Journal of Biogeography (J. Biogeogr.) (2011) 38, 1600–1612

ORIGINAL ARTICLE

Phylogeography of *Emerita analoga* (Crustacea, Decapoda, Hippidae), an eastern Pacific Ocean sand crab with long-lived pelagic larvae

Michael N Dawson[1,2], Paul H. Barber[3]*, Laura I. González-Guzmán[4], Robert J. Toonen[5], Jenifer E. Dugan[6] and Richard K. Grosberg[1]

[1]*Center for Population Biology, College of Biological Sciences, 1 Shields Avenue, University of California, Davis, CA 95616, USA,* [2]*School of Natural Sciences, 5200 North Lake Road, University of California, Merced, CA 95343, USA,* [3]*Department of Ecology and Evolutionary Biology, 621 Charles E. Young Drive South, University of California, Los Angeles, CA 90095, USA,* [4]*Section of Integrative Biology, University of Texas at Austin, 1 University Station A6700, Austin, TX 78712, USA,* [5]*The Hawai'i Institute of Marine Biology, University of Hawaii at Manoa, Coconut Island, PO Box 1346, Kane'ohe, HI 96744, USA,* [6]*Marine Science Institute, University of California, Santa Barbara, CA 93106, USA*

*Correspondence: Paul Barber, Department of Ecology and Evolutionary Biology, 621 Charles E. Young Drive South, University of California, Los Angeles, CA 90095, USA.
E-mail: paulbarber@ucla.edu

ABSTRACT

Aim Phylogeographic analyses have confirmed high dispersal in many marine taxa but have also revealed many cryptic lineages and species, raising the question of how population and regional genetic diversity arise and persist in dynamic oceanographic settings. Here we explore the geographic evolution of *Emerita analoga*, an inter-tidal sandy beach crab with an exceptionally long pelagic larval phase and wide latitudinal, amphitropical, distribution. We test the hypothesis that eastern Pacific *E. analoga* constitute a single panmictic population and examine the location(s), timing and cause(s) of phylogeographic differentiation.

Location Principally the eastern Pacific Ocean.

Methods We sequenced cytochrome *c* oxidase subunit I (COI) from 742 *E. analoga* specimens collected between 1997 and 2000 and downloaded homologous sequences of congeners from GenBank. We reconstructed a phylogeny for *Emerita* species using maximum likelihood and Bayesian methods and estimated times to most recent common ancestors (TMRCAs), using a COI divergence rate of 1% Myr^{-1} and timing of closure of the Central American Seaway. We constructed the COI haplotype network of *E. analoga* using statistical parsimony, calculated population genetic and spatial structure statistics in Arlequin, and estimated the demographic history of *E. analoga* using Bayesian skyline analysis.

Results Population subdivision and allele frequency differences were insignificant among north-eastern Pacific locations over 2000 km apart ($\Phi_{ST} = 0.00$, $P = 0.70$), yet two distinct phylogroups were recovered from the north-eastern and south-eastern Pacific ($\Phi_{CT} = 0.87$, $P < 0.001$). Amphitropical differentiation of these temperate clades occurred after TMRCA 1.9 ± 0.02 (mean ± SE) Ma and *E. analoga* has expanded into its present-day north-eastern Pacific range since *c.* 250 ka.

Main conclusions *Emerita analoga* is not panmictic but is very widely dispersed and approaching genetic homogeneity, i.e. 'eurymixis', in the north-eastern Pacific. North-eastern and south-eastern Pacific populations of *E. analoga* probably became isolated *c.* 1.5 Ma as the tropical eastern Pacific Ocean warmed and expanded, intensifying barriers to gene flow. The fragmentation of a widespread ancestral species previously connected by long-distance gene flow ('soft vicariance') coincident with changing oceanographic conditions may be a common theme in the evolution of *Emerita* species and in other highly dispersive taxa. Highly dispersive species may differentiate because of, not despite, the dynamic oceanographic setting.

Keywords

Dispersal, *Emerita*, gene flow, habitat, marine biogeography, Pacific Ocean, pelagic larval duration, phylogeography, sand crab, vicariance.

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/j.1365-2699.2011.02499.x

Journal of Biogeography

INTRODUCTION

Numerous molecular genetic analyses of population structure in marine taxa reveal a limited phylogeographic structure (e.g. Lessios & Robertson, 2006), consistent with planktonic dispersal driven by ocean currents (Hardy, 1962; Palumbi, 1994; Hellberg, 2009) over geographic scales considerably broader than in terrestrial systems (e.g. Kinlan & Gaines, 2003). However, the discovery of many cryptic lineages in what were once thought to be widely distributed species (Knowlton, 2000; Bickford *et al.*, 2007) raises the question of how population and regional genetic diversity arise and persist in dynamic oceanographic settings expected to thoroughly mix genes among distant localities (Palumbi, 1994; Hickerson & Meyer, 2008).

