inter-lineage distances exceeding 4 × intra-lineage distances.

The two divergent mtDNA COI lineages were also geographically structured (Fig. 1a). Overall, lineage N (104 haplotypes, 249 individuals) was dominant in the temperate zone of the NW Pacific, including the northern and central parts of the China coast (the Bohai and Yellow Seas) as well as the northern Japan Sea. By contrast, lineage S (132 haplotypes, 249 individuals) was mainly distributed in subtropical and tropical waters, including the East and South China Seas, and decreased dramatically in frequency along the southern Pacific coast of Japan. These two lineages were sympatric in their range with the southern Yellow Sea and south coast of Japan as overlapped zones (DF, IS and AI).

nrDNA ITS. A subset of 147 individuals from 14 populations were genotyped for nrDNA ITS to check the deep phylogenetic break observed by mtDNA COI. After exclusion of regions with mononucleotide and microsatellite repeats, the aligned nrDNA ITS ranged from 1171 bp to 1220 bp and had 307 polymorphic sites which defined 207 ribotypes in 295 clones. Three indels with sizes varying from 4 bp to 48 bp were inferred. According to the simple insertion/deletion coding method[47], three binary coded gap characters were appended at the end of the combined matrix for phylogenetic analyses. Using the seven algorithms in RDP4, we detected at least one recombination event in nrDNA ITS. Ten of the involved eleven recombinant sequences resembled the ITS1 version of lineage S while ITS2 was concordant to a typical ITS2 version of lineage N, which might be attributed to hybridization succeeded by intralocus recombination in ITS. Recombinant sequences were excluded from subsequent analysis because recombination may produce sequence regions with different evolutionary histories, ultimately impact phylogenies reconstruction and estimations of population genetic parameters[48]. All nrDNA ITS ribotypes were deposited in GenBank under Accession nos. KY197251–KY197457.

NrDNA ITS phylogenies produced concordant lineage structure with mtDNA COI (Fig. 1b), suggestive of two distinctive evolutionary lineages. These two lineages were further supported by the MST (Fig. 2b) which was
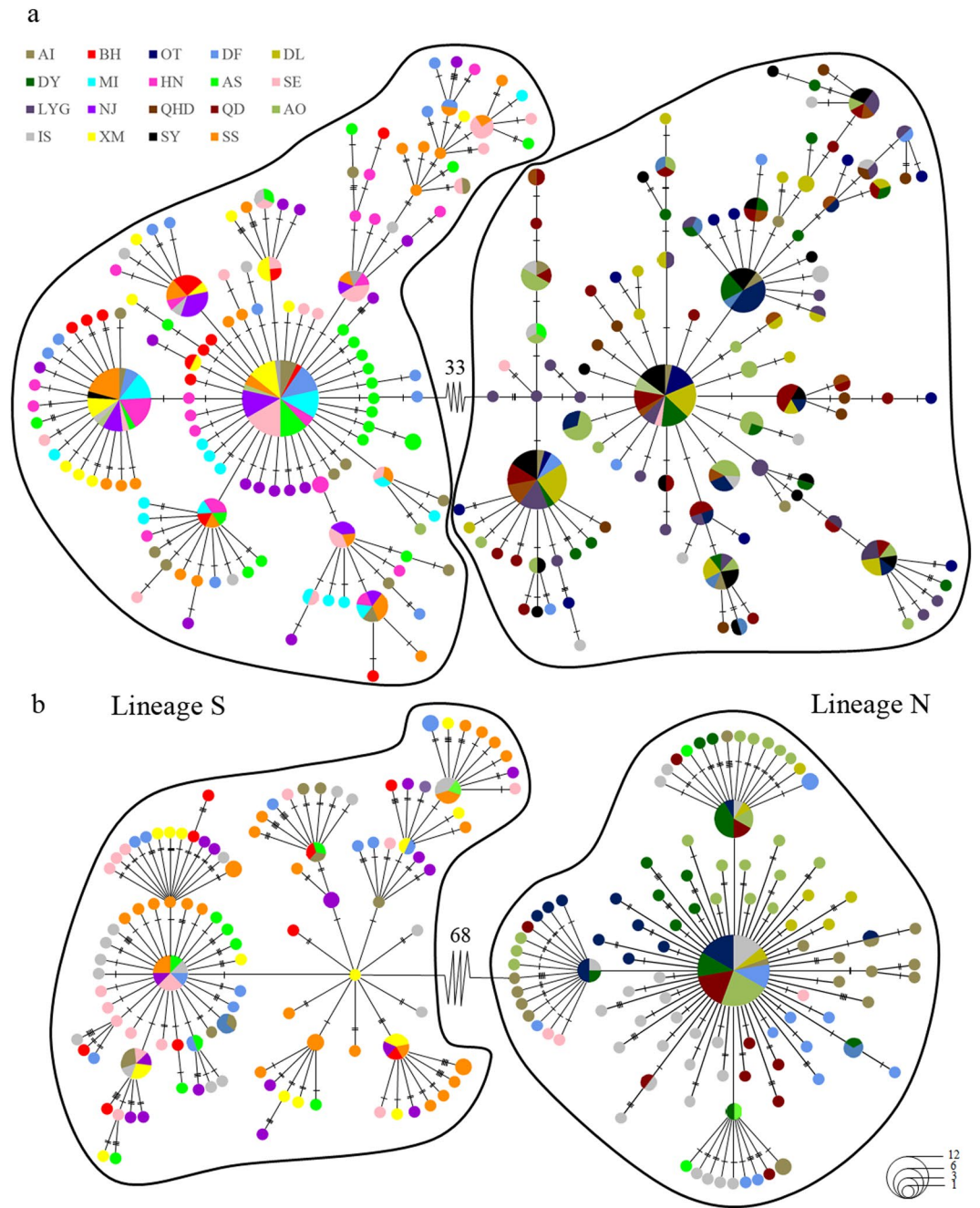
Figure 2. Minimum spanning trees constructed from mtDNA COI (**a**) and nrDNA ITS (**b**) with the sizes of circles proportionally to haplotype/ribotype frequency. The population origins of haplotypes/ribotypes are indicated by colours. Perpendicular tick marks on the lines joining haplotypes/ribotypes represent the number of nucleotide substitutions.

separated by 68 mutational steps. The MST topology of lineage N was characterized by a star-like polytomy with a dominant haplotype shared by all populations from Japan and northern of China, whereas two star-like polytomies were presented in lineage S. No distinct phylogeographic apportioning of localities was observed in each lineage. For nrDNA ITS, sequence divergence between lineages was 1.6%, and the intra-lineage level of divergence averaged 0.5% and 0.4%, respectively. Surprisingly, intra-individual ITS polymorphism was observed in 17 individuals (Supplementary Table S1), indicating possible past hybridization and introgression between these two evolutionary lineages. Intra-individual sequence divergence of nrDNA ITS was up to 1.8% (within sample AI20, IS11 and IS19), while the maximum intra-individual genetic distance between clones nested in the same lineage was 0.8% (within CS4). A similar pattern of spatial partitioning of ribotypes was observed in nrDNA ITS dataset, with exception that all ribotypes of population AO belonged to lineage N (Fig. 1b).