Developmental mode, especially the time that larvae spend in the plankton, undoubtedly plays a critical role in governing genetic and phylogeographic structure in the sea (e.g. Scheltema, 1988; Doherty *et al.*, 1995; Bohonak, 1999; Dawson, 2001; Lester *et al.*, 2007); however, oceanography and its effects on larval transport (e.g. Cowen *et al.*, 2000; Byers & Pringle, 2006; Galindo *et al.*, 2006), sea-level or land-level change (e.g. Lessios, 2008) and selection (e.g. Hilbish & Koehn, 1985; Gilg & Hilbish, 2003) may also influence the geographic distribution of genetic variation and spatial pattern of evolution (e.g. Hellberg, 1998; Dawson & Hamner, 2005; Faucci *et al.*, 2007; Hickerson & Meyer, 2008; Lessios, 2008; Rocha & Bowen, 2008). The emerging picture is that life history strongly influences gene flow, population structure and species ranges (e.g. Bohonak, 1999), but regional phylogeographic and biogeographic patterns emerge from the interactions, past and present, between larval life histories and shared environmental constraints (Lester *et al.*, 2007; Pelc *et al.*, 2009).

The vast majority of marine phylogeographic studies focus on rocky shore, estuarine and shallow reef macroinvertebrate epifauna and fishes (see, for example, meta-analyses and reviews by Dawson, 2001; Hellberg, 2009; and Kelly & Palumbi, 2009; see also Dugan *et al.*, 2010). Given these environmental biases, an important question is whether emerging phylogeographic generalities also apply to other groups living in more spatially and temporally dynamic habitats, such as infauna of exposed sandy beaches and holoplankton, or whether such under-represented biotas have a suite of quantitatively different characteristics?

The mole crab *Emerita analoga* (Stimpson, 1857) inhabits the inter-tidal zone of wave-exposed sandy beaches of the temperate eastern Pacific Ocean. It is the most abundant resident of these dynamic habitats, with up to 79,000 individuals per metre of shoreline (Dugan *et al.*, 2003). *Emerita analoga* also has extensive dispersal potential due to its 3- to 4-month-long planktonic larval phase (Johnson, 1939). *Emerita analoga* ranges widely, from Kodiak Island, Alaska (58° N), to Bahia Magdalena, Baja California (26° N; Haig & Abbott, 1980), with disjunct populations in the Gulf of California and in South America, where it has been reported on sandy beaches from Salavary, Peru, to False Bay, Argentina (Efford, 1970, 1976). Across this range, life-history parameters such as size and age at maturity, growth rate, maximum individual size and survival, all vary significantly with latitude (Osorio *et al.*, 1967; Dugan *et al.*, 1994). Recruitment, which can occur from April to November throughout California, also varies both locally and regionally (Diehl *et al.*, 2007).

The high abundance, long pelagic larval duration and broad range that encompasses environmental gradients across which many other taxa exhibit genetic discontinuities or range limits, suggest that *E. analoga* is an ecological generalist consisting of a single, large population (Dugan *et al.*, 2000). Alternatively, geographic variation in recruitment in California, coupled with the occurrence of disjunct populations in the Gulf of California and South America, imply that gene flow could be limited across at least some parts of the range, promoting differentiation and perhaps local adaptation in *E. analoga*. Molecular analyses to date have provided little clarification, with studies of different scale and scope variously evincing high gene flow (Beckwitt, 1985) or substantial regional differentiation (Tam *et al.*, 1996; Haye *et al.*, 2002). Here, we bridge prior analyses to describe phylogeographic structure in *E. analoga* of the eastern Pacific Ocean and to infer the location(s), timing and cause(s) of any phylogeographic differentiation.

MATERIALS AND METHODS

Sample collection, DNA extraction, amplification and sequencing

We collected *E. analoga* from North American beaches throughout California in October 1997, 1998 and 1999, from sites in Chile in 1998, and from Baja California to Oregon in 2000 (Table 1). We haphazardly sampled crabs from targeted sandy beaches, preserved them in ethanol, and when a large number of crabs were collected we arbitrarily chose 15–20 individuals from each beach for genetic analysis.

We extracted genomic DNA using a 10% Chelex solution (Bio-Rad, Hercules, CA, USA; Walsh *et al.*, 1991) or Gentra Systems' (Minneapolis, MN, USA) PureGene marine invertebrate (no. 00690) extraction protocol. We used polymerase chain reaction (PCR) to amplify cytochrome *c* oxidase subunit I (COI) using the generic primers LCO1490 and HCO2198 (Folmer *et al.*, 1994) or with *E. analoga*-specific internal primer pairs EA-LCO (5′-TAT ACT TCA TTT TCG GAG CC) with EA-HCO (5′-AGG TGT TGA TAT AGA ACT GG) or ECL1 (5′-GGA ACA TTA TAC TTC ATT TTC GGA) with ECH1 (5′-GGA TTC TTG ACT CAC AAT ATG G). Thermocycling began with a hotstart followed by 32–40 cycles of 30–45 s at 94 °C, 30–45 s at 48–55 °C, and 45–60 s at 72 °C, and concluded with a 3–10 min hold at 72 °C. We cleaned PCR products with Exo-SAP, labelled amplicons using BigDye terminator chemistry, and visualized sequences using recommended protocols (Applied Biosystems Inc., Foster City, CA, USA). We combined, edited and compiled forward and reverse

Table 1 A list of localities and years sampled in this study of *Emerita analoga* in the eastern Pacific Ocean, showing latitude (°N) and sample size (n), haplotype diversity (h), nucleotide diversity (π) and the composite estimator theta of population size and mutation rate (Θ_S) with estimates of standard error (SE) or standard deviation (SD). Numbers in column 1 correspond to localities in Fig. 3 with population sizes ≥ 8.

	Population	Year	n	Latitude	h	SE	π	SE	Θ_S	SD
	Oregon									
1	Tolovana Beach	2000	23	45.873	0.91	0.04	0.0055	0.003	6.50	2.45
2	Cape Lookout	2000	17	45.366	0.95	0.04	0.0041	0.003	4.14	1.78
3	Ocean Beach	2000	15	44.188	0.96	0.04	0.0044	0.003	3.69	1.66
	California									
4	Manchester State Beach	1999	16	38.970	0.93	0.05	0.0047	0.003	3.92	1.72
		1998	16		0.92	0.05	0.0051	0.003	5.42	2.26
		1997	16		0.88	0.08	0.0049	0.003	5.12	2.16
5	Anchor Bay	1998	8	38.803	0.89	0.11	0.0043	0.003	3.09	1.65
6	Doran Beach	2000	8	38.314	0.93	0.08	0.0050	0.003	3.09	1.65
	Pacifica	1999	1	37.655	–	–	–	–	–	–
		1998	3		1.00	0.27	0.0025	0.003	1.33	1.10
		1997	1		–	–	–	–	–	–
7	Gazos Creek	2000	8	37.166	1.00	0.06	0.0053	0.004	4.24	2.15
8	Rio Del Mar Beach	1999	16	36.969	0.90	0.05	0.0036	0.002	3.01	1.39
		1998	13		0.96	0.05	0.0040	0.003	3.22	1.53
		1997	14		0.97	0.04	0.0053	0.003	4.83	2.15
9	Monterey	2000	17	36.699	0.95	0.04	0.0045	0.003	4.73	1.99
		1999	15		0.97	0.03	0.0061	0.004	4.61	2.00
		1998	16		0.93	0.05	0.0041	0.003	3.92	1.72
		1997	8		0.75	0.14	0.0022	0.002	1.54	0.96
10	Pfeiffer Beach	1999	36	36.153	0.96	0.02	0.0060	0.004	8.68	2.90
		1997	16		0.88	0.06	0.0041	0.003	3.92	1.72
		1997	15		0.91	0.06	0.0066	0.004	7.07	2.90
11	Morro Strand State Beach	2000	17	35.433	0.92	0.06	0.0048	0.003	4.73	1.99
		1999	16		0.92	0.05	0.0032	0.002	2.71	1.28
		1998	13		0.96	0.05	0.0049	0.003	4.19	1.90
12	Morro Bay	2000	10	35.379	1.00	0.04	0.0057	0.004	4.60	2.18
13	Nipomo Dunes Reserve	2000	18	34.959	0.88	0.06	0.0037	0.002	3.49	1.53
		1999	17		0.88	0.07	0.0042	0.003	4.44	1.88
		1998	15		0.90	0.07	0.0057	0.004	5.84	2.45
		1997	17		0.85	0.07	0.0033	0.002	3.55	1.57
14	Jalama Beach	1999	16	34.510	0.94	0.05	0.0044	0.003	4.22	1.83
		1998	16		0.96	0.04	0.0042	0.003	3.92	1.72
		1997	3		1.00	0.27	0.0087	0.007	4.67	3.13
	Gaviota State Park	1999	3	34.471	1.00	0.27	0.0025	0.003	1.33	1.10
		1998	1		–	–	–	–	–	–
15	Goleta Beach	2000	20	34.416	0.93	0.04	0.0046	0.003	4.79	1.94
16	Leadbetter	2000	20	34.400	0.94	0.04	0.0043	0.003	4.51	1.85
17	Rincon Beach	2000	20	34.376	0.96	0.03	0.0055	0.003	6.20	2.42
18	Zuma Beach	1999	18	34.012	0.93	0.04	0.0052	0.003	5.23	2.14
		1998	18		0.92	0.05	0.0044	0.003	4.94	2.04
		1997	16		0.94	0.05	0.0040	0.003	4.22	1.83
19	Redondo Beach	2000	15	33.814	0.94	0.04	0.0045	0.003	3.69	1.66
	Baja California									
20	Playa Encantada	2000	17	32.315	0.97	0.03	0.0047	0.003	4.44	1.88
21	Punta Clara	2000	19	31.534	0.96	0.03	0.0047	0.003	4.58	1.89
22	Punta Baja	2000	20	29.949	0.99	0.02	0.0061	0.004	7.33	2.81
23	Playa Lobera	2000	20	29.432	0.94	0.04	0.0052	0.003	5.07	2.04
24	Bahia Tortugas	2000	19	27.681	0.96	0.03	0.0046	0.003	4.58	1.89
25	Bahia Asuncion	2000	20	27.149	0.87	0.06	0.0032	0.002	3.66	1.56
26	Punta Abreojos	2000	19	26.709	0.94	0.04	0.0045	0.003	4.58	1.89
	Chile, South America									
	Las Cruces	1998	2	−33.500	1.00	0.50	0.0019	0.003	1.00	1.00
	Matanzas	1998	19	−33.967	0.68	0.12	0.0028	0.002	3.15	1.39

sequences in SEQUENCHER 4.5 (GeneCodes Corp., Ann Arbor, MI, USA), which we then aligned according to amino acid translations using the invertebrate mitochondrial code in BEAST v.1.4.1 (Drummond & Rambaut, 2007). We estimated PCR + sequencing error rate (see Appendix S1 in Supporting Information) to revise the estimate of haplotype diversity. All nucleotide sequences were deposited in GenBank under accession numbers HQ340645–HQ341386.