Sample site	Abb	Date of collection	mtCOI					nrITS				
			N_c	n	h	π	L	N_i	n	h	π	L
[1] Otaru	OT	2015.06	30	20	0.947 ± 0.027	0.005 ± 0.003	N	18	12	0.895 ± 0.065	0.002 ± 0.001	N
[2] Aomori	AO	2015.10	30	18	0.959 ± 0.018	0.016 ± 0.008	N, S	27	19	0.917 ± 0.047	0.002 ± 0.001	N
[3] Qinhuangdao	QHD	2014.05	20	18	0.984 ± 0.024	0.006 ± 0.004	N					
[4] Dalian	DL	2014.05	30	17	0.933 ± 0.028	0.005 ± 0.003	N	8	7	0.964 ± 0.077	0.002 ± 0.001	N
[5] Dongying	DY	2014.11	23	17	0.961 ± 0.027	0.005 ± 0.003	N	19	12	0.906 ± 0.050	0.002 ± 0.001	N
[6] Qingdao	QD	2014.05	30	23	0.977 ± 0.015	0.005 ± 0.003	N	16	10	0.867 ± 0.079	0.002 ± 0.001	N
[7] Lianyungang	LYG	2015.09	30	23	0.977 ± 0.016	0.005 ± 0.003	N					
[8] Sheyang	SY	2015.05	30	20	0.961 ± 0.020	0.009 ± 0.005	N, S					
[9] Dafeng	DF	2015.05	30	24	0.972 ± 0.021	0.035 ± 0.018	N, S	32	25	0.980 ± 0.015	0.010 ± 0.005	N, S
[10] Ishikawa	IS	2013.10	20	18	0.990 ± 0.019	0.037 ± 0.019	N, S	44	39	0.988 ± 0.010	0.009 ± 0.005	N, S
[11] Aichi	AI	2015.10	20	18	0.984 ± 0.024	0.032 ± 0.016	N, S	26	24	0.994 ± 0.013	0.010 ± 0.005	N, S
[12] Setonaikai	SE	2015.10	30	19	0.936 ± 0.032	0.013 ± 0.007	N, S	20	19	0.995 ± 0.018	0.007 ± 0.004	N, S
[13] Ariake Sea	AS	2015.02	30	25	0.975 ± 0.021	0.010 ± 0.005	N, S	13	13	1.000 ± 0.030	0.008 ± 0.005	N, S
[14] Shengshan	SS	2015.07	30	23	0.961 ± 0.027	0.007 ± 0.004	S	31	26	0.989 ± 0.011	0.005 ± 0.003	S
[15] Nanji	NJ	2014.06	30	20	0.954 ± 0.023	0.006 ± 0.003	S	16	15	0.992 ± 0.025	0.004 ± 0.003	S
[16] Meizhou Island	MI	2014.04	22	15	0.931 ± 0.040	0.005 ± 0.003	S					
[17] Xiamen	XM	2014.04	20	13	0.926 ± 0.043	0.005 ± 0.003	S	16	14	0.983 ± 0.028	0.004 ± 0.002	S
[18] Beihai	BH	2015.07	16	14	0.975 ± 0.035	0.006 ± 0.003	S	9	9	1.000 ± 0.052	0.006 ± 0.003	S
[19] Xinying	XY	2007.11	10	9	0.978 ± 0.054	0.006 ± 0.004	S					
[20] Puqian	PQ	2007.12	3	2	0.677 ± 0.314	0.003 ± 0.003	S					
[21] Yinggehai	YGH	2007.12	8	8	1.000 ± 0.063	0.004 ± 0.003	S					
[22] Sanya	SYS	2008.03	6	5	0.933 ± 0.122	0.006 ± 0.004	S					
Total			498	236	0.981 ± 0.002	0.032 ± 0.016		295	207	0.982 ± 0.005	0.032 ± 0.015	

Table 1. Sample information and molecular diversity indices for *O. oratoria*. Abb., abbreviation of locality; N_c, number of sequenced individuals; n, number of haplotypes; N_i, number of clones; h, haplotype diversity; π, nucleotide diversity; L, lineage found in each locality.

Genetic diversity and differentiation of *O. oratoria*.

O. oratoria populations exhibited high haplotype diversity with overall value of 0.981 for mtDNA COI and 0.982 for nrDNA ITS (Table 1). The hybrid populations where two lineages coexisted had higher nucleotide diversity (mtCOI: 0.032–0.037; nrITS: 0.009–0.01) than other populations (mtCOI: 0.003–0.016; nrITS: 0.002–0.008).

Based on the distribution of haplotypes, we defined two distribution regions for the sampled *O. oratoria* populations: (T) temperate region (including the Bohai and Yellow Seas as well as the northern Japan Sea; site1–10), and (S) subtropical and tropical region (including the East and South China Seas as well as the southern Pacific coast of Japan; site 11–22). Pairwise F_{ST} values between populations of different regions were high (0.749–0.928 for mtDNA COI; 0.745–0.801 for nrDNA ITS) and statistically significant after Bonferroni corrections for both markers (Table 2). Moderate and significant F_{ST} values (0.127–0.653 for mtDNA COI; 0.094–0.39 for nrDNA ITS) were also observed for pairwise comparisons between hybrid populations (site 9–11) and other populations. In addition, most of the pairwise F_{ST} values between populations within each region were low and statistically non-significant. Further analysis by hierarchical AMOVA indicated that for mtDNA COI 72.32% of genetic variation was partitioned by regions, while for nrDNA ITS 38.72% of genetic variation occurred among regions (Table 3). An AMOVA analysis that excluded the hybrid populations showed more suitable variance partitioning with higher percentage of variance (88.24% for mtDNA COI, 71.28% for nrDNA ITS) among regions (P < 0.001, Table 3). Genetic subdivision was highly significant between regions ($F_{CT} = 0.882$, $P < 0.001$ for mtCOI; $F_{CT} = 0.713$, $P < 0.001$ for nrITS), indicating a high level of geographical population structure.

Divergence between lineages.

The pattern of reciprocal monophyly, deep phylogenetic divergence between lineages and limited genetic exchange between populations detected in *O. oratoria* indicated the existence of two cryptic species across its distribution. For each location, individuals belonging to the same lineage were grouped together to further explore the genetic structure of the two cryptic species. Population genetic differentiation between lineages was highly significant ($F_{ST} = 0.916$, $P < 0.001$ for mtCOI; $F_{ST} = 0.722$, $P < 0.001$ for nrITS). No genetic heterogeneity was detected among samples assigned to each of these two lineages (Supplementary Table S2). However, the generality of these results may be limited by the relatively low sample sizes of AI-N, SE-N and AS-N when compared with populations with a larger sample size. The calibrated divergence time between the two lineages was 3.85 Million Years ago (Mya) (95% HPD 2.23–6.07 Mya), dating back to the middle Pliocene. The coalescence time was estimated at 0.6 Mya (95% HPD 0.33–0.87 Mya) and 0.62 Mya (95% HPD 0.36–0.92 Mya) for lineage N and lineage S, respectively.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]
[1] OT	—	0.017		0.041	0.047*	0.017			0.254**	0.159**	0.172*	**0.543****	**0.549****	**0.745****	**0.792****		**0.801****	**0.792****
[2] AO	0.041	—		−0.016	−0.004	−0.010			0.286**	0.185**	0.210**	**0.572****	**0.577****	**0.755****	**0.791****		**0.799****	**0.792****
[3] QHD	−0.002	0.029	—															
[4] DL	0.019	0.036	0.004	—	−0.020	−0.028			0.214*	0.133*	0.144*	**0.499****	**0.475****	**0.721****	**0.760****		**0.772****	**0.743****
[5] DY	−0.012	0.033	−0.001	0.009	—	−0.002			0.264**	0.171*	0.192**	**0.546****	**0.542****	**0.741****	**0.778****		**0.788****	**0.774****
[6] QD	0.025*	0.036	−0.006	−0.009	0.02	—			0.247**	0.156*	0.175**	**0.536****	**0.531****	**0.738****	**0.779****		**0.789****	**0.775****
[7] LYG	0.017	0.033	−0.001	0.004	0.008	0.002	—											
[8] SY	0.001	−0.001	−0.004	−0.004	−0.013	−0.004	−0.01	—										
[9] DF	0.542**	0.400**	0.492**	0.543**	0.508**	0.540**	0.530**	0.487**	—	0.000	−0.002	0.094*	0.053	0.309**	0.275**		0.290**	0.254**
[10] IS	0.461**	0.287**	0.400**	0.463**	0.420**	0.459**	0.447**	0.392**	−0.011	—	−0.008	0.176**	0.124*	0.390**	0.366**		0.380**	0.348**
[11] AI	0.652**	0.492**	0.599**	0.653**	0.617**	0.648**	0.639**	0.591**	−0.029	0.020	—	0.156**	0.109*	0.379**	0.348**		0.364**	0.320**
[12] SE	**0.856****	**0.749****	**0.835****	**0.857****	**0.842****	**0.852****	**0.848****	**0.817****	0.179*	0.320**	0.127**	—	−0.025	0.101**	0.062*		0.079*	0.045
[13] AS	**0.886****	**0.783****	**0.871****	**0.888****	**0.877****	**0.883****	**0.879****	**0.849****	0.222**	0.371**	0.170**	−0.003	—	0.137**	0.103**		0.123**	0.068
[14] SS	**0.913****	**0.814****	**0.901****	**0.914****	**0.906****	**0.909****	**0.906****	**0.877****	0.265**	0.426**	0.230**	0.007	0.021*	—	0.013		0.016	−0.007
[15] NJ	**0.920****	**0.822****	**0.910****	**0.921****	**0.914****	**0.916****	**0.914****	**0.885****	0.274**	0.435**	0.237**	0.019	0.002	0.014	—		−0.028	−0.016
[16] MI	**0.927****	**0.817****	**0.917****	**0.928****	**0.921****	**0.922****	**0.920****	**0.887****	0.255*	0.412**	0.219**	0.015	−0.001	−0.009	−0.008	—		
[17] XM	**0.924****	**0.810****	**0.913****	**0.926****	**0.918****	**0.920****	**0.917****	**0.883****	0.244**	0.394**	0.212**	0.036	0.012	0.024	0.011	0.008	—	−0.007
[18] BH	**0.923****	**0.800****	**0.910****	**0.925****	**0.916****	**0.918****	**0.915****	**0.878****	0.221**	0.366**	0.182*	0.011	−0.012	0.004	−0.021	−0.015	−0.011	—
[19–22] HN	**0.924****	**0.821****	**0.914****	**0.925****	**0.918****	**0.919****	**0.917****	**0.887****	0.265**	0.428**	0.229**	0.019	0.009	−0.014	−0.001	−0.018	0.015	−0.015