Phylogenetic analyses

COI sequences for *Emerita* species [*Emerita austroafricana* Schmitt, 1937, *E. analoga*, *Emerita benedicti* Schmitt, 1935, *Emerita brasiliensis* Schmitt, 1935, *Emerita emeritus* (Linnaeus, 1767), *Emerita rathbunae* Schmitt, 1935, *Emerita talpoida* (Say, 1817)] and *Hippa pacifica* (Dana, 1852) were downloaded from GenBank and aligned by eye using invertebrate mitochondrial (mt) DNA amino acid translations as a guide. Sequences were unavailable for three species (*Emerita holthuisi* Sankolli, 1965, *Emerita karachiensis* Niazi & Haque, 1974, *Emerita portoricensis* Schmitt, 1937) that are geographically and phylogenetically distant from New World *Emerita* species or nested within a western Atlantic clade that was otherwise sampled completely (Tam *et al.*, 1996; Haye *et al.*, 2002). We analysed this 10-taxon 444-position matrix, the longest dataset possible excluding terminal positions missing data, using maximum likelihood in GARLI v.1.0 (Zwickl, 2006) run on the CIPRES Portal 2.0 (http://www.phylo.org/). GARLI analyses used the GTR+G model of sequence evolution assuming four rate categories and empirical base frequencies. Starting trees were chosen stepwise for each of 100 heuristic searches, and default settings were used for all other parameters. Bootstrap analyses employed the same parameter settings except that two heuristic searches were completed on each of 200 bootstrap replicates.

We then added to the 10-taxon dataset an additional *E. analoga* sequence from GenBank (L43099), a second *E. rathbunae* sequence transcribed from Tam *et al.* (1996), and four new sequences; two selected randomly from western North America and two from Chile. Terminal positions missing data were deleted, resulting in a 16-taxon 361-nucleotide data matrix which we analysed using Bayesian Markov chain Monte Carlo (MCMC) phylogenetic techniques in the BEAST v.1.5.3 package (Drummond & Rambaut, 2007). Preliminary analyses were run in BEAST to estimate burn-in times, run times and clades for which origination times (i.e. the time to most recent common ancestor, TMRCA) might meaningfully be estimated. Phylogenetic reconstruction in BEAST employed the HKY model with empirical base frequencies, gamma site heterogeneity with four rate categories, partitioning nucleotides according to the SRD06 model (Shapiro *et al.*, 2006) with all parameters unlinked and an uncorrelated exponential relaxed clock with substitution rate = 1.0 and all tip dates set to zero. Starting trees were randomly generated and the tree prior assumed the Yule speciation process. Taxon sets for estimating TMRCA were not constrained to be monophyletic. The MCMC was 10^7 steps long and parameters were sampled every 10^3 steps after a burn-in of 10^6 steps. We ran the same analysis three times using different random seeds to confirm convergence, and we report the combined results of these runs. We subsequently repeated these phylogenetic analyses using the same settings but including only the 13 sequences from *Emerita* species of North and South America (i.e. *E. analoga*, *E. benedicti*, *E. brasiliensis*, *E. rathbunae*, *E. talpoida*). We generated summary statistics and the maximum clade credibility tree with median node heights and > 50% posterior probabilities from 27,003 sampled trees.

Phylogeographic, population genetic and historical demographic analyses

To portray the relationships among, and the diversity of, haplotypes, we constructed a minimum spanning tree of all haplotypes using the MINSPNET algorithm in ARLEQUIN 3.0 (Excoffier *et al.*, 2005) and statistical parsimony network in TCS v.1.2.1 (Clement *et al.*, 2000). We then plotted major groupings revealed in the minimum spanning tree against their geographic and temporal origin for all locations and years sampled and compared their frequencies using contingency tests to complement non-tree based analyses.

We further characterized spatial patterns in genetic diversity that might arise from variation in demographic history across the north-eastern Pacific range of *E. analoga* by correlating latitude with haplotype diversity (h), nucleotide diversity (π) and theta [a composite estimator of population size and mutation rate ($\Theta_s = N\mu$)], calculated in ARLEQUIN 3.0 (Schneider *et al.*, 2000). We examined regional structure in *E. analoga* using analysis of molecular variance (AMOVA) in ARLEQUIN 3.0 to compare samples, using uncorrected pairwise (p)-distances, within each of the years 1997, 1998, 1999 and 2000 and subsequently defined regional groups suggested by significant pairwise F_{ST} (the fixation index, a measure of population genetic differentiation). We analysed temporal changes in genetic diversity by comparing h, π and Θ_s across years and by using AMOVA to compare samples from different years within locations. Finally, we simultaneously examined regional and temporal structure by including all samples across all years in a single AMOVA. Because large numbers of singleton and low-frequency haplotypes arising only from the sampling process may obscure weak signals of genetic structure, we repeated the spatio-temporal AMOVA including only haplotypes observed at a frequency ≥ 5. Test significance was determined by comparison of the observed data against 3000 randomized permutations. Significance levels for pairwise F_{ST} values were also considered after Bonferroni correction. Finally, because F_{ST}-like statistics may not perform well when diversity is high (Jost, 2008) we also: (1) used a chi-square test to determine whether the frequency of the most common haplotypes at each location differed from random, based on the frequencies of these haplotypes in the entire data set, and (2) calculated D_{EST} (a nearly unbiased estimator of the relative

genetic differentiation of samples) with 1000 bootstraps in SPADE (Chao & Shen, 2009).