Table 2. Pairwise F_{ST} comparisons between populations inferred from mtDNA COI (below diagonal) and nrDNA ITS (above diagonal). *$P < 0.05$; **$P < 0.001$.

	mtDNA COI				nrDNA ITS			
Structure tested	df	Variance	%total	F statistics	df	Variance	%total	F statistics
Two gene pools	(Group I: 1–10; II: 11–22)				(Group I: 1–2, 4–6, 9–10; II: 11–15, 17–18)			
Among regions	1	18.35	72.32	$F_{CT} = 0.723^{**}$	1	2.79	38.72	$F_{CT} = 0.387^{*}$
Among populations with region	16	1.32	7.15	$F_{SC} = 0.258^{**}$	12	0.81	11.26	$F_{SC} = 0.184^{*}$
Wthin populations	450	3.79	20.53	$F_{ST} = 0.795^{**}$	281	3.60	50.02	$F_{ST} = 0.500^{**}$
Two gene pools	(Group I: 1–8; II: 12–22)				(Group I: 1–2, 4–6; II: 12–15, 17–18)			
Among regions	1	18.55	88.24	$F_{CT} = 0.882^{**}$	1	6.34	71.28	$F_{CT} = 0.713^{**}$
Among populations with region	13	0.02	0.12	$F_{SC} = 0.010^{*}$	9	0.07	0.83	$F_{SC} = 0.029^{*}$
Wthin populations	383	2.45	11.64	$F_{ST} = 0.884^{**}$	182	2.48	27.89	$F_{ST} = 0.721^{**}$

Table 3. Results of analysis of molecular variance (AMOVA) for different hierarchical analyses of *O. oratoria* populations. *$P < 0.05$; **$P < 0.001$.

					Tajima's D		Fu's Fs		Mismatch distribution			
Groups	N	n	h	π	D	P	Fs	P	θ_0	θ_1	P(SSD)	P(HRI)
mtCOI												
Lineage N	249	104	0.969 ± 0.005	0.005 ± 0.003	−2.119	0.000	−25.842	0.000	0.689	9999	0.376	0.548
Lineage S	249	132	0.954 ± 0.008	0.006 ± 0.003	−2.345	0.000	−25.654	0.000	1.182	68.672	0.965	0.727
nrITS												
Lineage N	156	100	0.941 ± 0.015	0.003 ± 0.001	−2.865	0.000	−26.409	0.000	0.682	271.895	0.910	0.676
Lineage S	139	107	0.992 ± 0.003	0.005 ± 0.002	−2.650	0.000	−25.135	0.000	0.796	121.719	0.782	0.689

Table 4. Summary of molecular diversity and demographic analyses for two lineages of *O. oratoria*. Number of individuals (*N*), number of haplotypes or ribotypes (*n*), haplotype diversity (*h*), nucleotide diversity (*π*), Tajima's *D* and Fu's F_S with corresponding P values for each lineage are shown. θ_0 and θ_1 are θ parameter before and after expansion; *P* values for the sum of squared deviations (SSD) and raggedness index (HRI) under the hypothesis of sudden expansion of each lineage are also shown.

Historical demography. Tajima's *D* and Fu's F_S calculated for each lineage were negative and statistically significant (Table 4), indicating recent demographic expansions. The sequential mismatch distributions of the two lineages (Fig. 3) and the permutation tests with the SSD and HRI statistics (Table 4) all suggested a unimodal distribution, providing further evidence of population expansion. BSP analysis indicated that two lineages have
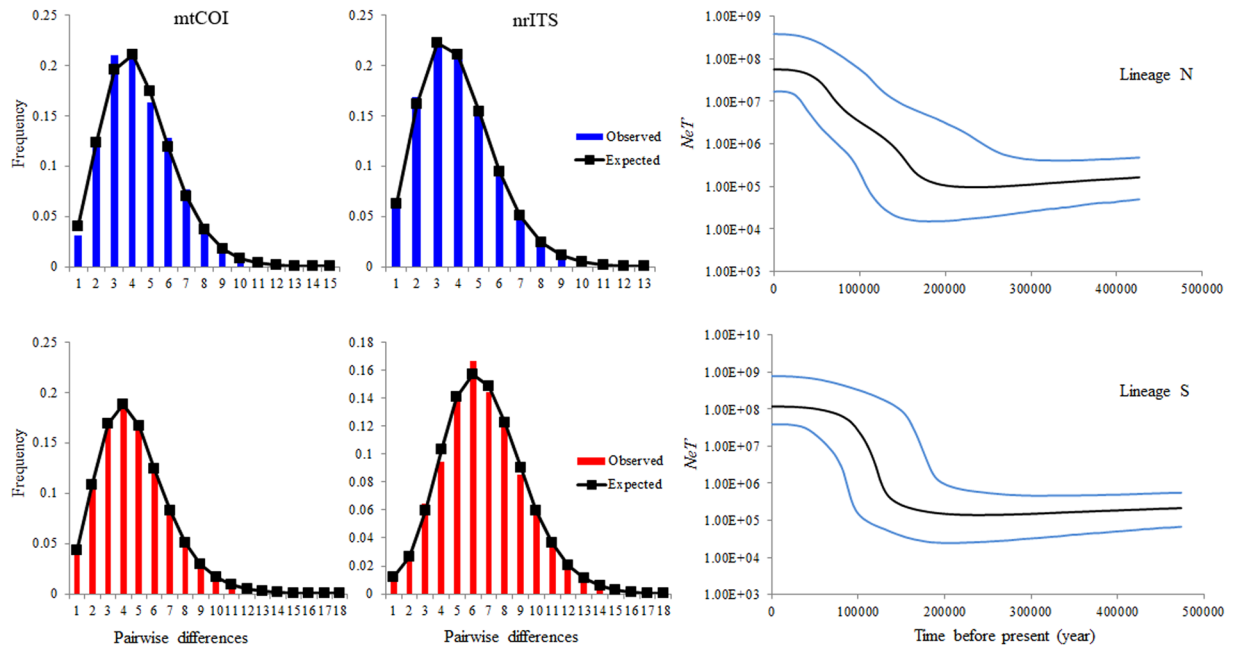
Figure 3. Mismatch distributions and Bayesian skyline plots for each genetic lineage. In mismatch distributions, bars represent the observed frequency of pairwise differences, whereas the solid lines show the expected values under the sudden expansion model. The BSPs show N_eT (N_e = effective population size; T = generation time) changes through time. Black lines represent median estimates while the blue lines show the 95% highest posterior density (HPD) limits.

experienced exponential demographic increases in the late Pleistocene. Specifically, lineage N experienced expansion at *c*. 200 000 years ago, whereas lineage S was dramatic in its increase in both rate and timing and showed a more recent population expansion at *c*. 180 000 years ago (Fig. 3).

Discussion

Allopatric diversification and biogeographic processes. Two major genealogical lineages and a strong phylogeographical structure were detected in our mtDNA analyses of *O. oratoria*. The two divergent lineages may represent vicariant relicts of an ancestral population which was isolated in two allopatric areas during the middle Pliocene. The geographical distribution of lineage S implies an origin in the South China Sea. Although it is difficult to identify potential origin for lineage N due to the lack of samples from Okinawa, its broad distribution in the Bohai Sea, Yellow Sea and northern Japan Sea, in combination with the dated ages concur in suggesting a possible origin in the Sea of Japan. As a composite marginal sea in the NW pacific, the Sea of Japan is a semi-enclosed sea and it connects with nearby water bodies through narrow straits that are shallow and less than 130 m deep. During the middle Pliocene, the Sea of Japan was strongly influenced by a cold northern surface water until 3.5 Mya when the warm Tsushima Current first reappeared corresponding to the reopening of the Tsushima Strait[49, 50]. Meanwhile, the South China Sea was isolated from the East China Sea by landmasses formed between continental China and Taiwan[51]. This process may have been the powerful vicariant event for allopatric differentiation of the ancestral *O. oratoria* populations. During each glacial episode, ancestors of the lineage N that were trapped in the Sea of Japan probably evolved to adapt in cold water environment. By contrast, relatively high-temperature adaptation may have evolved in ancestors of lineage S. In reality, similar phylogeographic pattern was also observed in several other marine species in this region, such as two varieties of *Sargassum*[52], flathead mullet *Mugil cephalus*[9] and the bivalve *Cyclina sinensis*[24], although the magnitude of genetic subdivision varies among species. Taken together, complex coastal topography and habitat patchiness in the NW Pacific driven by the recurrent drops in sea level over the past few million years probably structured biogeographic shifts of marine organisms at various temporal and spatial scales[53].

Several of our statistical analyses yielded signatures of major demographic expansions for both lineages in *O. oratoria* in the late Pleistocene (Table 4, Fig. 3). Glacial-interglacial climate fluctuations during the Pleistocene epoch led to changes of sea level and then caused habitat contractions or expansions, which might have reinforced ancient vicariance and postglacial dispersal of *O. oratoria*. The rising sea-level and temperature driven by climate changes during the interglacial period might have brought drastic events of postglacial marine transgression in the Yellow and East China seas due to the shallow continental shelves[18, 54], generating large amounts of shallow coastal habitat for *O. oratoria*. In this context, demographic expansions and range extension of ancestral *O. oratoria* populations would have accompanied a series of colonizing events in previously unoccupied habitats. Specifically, the Bohai and Yellow Seas were colonized by lineage N from a northern route of dispersal through the Tsushima Strait, whereas the East China Sea and the Pacific coast of Japan were recolonized by lineage S from southern route that followed the Kuroshio Current. Lastly, sympatric distribution of two lineages

in the Sea of Japan might ascribe to secondary contact sometime after a geographically northward expansion or range shifts of lineage S that was facilitated by the Tsushima Current. This scenario is also supported by the observation that warm-water planktonic foraminifera and mollusks migrated into the Sea of Japan under the same effects of the Tsushima Current[55]. Although glacial influence to population fluctuation was observed in late Pleistocene, it seems population size of *O. oratoria* was persistent through the Last Glacial Maximum (LGM, 26 000–19 000 years ago). Previous paleoclimate evidence demonstrated that the most substantial global glacial extension occurred in the Marine Isotope Stages 16–18 (MIS 16-MIS 18, 0.6–0.7 Mya)[56, 57]. Since then, environmental changes seem to be moderate in subsequent climate oscillations. Recent analyses of demographic history of marine invertebrate species in the NW Pacific[58, 59] also converged on similar findings that majority of the species had potential abilities to survive the LGM and regional persistence maybe more prevalent.

Cryptic species and hybridization in *O. oratoria* complex. Two reciprocally monophyletic, highly supported lineages together with large inter-lineage relative to intra-lineage divergences recovered from mtDNA COI data allowed us to propose the existence of two candidate cryptic species under the "4 × rule" for species detection[46]. Although initially designed for asexual organisms, the "4 × rule" can also be applied to mitochondrial genes in sexual organisms and has become a promising tool for cryptic species delimitation in crustaceans[60, 61]. In this study, the divergence level between *O. oratoria* lineages (6%) is comparable to that calculated between some cryptic species of stomatopods species (*Haptosquilla pulchella*: 2.9–6.6%; *Haptosquilla glyptocercus*: 4.7–11.7%; *Gonodactylellus viridis*: 3.0%–10.6%[62]) or barnacle species (*Chthamalus moro*: 3.9–8.3%[6]). This COI divergence value also exceeded the COI species screening threshold (5% divergence) applied for stomatopods[63], a value also supported by data from alpheid shrimp in which most sister species showed more than 5% sequence divergence[49]. Nuclear evidence of two monophyletic lineages in combination with a relatively long period of divergence between two *O. oratoria* lineages (since middle Pliocene) concur in suggesting that lineage sorting of ancestral polymorphisms has been completed. Moreover, the clear allopatric distribution of the two lineages also tend to support the presence of cryptic species: lineage N was almost exclusively temperate whereas lineage S was mostly tropical and subtropical and extended into temperate regions of Japan despite being found in sympatry at their geographical boundaries. Lastly, further evidence for species distinctiveness is that the gene flow appeared to be more limited among sympatric individuals belonging to different lineages than between geographically distant individuals sharing the same lineage (Supplementary Table S2). Given the marked genetic isolation and limited gene flow between natural populations of lineage N and S as revealed from evidence of both mitochondrial and nuclear gene, we argue that they could actually constitute different species.

Natural hybridization occurs relatively frequently when previously allopatric populations or species achieve secondary contact[64]. Unless reproductive isolation is complete, any secondary contact would be followed by hybridization[65, 66]. In this study, we identified species-specific nuclear ribotypes as they are highly divergent and exclusively present in allopatric populations of one cryptic species. Presence of such species-specific alleles in an individual genome reflected either a past hybridization or a gene duplication event[67]. For nrDNA ITS sequencing, approximately half samples were sequenced with more than two clones, and most of them were represented by more than two alleles, indicating that the intragenomic variance observed in *O. oratoria* could not be an occasional gene duplication event. Instead, the observation of recombinant nrDNA ITS sequence, intra-individual polymorphism of nrDNA ITS in 17 individuals and hybrids sharing their nrDNA ITS ribotypes with nonhybridization individuals all suggested recent past hybridization has occurred. However, absence of morphological divergence between the two *O. oratoria* cryptic species makes it difficult to decipher the direction and magnitude of introgression. In our study, most putative hybrids have been observed to possess two types of nrDNA ITS ribotypes, implying bidirectional nuclear introgression. The three hybrids in the population AO showed an exceptional trend; all their nrDNA ITS ribotypes (28 clones in total) belonged to lineage N, but their mtDNA COI sequences were all grouped into lineage S (Supplementary Table S1). A reasonable explanation was the mitochondrial introgression from lineages S to N. Further studies are warranted to address the rate and pattern of hybridization, which in return can offer further insights into the origin and driving forces of hybridization between *O. oratoria* species complex.

Population structure and distribution pattern. Recent studies reported different pattern of genetic differentiation of *O. oratoria* populations in China Sea and identified two genetic groups as a consequence of long-term isolation by the Taiwan Strait[43, 44]. Instead, our study detected a sharp genetic discontinuity among China coastal populations in line with the Changjiang Estuary, a finding generally consistent with the mtDNA analysis by Du *et al.*[45]. In the sea, it is well recognized that the movements of pelagic stages are directly influenced by ocean currents that can lead to dispersal of hundreds of kilometres[68, 69]. As benthic habitat for adults, genetic exchanges among *O. oratoria* populations mainly occur in the planktonic larval stage. Laboratory-reared larvae of *O. oratoria* have a planktonic stage of between 36 and 59 days[70], indicating high potential for dispersal during its planktonic larval stage. Thus, the lack of genetic structure in each *O. oratoria* cryptic species is not surprising because oceanographic current systems in the NW Pacific are expected to facilitate larvae dispersal over considerable distances, eventually homogenizing widely spaced coastal populations. On the other hand, dynamic oceanographic processes, such as freshwater outflow, local gyres, upwelling systems and fronts, can also profoundly restricted connectivity among populations. Mounting genetic evidence shows that limited connectivity was found in coastal species despite long-lived pelagic larvae, which generally arise from restricted oceanographic exchange among geographically proximate locations[29, 71, 72]. The Changjiang River outflow influences its surrounding hydrological condition and causes a decline in the salinity of the upper layer of the Kuroshio Current[73], which may have some impacts on gene flow between the Yellow Sea and East China Sea. For *O. oratoria*, limited larval transport influenced by the Changjiang River outflow may act as a mechanism promoting allopatric diversification, and may explain the observed genetic discontinuity along the China Coast. Similar genetic breaks

that resulted from the Changjiang River freshwater discharge have been reported in some other marine species, including crustacean *Trachypenaeus curvirostris*[36], macroalga *Sargassum hemiphyllum*[52] and gastropod *Cellana toreuma*[74]. However, the barrier effect of the Changjiang River outflow on gene flow of costal species might have weakened as a consequence of global warming[59]. The present-day occurrence of lineage S in the southern of Yellow Sea probably results from its northward expansion as the frontal zone between the Kuroshio Current and the Changjiang River outflow shifts northwards according to the strength of the Kuroshio Current[75, 76].