We characterized the demographic history of north-eastern Pacific *E. analoga* using mismatch analyses in ARLEQUIN 3.0 and by skyline analyses (Minin *et al.*, 2008) in BEAST v.1.5.3, although use of a single locus can detect only general trends (Heled & Drummond, 2008). Skyline analyses employed the HKY model with empirical base frequencies, gamma site heterogeneity with four rate categories (of comparable complexity with the models suggested by MODELTEST 3.7; Posada & Buckley, 2004), a strict clock with substitution rate = 1.0 and all tip dates set to zero. Starting trees were randomly generated, and the tree prior assumed the coalescent GMRF Bayesian skyride speciation process with uniform smoothing, which performs well in diverse scenarios without strong priors (Minin *et al.*, 2008). The MCMC was 2.0×10^7 steps long and parameters were sampled every 2.0×10^3 steps after a burn-in of 2.0×10^6 steps. The same analysis was run three times using different random seeds to confirm convergence and we report the combined results of these runs. The skyline plot was drawn using TRACER v.1.5 (Rambaut & Drummond, 2007).

RESULTS

Sample collection and sequence diversity

We sequenced 537 nucleotides of mtDNA COI from a total of 742 specimens of *E. analoga* collected in 1997 (n = 106), 1998 (n = 140), 1999 (n = 154) and 2000 (n = 342). Of these, 297 haplotypes were unique (all uninterrupted open reading frames), representing a total of 156 variable sites. Accounting for PCR and sequencing error, 237 unique haplotypes is the minimum estimate of haplotype richness in our sample of *E. analoga*. Mean (± SD) pairwise sequence difference among all *E. analoga* was 3.42 ± 1.75. Within South America pairwise sequence difference was 1.48 ± 0.93, within North America 2.51 ± 1.35, and between South America and North America 18.86 ± 2.40.

Phylogenetic analyses

In all major respects, Bayesian MCMC and maximum likelihood analyses reconstructed the same relationships between species of mole crabs (Fig. 1). All analyses of nucleotide sequences recovered a monophyletic clade of *Emerita* species in the Americas, within which Atlantic and Pacific species formed reciprocally monophyletic clades. The only topological difference among any analyses was in the relationships of Atlantic–Caribbean mole crabs. Bayesian analyses of nucleotides recovered the group '*E. benedicti* + *E. brasiliensis*' as the sister taxon to *E. talpoida*, whereas maximum likelihood analysis of nucleotides placed this group as paraphyletic with respect to *E. talpoida*, but statistical support was weak for both relationships. The estimated mean time, measured in mutations per nucleotide, to most recent common ancestor (± SE, effective sample size, n_{ESS}), for *Emerita* species in the Americas was ^{Amer}TMRCA = 0.217 (± 0.001, n_{ESS} = 5674; Fig. 1). For *Emerita* species of Pacific American shorelines, ^{Ear}TMRCA = 0.166 (± 0.001, n_{ESS} = 5467) and for *E. analoga*, ^{Ea}TMRCA = 0.0372 (± 0.0002, n_{ESS} = 7443). All analyses recovered two geographically distinct reciprocally monophyletic clades of *E. analoga* in North America and South America.

Phylogeographic, population genetic and historical demographic analyses

Haplotype diversity was high (mean = 0.93, SD = 0.045) across all locations, ranging from 0.68 in Matanzas, Chile, to 1.00 in North American samples after excluding samples with $n < 5$ (Table 1). In contrast, nucleotide diversity was low, ranging from 0.0022 to 0.0066, with a mean (± SD) of 0.0046 ± 0.001 (excluding samples with $n < 5$). Θ_s ranged from 1.54 in the 1997 sample from Monterey to 8.68 for the 1999 Pfeiffer Beach sample. There was no correlation between latitude and any measure of genetic diversity ($r \leq 0.021$, $P \geq 0.88$).

The statistical parsimony network revealed two clusters within *E. analoga*, one in North America and the other in South America, separated by 15 mutational steps. Within the North America cluster, the most common haplotypes (H), numbered 1 and 17 in Fig. 2, were separated by one mutational step. Of these, haplotype H1 is numerically dominant, observed in 144 individuals (19.5% of the total sample), and related by one mutational step to numerous singleton haplotypes and five haplotypes of moderate frequency (found in 6 to 37 individuals, i.e. frequencies of 0.8% to 4.99%); a single additional step includes all closely related haplotypes sampled two or more times. TCS inferred H1 to be ancestral. Haplotype H17 was the second most common haplotype, observed in 119 individuals (16.0%), but was connected by one mutational step only to singletons; the five haplotypes of moderate frequency were two or three mutational steps distant from H17. Eighty-two per cent of the remaining 297 haplotypes were singletons (of which *c.* 24 are likely due to PCR sequencing error). The major topology of the network is therefore well represented by the MINSPNET minimum spanning tree shown in Fig. 2 which forms the core of any network in which only the most parsimonious evolutionary history of a minority of singleton haplotypes is ambiguous. We use this minimum spanning tree topology for further analyses of the geographic evolution of *E. analoga* in North America.