A well-characterized natural environmental gradient in the NW Pacific is its steep thermal cline. The distributional ranges of the two *O. oratoria* cryptic species closely parallel the spatial distribution pattern of Sea Surface Temperature (SST) in the NW Pacific governed by the oceanic currents system in this region (Fig. 1). The temperature species inhabits regions influenced by cold China Costal Current and Korean Coastal Current which facilitate cold water flowing southward from the Yellow Sea with average of water temperature from 12 °C in March to 20 °C in September[77]. By contrast, the distribution of the subtropical and tropical species corresponds to the route of the warm Kuroshio Current and its branches. The Kuroshio Current remains warm throughout the year (23–26 °C), and facilitates a great number of warm-water marine species migrate northward from their tropical center and expand their distribution ranges[78]. The strong correlation between species distributions and flows of oceanographic currents associate with their temperatures implies an adaptation of *O. oratoria* species complex to local thermal regimes. This finding is not unusual because recent genetic studies of *M. cephalus* species complex revealed that water temperature is one of the main factors that affect the distribution of three NW Pacific cryptic species[9]. Such close relationships between oceanic currents system and geographical distribution of genetic variation were also observed in the king weakfish, *Macrodon ancylodon*, along Atlantic coastal waters of South America[14]. Further researches on adaptive selection are necessary to better understand the role of temperature gradients in the NW Pacific in generating species boundaries and in shaping the contemporary spatial distribution of *O. oratoria* species complex, and their adaption to local environment.

Conclusions

The physical and environmental attributes of the Northwestern (NW) Pacific provide an ideal model to explore mechanisms underpinning speciation and evolutionary processes of marine species. Our results revealed cryptic speciation in the Japanese mantis shrimp *O. oratoria* by integrating mitochondrial and nuclear evidence. These two cryptic species likely diverged in allopatry during the middle Pliocene epoch, followed by range extensions and secondary contact that resulted in their present-day distribution. Our analyses also revealed recent hybridization event of the two cryptic species occurred in sympatric areas. The distribution range of the two *O. oratoria* cryptic species exhibits a remarkable latitudinal cline, which might be correlated to distinct temperature preferences of the species. This study highlights the interactive role of paleoclimate changes and environmental heterogeneity in driving genetic diversification of coastal species in the NW Pacific and underscores the importance of phylogeographic data in revealing cryptic marine biodiversity. From a conservation perspective, management efforts will need to maintain the genetic integrity and evolutionary potential of each cryptic species in their native regions.

Methods

The methods were carried out in accordance with the approved guidelines of the Good Experimental Practices adopted by Institute of Oceanology, Chinese Academy of Sciences. All experimental protocols were conducted under the permits approved by Institute of Oceanology, Chinese Academy of Sciences, China.

Sampling, sequencing and alignment. A total of 498 *O. oratoria* individuals were collected from 22 localities along the coast of China and Japan (Fig. 1, Table 1), which almost covers the entire distribution range of this species. Four individuals of *Oratosquilla kempi* were also collected from South China Sea and used as out-group for phylogenetic analyses. Specimens were preserved in 95% ethanol for subsequent analysis.

Total genomic DNA was extracted from abdominal muscle using a DNeasy tissue kit (Qiagen) following the manufacturer's protocol. The 5' end of the mtDNA COI gene (658 bp) was amplified for all samples following cheng *et al.*[79] with the universal primers LCO1490 and HCO2198[80]. Considering the extensive sampling and patterns revealed by mtDNA COI in coastal waters of China, the nrDNA ITS spacer (ITS1-5.8S-ITS2) was amplified for a subset of samples (14 localities, 147 specimens) following Larsen[81] with the primers ITS1 and ITS4[82]. PCR products were purified using the QIA-quick gel purification kit (QIAGEN), in accordance with manufacturer's instructions. The purified PCR products for mtDNA COI were used as template for direct sequencing on an ABI Prism 3730 (Applied Biosystems) automatic sequencer. Sequencing reactions were performed on both forward and reverse strands. Direct sequencing of nrDNA ITS region yielded unreadable sequence data due to intra-genomic variation. In this case, the purified PCR products were cloned with the pMD19-T Vector Cloning kit (Takara Biotechnology). Positive colonies were subjected to direct PCR with the M13 primers using the same PCR conditions for nrDNA ITS. More than three clones were picked for each individual that showed incongruence between mitochondrial and nuclear markers in the primary analysis.

Sequences were assembled in MEGA v6.06[83], aligned with Muscle default settings and further refined manually. The nrDNA ITS had highly variable regions and the alignments needed introducing several gaps (insertions/ deletions). In this case, regions with adjacent mononucleotide or microsatellite repeats were excluded because uncertainty of homology could be exacerbated by potential inaccuracies of enzymatic processes during PCR amplification and sequencing[84, 85]. Indels introduced into the alignment were coded as binary characters following the simple insertion/deletion coding method[47] using FastGap v1.2[86].

We used multiple recombination detection methods implemented in RDP4[87] to identify putative recombination events between the nrDNA ITS sequences. Seven methods (RDP, Chimaera, BootScan, 3Seq, GENECONV,

MaxChi and SiScan) were applied to minimize the risk of false positives. The threshold P-value was set at 0.05 using Bonferroni correction for multiple comparisons.

Phylogenetic analyses. MtDNA COI-inferred haplotype phylogenies were constructed using maximum-likelihood (ML) analysis and Bayesian inference (BI) implemented in MEGA v6.06 and MrBayes v3.2.6[88], respectively. The haplotypes from *O. Kempi* were used as out-group. The best-fitting substitution model was determined using the program jModeltest[89] under Akaike's information criterion. The ML analysis was conducted under the chosen substitution model (TrN+I+G) with the Nearest-Neighbour Interchange tree-swapping operation. The starting tree was estimated by BioNJ. The robustness of the ML topology was tested with 1000 bootstrap replicates. For Bayesian analyses, two independent runs were carried out for 30 million generations with a sampling frequency of 1000. Convergence was assessed by monitoring average standard deviations of split frequencies between two simultaneous runs (<0.01) and potential scale reduction factor (PSRF, close to 1.0). The program Tracer v1.6[90] was applied to check all parameters for effective sampling size and unimodel posterior distribution. The first 25% of sampled trees were discarded as burn-in and the posterior probabilities were calculated from the remaining trees. In addition, genealogical relationships among haplotypes were assessed using a minimum spanning tree constructed by ARLEQUIN V3.5[91].

To test the concordance across nuclear and mitochondrial markers, a subset of specimens (n = 147) was sequenced for nrDNA ITS and phylogenetic analyses were performed under the same settings described above. The phylogenetic trees were unrooted because no suitable out-group was available owing to numerous ambiguous segments during alignments with other available stomatopod nrDNA ITS sequences. For Bayesian analyses, the best fitted model and its associated parameters were estimated separately for DNA and gap data partitions. The GTR+I+G model was chosen for DNA data, and the restriction site (binary) model with variable coding bias was applied for the gap characters assuming that the rates varied over such positions according to a gamma distribution. For ML analysis, the gaps were treated as missing data.

MEGA v6.06 was explored to calculate pairwise sequence divergences (uncorrected p-distances) between and within lineages as indicated by the phylogenetic analyses. We applied the species criterion ("$4 \times$ rule") to delimit cryptic species[46], which states that monophyletic lineages represent independent evolutionary units when the mean sequence difference between lineages is more than four times greater than the average variation within the lineage.

Population structure and diversity. Considering small sampling size around Hainan Island, samples from four sites (XY, PQ, YGH and SYS) in South China Sea were grouped as a single population (HN) after confirming non-significant between-site differentiation. Population-level pairwise genetic divergence was evaluated by the fixation index F_{ST}[92]. The significance of F_{ST} was tested with 1000 permutations after sequential Bonferroni adjustments. Furthermore, we conducted a hierarchical analysis of molecular variance (AMOVA) to estimate population structure among putative regional grouping according to the structured haplotypes and F_{ST} values between regions. For all calculations, significance was assessed by 1000 permutations and P values from multiple comparisons were Bonferroni-adjusted. All genetic structure calculations were performed in ARLEQUIN v3.5. Molecular diversity was estimated by the number of haplotypes (n), haplotype diversity (h) and nucleotide diversity (π) for each population and each lineage with ARLEQUIN, and confidence intervals were calculated with 1000 permutations.

Divergence time estimation. Bayesian molecular dating method was employed to estimate the divergence time between *O. oratoria* mtDNA lineages implemented in BEAST v.1.7.5[93]. Prior to the analysis, we tested the null hypothesis of equal molecular clock rate between lineages using Tajima's relative rate test[94] implemented in MEGA v6.06. Given that the result failed to reject the clock model ($P > 0.5$), we used a strict clock model and a coalescent tree model with constant population size in the BEAST analysis. As no mutation rates have been reported for stomatopods, this uncertainty was taken into our analyses such that a range of mutation rates (1.4–2.33% per million years estimated for crustacean species[95, 96]) was used as an a prior uniform distribution of the mutation rate of COI in BEAST to get a rough estimate of divergence time. We performed three independent runs in BEAST with a Markov chain Monte Carlo (MCMC) chain length of 10 million generations and sampling every 1000 generations. Three runs were then combined with logCombiner v1.7.5[93] and the first 1 million generations of each run were discarded as burn-in. The program Tracer v1.6[90] was used to ascertain that analyses converged to similar posterior probabilities and that the effective sampling size for each parameter exceeded 200. Node ages and lower bounds of the 95% highest posterior density intervals for divergence time were calculated using TreeAnnotator v1.7.5[93] and visualized in FigTree v1.4.1[93]. While the assumption underlying the molecular clock approach in estimating divergence time was called into question[97], we would regard time estimates as approximations on the scale of geological eras.

Demographic inference. Several alternative methods were used to examine the historical demography of *O. oratoria*. First, Tajima's D test[98] and Fu's F_S test[99] were calculated to test for neutrality. Significant negative D and F_S statistics can be interpreted as signatures of population expansion. Second, historical demography was further investigated by examining the frequency distributions of pairwise differences between sequences (mismatch distribution). The distribution is usually unimodel for lineages following a recent bottleneck or population expansion and multimodel in samples drawn from populations at demographic equilibrium. In addition, we tested the goodness-of-fit of the actual distributions with the expected distributions under a sudden expansion model by calculating the sum of squared deviations (SSD) and Harpending's raggedness index (RI) following 1000 coalescent simulations. Both neutrality tests and mismatch analysis were performed in ARLEQUIN. Third, we conducted Bayesian skyline plots (BSPs) with BEAST v.1.7.5 to infer past demographic changes using 50

million MCMC steps, sampled every 1000 generations under the assumption of a strict clock and the nucleotide substitution model inferred with jModeltest. Tracer v1.6 was used to check the convergence to the stationary distribution and sufficient effective sampling sizes for each estimated parameter. After discarding the first 10% of the trees as the burn-in, the same software was then used to perform the Bayesian skyline reconstruction. The mitochondrial COI was utilized as the reference marker in the BSP analysis considering the complicated evolutionary processes of nuclear ITS[24]. We used the same substitution rate as in the previous analyses to convert the parameters to actual time quantities. These analyses were performed separately for each genetic lineage identified in this study.

References

1. Knowlton, N. Sibling species in the sea. *Annu. Rev. Ecol. Syst.* **24**, 189–216 (1993).
2. Byrne, M., Hart, M. W., Cerra, A. & Cisternas, P. Reproduction and larval morphology of broadcasting and viviparous species in the Cryptasterina species complex. *Biol. Bull.* **205**, 285–294, doi:10.2307/1543292 (2003).
3. Goetze, E. Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae. *Proc. R. Soc. Lond. B* **270**, 2321–2331, doi:10.1098/rspb.2003.2505 (2003).
4. Andrews, K. R., Norton, E. L., Fernandez-Silva, I., Portner, E. & Goetze, E. Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. *Mol. Ecol.* **23**, 5462–5479, doi:10.1111/mec.12950 (2014).
5. Chan, B. K. K., Tsang, L. M. & Chu, K. H. Cryptic diversity of the *Tetraclita squamosa* complex (Crustacea: Cirripedia) in Asia: description of a new species from Singapore. *Zool. Stud.* **46**, 46–56 (2007).
6. Wu, T. H., Tsang, L. M., Chan, B. K. K. & Chu, K. H. Cryptic diversity and phylogeography of the island associated barnacle *Chthamalus moro* in Asia. *Mar. Ecol.* **36**, 368–378, doi:10.1111/maec.12146 (2015).
7. Taylor, J. D. & Glover, E. A. Cryptic diversity of chemosymbiotic bivalves: a systematic revision of worldwide *Anodontia* (Mollusca: Bivalvia: Lucinidae). *Syst. Biodivers.* **3**, 281–338, doi:10.1017/S1477200005001672 (2005).
8. Liu, J., Li, Q., Kong, F. & Zheng, X. Cryptic diversity in the pen shell *Atrina pectinata* (Bivalvia: Pinnidae): high divergence and hybridization revealed by molecular and morphological data. *Mol. Ecol.* **20**, 4332–4345, doi:10.1111/j.1365-294X.2011.05275.x (2011).
9. Shen, K. N., Jamandre, B. W., Hsu, C. C., Tzeng, W. N. & Durand, J. D. Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugil cephalus*. *BMC Evol. Biol.* **11**, 83, doi:10.1186/1471-2148-11-83 (2011).
10. Hubert, N. *et al.* Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. *PLoS ONE* **7**, e28987, doi:10.1371/journal.pone.0028987 (2012).
11. Palumbi, S. R. Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.* **25**, 547–572, doi:10.1146/annurev.es.25.110194.002555 (1994).
12. Knowlton, N. Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* **420**, 73–90, doi:10.1023/A:1003933603879 (2000).
13. Fé´ral, J. P. How useful are the genetic markers in attempts to understand and manage marine biodiversity? *J. Exp. Mar. Biol. Ecol.* **268**, 121–145, doi:10.1016/S0022-0981(01)00382-3 (2002).
14. Santos, S., Hrbek, T., Farias, I. P., Schneider, H. & Sampaio, I. Population genetic structuring of the king weakfish, *Macrodon ancylodon* (Sciaenidae), in Atlantic coastal waters of South America: deep genetic divergence without morphological change. *Mol. Ecol.* **15**, 4361–4373, doi:10.1111/j.1365-294X.2006.03108.x (2006).
15. Wilson, A. B. & Veraguth, I. E. The impact of Pleistocene glaciation across the range of a widespread European coastal species. *Mol. Ecol.* **19**, 4535–4553, doi:10.1111/j.1365-294X.2010.04811.x (2010).
16. Ho, P. T., Kwan, Y. S., Kim, B. & Won, Y. J. Postglacial range shift and demographic expansion of the marine intertidal snail *Batillaria attramentaria*. *Ecol. Evol.* **5**, 419–435, doi:10.1002/ece3.1374 (2015).
17. Tamaki, K. & Honza, E. Global tectonics and formation of marginal basins: role of the western Pacific. *Episodes* **14**, 224–230 (1991).
18. Wang, P. X. Response of Western Pacific marginal seas to glacial cycles: paleoceanographic and sedimentological features. *Mar. Geol.* **156**, 5–39, doi:10.1016/S0025-3227(98)00172-8 (1999).
19. Kitamura, A., Takano, O., Takata, H. & Omote, H. Late Pliocene-early Pleistocene paleoceanographic evolution of the Sea of Japan. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **172**, 81–98, doi:10.1016/S0031-0182(01)00272-3 (2001).
20. Kimura, M. Paleogeography of the Ryukyu Islands. *Tropics* **10**, 5–24, doi:10.3759/tropics.10.5 (2000).
21. Voris, H. K. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J. Biogeogr.* **27**, 1153–1167, doi:10.1046/j.1365-2699.2000.00489.x (2000).
22. Xu, J., Chan, T. Y., Tsang, L. M. & Chu, K. H. Phylogeography of the mitten crab *Eriocheir sensu stricto* in East Asia: Pleistocene isolation, population expansion and secondary contact. *Mol. Phylogenet. Evol.* **52**, 45–56, doi:10.1016/j.ympev.2009.02.007 (2009).
23. Liu, J. X., Gao, T. X., Wu, S. F. & Zhang, Y. P. Pleistocene isolation in the Northwestern Pacific marginal seas and limited dispersal in a marine fish, *Chelon haematocheilus* (Temminck & Schlegel, 1845). *Mol. Ecol.* **16**, 275–288, doi:10.1111/j.1365-294X.2006.03140.x (2007).
24. Ni, G., Li, Q., Kong, L. & Zheng, X. Phylogeography of bivalve *Cyclina sinensis*: testing the historical glaciations and Changjiang River outflow hypotheses in northwestern Pacific. *PLoS ONE* **7**, e49487, doi:10.1371/journal.pone.0049487 (2012).
25. Lin, W. *et al.* Phylogeography of the finless porpoise (genus *Neophocaena*): testing the stepwise divergence hypothesis in the northwestern Pacific. *Sci. Rep.* **4**, 6572, doi:10.1038/srep06572 (2014).
26. Zhan, A. *et al.* Fine-scale population genetic structure of Zhikong scallop (*Chlamys farreri*): do local marine currents drive geographical differentiation? *Mar. Biotechnol.* **11**, 223–235, doi:10.1007/s10126-008-9138-1 (2009).
27. Hu, Z. M. *et al.* Phylogeographic heterogeneity of the brown macroalga *Sargassum horneri* (Fucaceae) in the northwestern Pacific in relation to late Pleistocene glaciation and tectonic configurations. *Mol. Ecol.* **20**, 3894–3909, doi:10.1111/j.1365-294X.2011.05220.x (2011).
28. Hu, Z. M., Zhang, J., Lopez-Bautista, J. & Duan, D. L. Asymmetric genetic exchange in the brown seaweed *Sargassum fusiforme* (Phaeophyceae) driven by oceanic currents. *Mar. Biol.* **160**, 1407–1414, doi:10.1007/s00227-013-2192-x (2013).
29. Han, Z. *et al.* A barrier to gene flow in the Asian paddle crab, *Charybdis japonica*, in the Yellow Sea. *ICES J. Mar. Sci.* **72**, 1440–1448, doi:10.1093/icesjms/fsv020 (2015).
30. Li, J. J. *et al.* Oceanic currents drove population genetic connectivity of the brown alga *Sargassum thunbergii* in the north-west Pacific. *J. Biogeogr.* **44**, 230–242, doi:10.1111/jbi.12856 (2017).
31. Hsueh, Y., Wang, J. & Chern, C. S. The intrusion of the Kuroshio across the continental shelf northeast of Taiwan. *J. Geophys. Res.* **97**, 14323–14330, doi:10.1029/92JC01401 (1992).
32. Hu, J., Kawamura, H., Hong, H. & Qi, Y. A review on the currents in the South China Sea: seasonal circulation, South China Sea Warm Current and Kuroshio intrusion. *J. Oceanogr.* **56**, 607–624, doi:10.1023/A:1011117531252 (2000).
33. Wu, B. Some problems on circulation study in Taiwan Strait. *Taiwan Str.* **1**, 1–7 (1982).
34. Wang, B., Wang, X. & Zhan, R. Nutrient conditions in the Yellow Sea and the East China Sea. *Estuar. Coast. Shelf S.* **58**, 127–136, doi:10.1016/S0272-7714(03)00067-2 (2003).
35. Chen, C. C., Shiah, F. K., Chaing, K., Gong, G. C. & Kemp, W. M. Effects of the Changjiang (Yangtze) River discharge on planktonic community respiration in the East China Sea. *J. Geophys. Res.* **114**, C03005, doi:10.1029/2008JC004891 (2009).