Strong spatial genetic structure was evident in AMOVA of the 1998 North and South American samples (Φ_{ST} = 0.88, $P < 10^{-3}$), with 88.3% of genetic variance partitioned between North versus South American samples, 12.7% of variance within populations and none among populations within groups. Without the South American samples, $\Phi_{ST} = 0$ and all variation was apportioned within populations throughout North America. The identical pattern occurred in the North American collections from 1997, 1999 or 2000 (Table 2). AMOVA analyses excluding low-frequency ($n < 5$) haplotypes yielded identical patterns (results not shown).

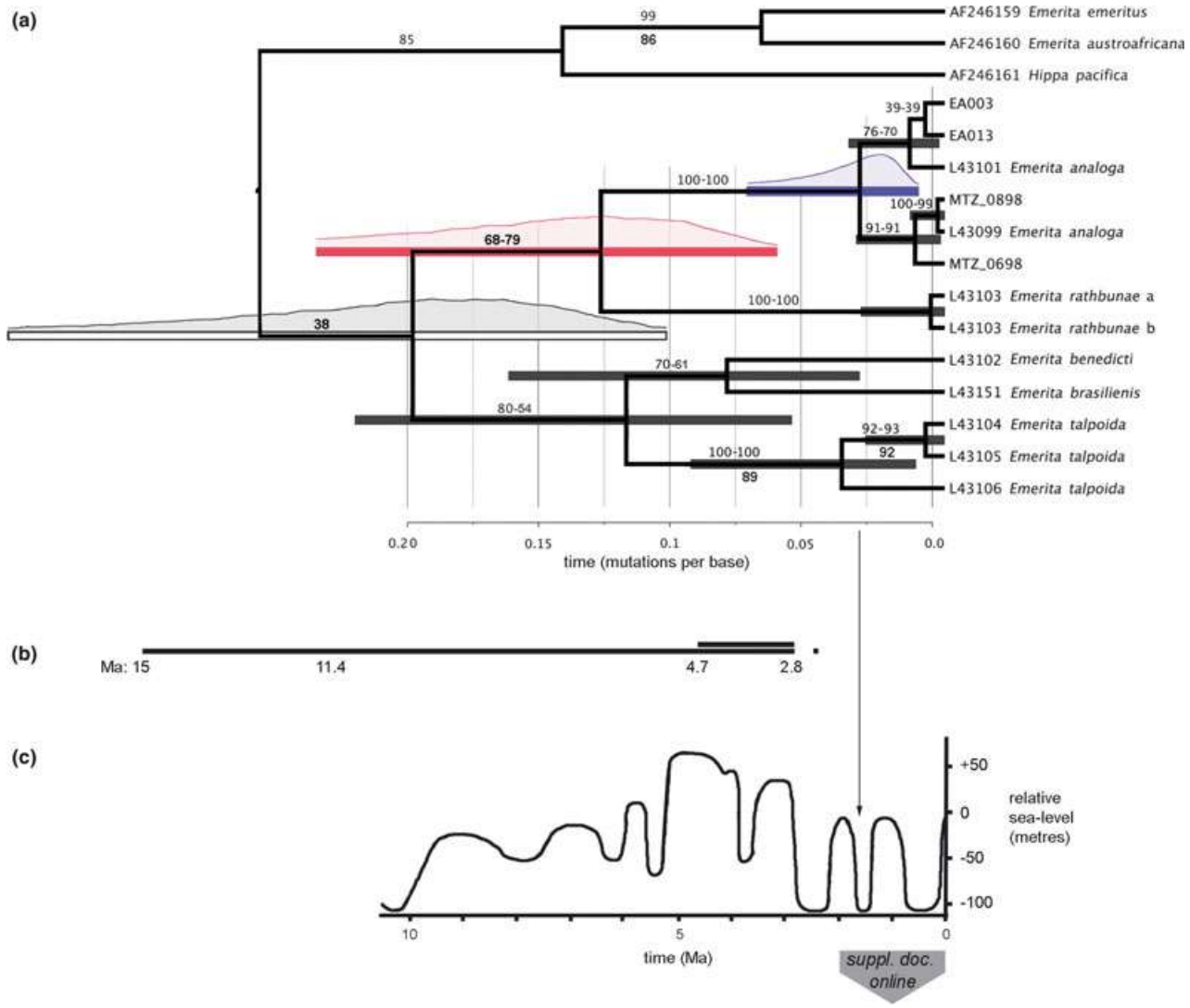

Figure 1 (a) Cytochrome *c* oxidase subunit I (COI) gene tree for species of mole crabs (*Emerita* spp. + *Hippa pacifica*) from the Atlantic, Indian and Pacific oceans reconstructed using Bayesian Markov chain Monte Carlo (MCMC) analysis (tree posterior probability = −1153, n_{ESS} = 53.4). Numbers above branches are posterior probabilities generated by analyses of, respectively, all species or only species from the Americas. Numbers below branches are maximum likelihood bootstrap support values (if > 50%). Nodes are plotted at the median height of 27,003 trees with relative marginal density distributions shown for the three nodes in the evolution of *E. analoga*. Time-scales are presented in units of mutations per base, counting from the present back in time. Haplotypes EA003 and EA013 are from, respectively, Punta Baja and Bahia Asuncion. (b) Scale bar showing the period of closure of the Central American Seaway due to emergence of the Central American Isthmus, following Coates *et al.* (2005), Harmon (2005) and Lessios (2008). Dates emphasized are 15 Ma, when closure began and extensive deepwater passages still existed; 11.4 Ma, which is the estimated earliest age of divergence in molluscs due to emergence of the Central American Isthmus; 4.7 Ma, when marine conditions began to diverge in the Caribbean and eastern Pacific; 2.8 Ma, at which the Central American Isthmus finally closed; and 2.0–1.9 Ma, when there may have been a brief breach in the isthmus. (c) Reconstructed short-term eustatic sea-level curve from 10 Ma to present, redrawn from Haq *et al.* (1987). Mutational and absolute time lines are drawn to scale assuming the rule-of-thumb evolutionary rate for COI of 1% divergence Myr^{-1}. Drop line from (a) to (c) indicates the median estimate for the most recent common ancestor of eastern Pacific *E. analoga*.

While $\Phi_{ST} = 0$ in all North American comparisons, some significant F_{ST} values were recovered. In 2000, significant ($P < 0.05$) pairwise F_{ST} values ($0.018 \leq F_{ST} \leq 0.12$) occurred between Cape Lookout, Oregon, and most populations north of Redondo Beach, California. There also was one significant F_{ST} value between Manchester and Rio del Mar, California, in 1997. These values, however, were not significant following Bonferroni correction. All other F_{ST} values were non-significant, whether comparing samples from different locations within years or from the same location in different years, indicating a lack of both geographic and temporal genetic structure within North American *E. analoga*. The null hypothesis that each of the two primary North American haplotype clusters is distributed uniformly across the range in

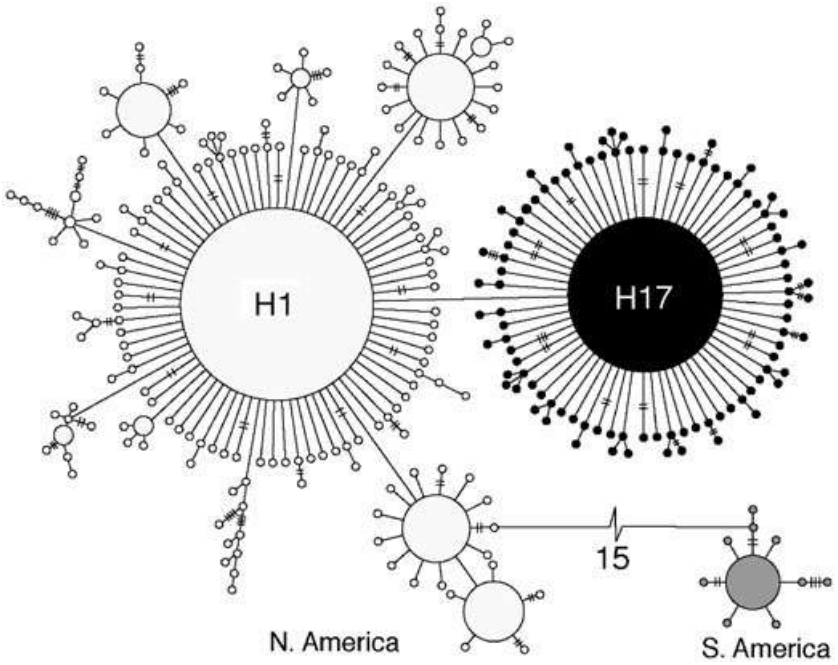

Figure 2 The minimum spanning tree depicting relationships and relative frequencies of the 297 mitochondrial cytochrome oxidase *c* subunit I (mtCOI) haplotypes (circles) sequenced from *Emerita analoga* in the eastern Pacific Ocean. The frequency of each haplotype is proportional to the area of the circle (smallest = frequency 1, largest = frequency 144). A single mutational difference between haplotypes is represented by a line without cross-marks, or by each cross-mark on a line, connecting circles. A total of 15 mutational steps separate populations from North America and South American (grey). Alternative connections among haplotypes are not shown (see text for details). Star-like clusters of North American haplotypes radiating from haplotypes H1 and H17 are coloured white and black, and their relative frequencies at each site are illustrated in Fig. 3.

the year 2000 was not rejected (H1: d.f. = 42, χ^2 = 11.30, $P > 0.05$; H17: d.f. = 42, χ^2 = 12.32, $P > 0.05$), although the relative frequencies of the clusters clearly vary in space and/or time (Fig. 3). Pairwise comparisons of shared allele frequencies also revealed differentiation of North American from South American *E. analoga* [D_{EST} = 0.150; lower 95% confidence interval (CI) = 0.082, upper 95% CI = 0.224] and no significant structure within North America (D_{EST} = 0.044; lower 95% CI = 0.000, upper 95% CI = 0.113).

The mismatch distribution for North American *E. analoga* differs statistically from that expected for a single rapid demographic expansion (P = 0.034) and rapid spatial expansion (P = 0.005) due to an excess of two-, three- and ≥ seven-step values for tau (τ) in the observed dataset relative to the model. The small mean value of τ and steep leading face of the distribution are consistent with a relatively recent bottleneck or founder event. Bayesian skyline analyses indicate that the most recent wave of population growth occurred within the last 250 kyr, which the median and lower estimates suggest may have occurred in two major stages, before and after *c.* 110–100 ka (see Appendix S1). Overall, effective population size increased by four orders of magnitude between *c.* 250–200 ka and 50 ka. The Bayesian skyline analyses also indicate a small decrease in effective population since 20–10 ka.