36. Han, Z. Q., Zhu, W. B., Zheng, W., Li, P. & Shui, B. N. Significant genetic differentiation between the Yellow Sea and East China Sea populations of cocktail shrimp *Trachypenaeus curvirostris* revealed by the mitochondrial DNA COI gene. *Biochem. Syst. Ecol.* **59**, 78–84, doi:10.1016/j.bse.2014.12.028 (2015).

37. Komai, T. Stomatopoda of Japan and adjacent localities. *Mem. Coll. Sci. Kyoto Imp. Univ. Ser. B* **3**, 307–354 (1927).

38. Manning, R. B. Keys to the species of *Oratosquilla* (Crustacea, Stomatopoda), with descriptions of two new species. *Smithson. Contrib. Zool.* **71**, 1–16, doi:10.5479/si.00810282.71 (1971).

39. Huang, Z. G. *Marine Species and Their Distribution in China.* (Ocean Press, Beijing, 2008).

40. Wang, C. L. & Xu, S. L. A biological basic character of *Oratosquilla oratoria. J. Zhejiang Coll. Fish* **15**, 60–62 (1996).

41. Kodama, K., Yamakawa, T., Shimizu, T. & Aoki, I. Age estimation of the wild population of Japanese mantis shrimp *Oratosquilla oratoria* (Crustacea: Stomatopoda) in Tokyo Bay, Japan, using lipofuscin as an age marker. *Fish. Sci.* **71**, 141–150, doi:10.1111/j.1444-2906.2005.00941.x (2005).

42. Lin, Y. J., Liu, H. Y., Xu, H. L. & Gu, D. X. Morphometry of mantis shrimp *Oratosquilla oratoria* in Dalian Coast. *J. Dalian Fish. Univ.* **23**, 215–217 (2008).

43. Zhang, D. *et al.* Comparative phylogeography of two marine species of crustacean: recent divergence and expansion due to environmental changes. *Gene* **550**, 141–147, doi:10.1016/j.gene.2014.08.006 (2014).

44. Zhang, D. *et al.* Geographical distribution, dispersal and genetic divergence of the mantis shrimp *Oratosquilla oratoria* (Stomatopoda: Squillidae) in China Sea. *Biochem. Syst. Ecol.* **65**, 1–8, doi:10.1016/j.bse.2016.01.009 (2016).

45. Du, X. *et al.* Population genetic structure of mantis shrimps *Oratosquilla oratoria*: testing the barrier effect of the Yangtze River outflow. *Biochem. Syst. Ecol.* **66**, 12–18, doi:10.1016/j.bse.2016.02.033 (2016).

46. Birky, C., Wolf, C., Maughan, H., Herbertson, L. & Henry, E. Speciation and selection without sex. *Hydrobiologia* **546**, 29–45, doi:10.1007/s10750-005-4097-2 (2005).

47. Simmons, M. P. & Ochoterena, H. Gaps as characters in sequence-based phylogenetic analyses. *Syst. Biol.* **49**, 369–381, doi:10.1093/sysbio/49.2.369 (2000).

48. Posada, D., Crandall, K. A. & Holmes, E. Recombination in evolutionary genomics. *Annu. Rev. Genet.* **36**, 75–97, doi:10.1146/annurev.genet.36.040202.111115 (2002).

49. Tada, R. Paleoceanographic evolution of the Japan Sea. *Palaeogeogr. Palaeocl.* **108**, 487–508, doi:10.1016/0031-0182(94)90248-8 (1994).

50. Kitamura, A., Takano, O., Takata, H. & Omote, H. Late Pliocene–early Pleistocene paleoceanographic evolution of the Sea of Japan. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **172**, 81–98, doi:10.1016/S0031-0182(01)00272-3 (2001).

51. Ota, H. Geographic patterns of endemism and speciation in amphibians and reptiles of the Ryukyu Archipelago, Japan, with special reference to their paleogeographical implications. *Popul. Ecol.* **40**, 189–204, doi:10.1007/BF02763404 (1998).

52. Cheang, C. C., Chu, K. H. & Ang, J. P. O. Phylogeography of the marine macroalga *Sargassum hemiphyllum* (Phaeophyceae, Heterokontophyta) in northwestern Pacific. *Mol. Ecol.* **19**, 2933–2948, doi:10.1111/j.1365-294X.2010.04685.x (2010).

53. Hewitt, G. M. Some genetic consequences of ice ages, and their role, in divergence and speciation. *Biol. J. Linn. Soc.* **58**, 247–276, doi:10.1111/j.1095-8312.1996.tb01434.x (1996).

54. Xu, X. & Oda, M. Surface-water evolution of the eastern East China Sea during the last 36,000 years. *Mar. Geol.* **156**, 285–304, doi:10.1016/S0025-3227(98)00183-2 (1999).

55. Kitamura, A., Takano, O., Takata, H. & Omote, H. Late Pliocene–early Pleistocene paleoceanographic evolution of the Sea of Japan. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **172**, 81–98, doi:10.1016/S0031-0182(01)00272-3 (2001).

56. Ruddiman, W. F., Raymo, M. E., Martinson, D. G., Clement, B. M. & Backman, J. Pleistocene evolution: Northern hemisphere ice sheets and North Atlantic Ocean. *Paleoceanography* **4**, 353–412, doi:10.1029/PA004i004p00353 (1989).

57. Wu, G. J., Pan, B. T., Guan, Q. Y. & Gao, H. S. The maximum glaciation and desert expansion in China during MIS16. *J. Glaciol. Geocryol.* **24**, 544–549 (2002).

58. Ni, G., Li, Q., Kong, L. & Yu, H. Comparative phylogeography in marginal seas of the northwestern Pacific. *Mol. Ecol.* **23**, 534–548, doi:10.1111/mec.12620 (2014).

59. Wang, J., Tsang, L. M. & Dong, Y. W. Causations of phylogeographic barrier of some rocky shore species along the Chinese coastline. *BMC Evol. Biol.* **15**, 114, doi:10.1186/s12862-015-0387-0 (2015).

60. Morrone, F., Lo Brutto, S. & Arculeo, M. Molecular evidence for the presence of cryptic evolutionary lineages in the freshwater copepod genus *Hemidiaptomus* G.O. Sars, 1903 (Calanoida, Diaptomidae). *Hydrbiologia* **644**, 115–125, doi:10.1007/s10750-010-0101-6 (2010).

61. Baird, H. P., Miller, K. J. & Stark, J. S. Evidence of hidden biodiversity, ongoing speciation and diverse patterns of genetic structure in giant Antarctic amphipods. *Mol. Ecol.* **20**, 3439–3454, doi:10.1111/j.1365-294X.2011.05173.x. (2011).

62. Barber, P. H., Erdmann, M. V. & Palumbi, S. R. Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the Coral Triangle. *Evolution* **60**, 1825–1839, doi:10.1111/j.0014-3820.2006.tb00526.x (2006).

63. Barber, P. H. & Boyce, S. L. Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proc. R. Soc. Lond. B* **273**, 2053–2061, doi:10.1098/rspb.2006.3540 (2006).

64. Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **20**, 229–237, doi:10.1016/j.tree.2005.02.010 (2005).

65. Harrison, R. G. *Hybrid Zones and the Evolutionary Process.* (Oxford University Press, Oxford, 1993).

66. Coyne, J. & Orr, H. A. *Speciation.* (Sinauer Associates Inc., Sunderland, 2004).

67. O'Donnell, K. & Cigelnik, E. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus Fusarium are nonorthologous. *Mol. Phylogenet. Evol.* **7**, 103–116, doi:10.1006/mpev.1996.0376 (1997).

68. Sponaugle, S. *et al.* Predicting self-recruitment in marine populations: biophysical correlates and mechanisms. *Bull. Mar. Sci.* **70**, 341–375 (2002).

69. Palumbi, S. R. Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. *Annu. Rev. Env. Resour.* **29**, 31–68, doi:10.1146/annurev.energy.29.062403.102254 (2004).

70. Hamano, T. & Matsuura, S. Egg size, duration of incubation, and larval development of the Japanese mantis shrimp in the laboratory. *Nip. Suis. G.* **53**, 23–39, doi:10.2331/suisan.53.23 (1987).

71. Buonaccorsi, V. P., Kimbrell, C. A., Lynn, E. A. & Vetter, R. D. Population structure of copper rockfish (*Sebastes caurinus*) reflects postglacial colonization and contemporary patterns of larval dispersal. *Can. J. Fish. Aquat. Sci.* **59**, 1374–1384, doi:10.1139/f02-101 (2002).

72. Galarza, J. A. *et al.* The influence of oceanographic fronts and early-life history traits on connectivity among littoral fish species. *Proc. Natl. Acad. Sci. USA* **106**, 1473–1478, doi:10.1073/pnas.0806804106 (2009).

73. Wu, Z. Y., Chen, H. X. & Liu, N. Relationship between east China Sea Kuroshio and climatic elements in East China. *B. Mar. Sci.* **12**, 1–9 (2010).

74. Dong, Y. W. *et al.* The impact of Yangtze River discharge, ocean currents and historical events on the biogeographic pattern of *Cellana toreuma* along the China coast. *PLoS ONE* **7**, e36178, doi:10.1371/journal.pone.0036178 (2012).

75. Yang, D. *et al.* Numerical study on the pattern and origins of Kuroshio branches in the bottom water of southern East China Sea in summer. *J. Geophys. Res.* **117**, C02014, doi:10.1029/2011JC007528 (2012).

76. Yang, D., Yin, B., Sun, J. & Zhang, Y. Numerical study on the origins and the forcing mechanism of the phosphate in upwelling areas off the coast of Zhejiang province, China in summer. *J. Mar. Syst.* **123–124**, 1–18, doi:10.1016/j.jmarsys.2013.04.002 (2013).

77. Teague, W. J. *et al.* Connectivity of the Taiwan, Cheju, and Korea straits. *Cont. Shelf Res.* **23**, 63–77, doi:10.1016/S0278-4343(02)00150-4 (2003).

78. Liu, J. Y. Status of marine biodiversity of the China Seas. *PLoS ONE* **8**, e50719, doi:10.1371/journal.pone.0050719 (2013).

79. Cheng, J., Sha, Z. & Liu, R. DNA barcoding of genus *Metapenaeopsis* (Decapoda: Penaeidae) and molecular phylogeny inferred from mitochondrial and nuclear DNA sequences. *Biochem. Syst. Ecol.* **61**, 376–384, doi:10.1016/j.bse.2015.07.005 (2015).
80. Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I form diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299 (1994).
81. Larsen, K. Morphological and molecular investigation of polymorphism and cryptic species in tanaid crustaceans: implications for tanaid systematics and biodiversity estimates. *Zool. J. Linn. Soc.* **131**, 353–379, doi:10.1111/j.1096-3642.2001.tb02241.x (2001).
82. White, T. J., Bruns, T., Lee, S. & Taylor, J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. in *PCR Protocols: a Guide to Methods and Applications* (eds Innis, M., Gelfand, D., Sninsky, J. & White, T.) 315–322 (Academic Press, New York, 1990).
83. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729, doi:10.1093/molbev/mst197 (2013).
84. Kelchner, S. A. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Ann. Mo. Bot. Gard.* **87**, 482–498, doi:10.2307/2666142 (2000).
85. Zhu, Q. & Ge, S. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytol.* **167**, 249–265, doi:10.1111/j.1469-8137.2005.01406.x (2005).
86. Borchsenius, F. FastGap 1.2. *Department of Biological Sciences. University of Aarhus, Denmark* http://www.aubot.dk/FastGap_home.htm (2009).
87. Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **1**, vev003, doi:10.1093/ve/vev003 (2015).
88. Ronquist, F. *et al.* Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542, doi:10.1093/sysbio/sys029 (2012).
89. Posada, D. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256, doi:10.1093/molbev/msn083 (2008).
90. Rambaut, A., Suchard, M. & Drummond, A. J. Tracer v1.6. Available from http://tree.bio.ed.ac.uk/software/tracer/ Date of access: 11/12/2013 (2013).
91. Excoffier, L. & Lischer, H. E. L. 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, doi:10.1111/j.1755-0998.2010.02847.x (2010).
92. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370, doi:10.2307/2408641 (1984).
93. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973, doi:10.1093/molbev/mss075 (2012).
94. Tajima, F. Statistical analysis of DNA polymorphism. *Jpn. J. Genet.* **68**, 567–595, doi:10.1266/jjg.68.567 (1993).
95. Knowlton, N. & Weigt, L. A. New dates and new rates for divergence across the Isthmus of Panama. *Proc. R. Soc. Lond. B* **265**, 2257–2263, doi:10.1098/rspb.1998.0568 (1998).
96. Schubart, C. D., Diesel, R. & Hedges, S. B. Rapid evolution to terrestrial life in Jamaican crabs. *Nature* **393**, 363–365, doi:10.1038/30724 (1998).
97. Ayala, F. J. Vagaries of the molecular clock. *Proc. Nati. Acad. Sci. USA* **94**, 7776–7783, doi:10.1073/pnas.94.15.7776 (1997).
98. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
99. Fu, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925 (1997).

Acknowledgements

Author Contributions

J.C. performed the molecular experiment, conducted data analyses and drafted the manuscript. Z.S. revised the manuscript and managed the research project. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-02059-7

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.