DISCUSSION

Phylogeographic structure of *Emerita analoga* in the eastern Pacific

In the absence of strong extrinsic barriers to dispersal, the exceptionally long pelagic larval phase of *E. analoga* should result in very high gene flow over large distances. We found insignificant mtCOI allele frequency differences and insignificant population subdivision among *E. analoga* samples collected at north-eastern Pacific locations over 2000 km apart, suggesting that across single or multiple generations dispersal and gene flow in *E. analoga* can occur over very long distances. Indeed, the occurrence of *E. analoga* in the north-eastern and south-eastern Pacific, separated geographically by the sister taxon *E. rathbunae*, must reflect trans-tropical long-distance dispersal during the evolutionary history of *E. analoga*.

However, *E. analoga* in the eastern Pacific Ocean is not a single panmictic population. *Emerita analoga* consists of two distinct phylogroups, one along the coast of Pacific North America and one on the coast of Pacific South America (Fig. 2, Table 2). Concordant zoogeographic and phylogeographic patterns (Tam *et al.*, 1996; our data) indicate that these groups diverged due to separation by a long-term biogeographic filter located in the tropical eastern Pacific Ocean. The two questions that remain to be answered, then, are when and how did this phylogeographic differentiation originate?

Timing of divergence within *Emerita analoga*

The speciation history of *Emerita* from the Americas provides an opportunity to estimate rates of molecular evolution by comparison with other taxa co-distributed on both sides of the Central American Isthmus (Lessios, 2008). Thus, the relative timing and environmental setting of the divergence of eastern Pacific *E. analoga* can be inferred with reasonable confidence, despite uncertainty introduced by lineage-specific or heterotachous (temporally varying) DNA substitution rates, using a single marker and coarse-grained geological reconstruction. Whether a rule-of-thumb substitution rate for COI of 1%

Table 2 Tests of regional and population genetic structure in *Emerita analoga* in the eastern Pacific Ocean. Analyses of molecular variance (AMOVA) were either run with North American (NA) and South American (SA) samples defined as two groups, or excluding South American samples. Analyses were run combining samples from 1997, 1998, 1999, 2000 (all years) or individually. South American samples were only available from 1998. AMOVA tables report the degrees of freedom (d.f.), sum of squares (Sum of sq.), Variance Components (Var. com.) and percentage of total molecular variation (% Var.) among North American and South American regions, as well as within and among North American populations.

	d.f.	Sum of sq.	Var. com.	% Var.
All years, NA versus SA				
Among NA and SA	1	349.21	8.53	87.3
Within NA or SA	49	58.67	0.00	0.0
Within populations	690	855.18	1.24	12.7
Φ_{ST}	0.87	$P < 0.00001$		
Φ_{CT}	0.87	$P < 0.00001$		
All years, no SA				
Among populations	48	58.09	0.00	−0.2
Within populations	671	841.00	1.25	100.2
Φ_{ST}	0.00	0.70		
2000				
Among populations	19	24.34	0.00	0.1
Within populations	322	403.24	1.25	99.9
Φ_{ST}	0.00	0.37		
1999				
Among populations	9	10.55	−0.01	−0.7
Within populations	144	186.61	1.30	100.7
Φ_{ST}	−0.01	0.70		
1998, NA versus SA				
Among groups	1	312.27	8.57	88.3
Within groups	11	10.77	−0.02	−0.2
Within populations	143	164.46	1.15	11.9
Φ_{ST}	0.88	$P < 0.00001$		
Φ_{CT}	0.88	$P < 0.00001$		
1998, no SA				
Among populations	10	10.19	−0.02	−1.3
Within populations	124	150.27	1.21	101.3
Φ_{ST}	−0.01	n.s.		
1997				
Among populations	7	10.18	0.02	1.6
Within populations	81	100.88	1.25	98.5
Φ_{ST}	0.02	n.s.		

Φ_{ST} is the correlation of random haplotypes within populations, and Φ_{CT} is the correlation of random haplotypes within a group of populations (i.e. region), relative to that of random pairs of haplotypes drawn from the whole species.

Myr^{-1} (Knowlton & Weigt, 1998; Marko, 2002; Fig. 1) or a rate two-fold different (e.g. Lessios, 2008) is applied, eastern Pacific (*E. analoga*, *E. rathbunae*) and western Atlantic–Caribbean (*E. benedicti*, *E. brasiliensis*, *E. talpoida*) mole crabs diverged many millions of years before closure of the Central American Seaway (Fig. 1). Pre-closure divergence of *Emerita* also agrees with Kimura two-parameter sequence differences of > 9% shown by a broad suite of trans-isthmian species pairs of other crustaceans (Lessios, 2008). Using these two metrics, speciation of *E. analoga* and *E. rathbunae* in the eastern Pacific coincided with shoaling of the Central American Seaway to *c.* 100–200 m between 8 and 6 Ma or oceanographic changes that began 4.7 Ma, but certainly preceded complete closure of the seaway 2.8 Ma (Fig. 1). Likewise, with a COI substitution rate of 1% Myr^{-1}, or twice that rate, speciation of *E. analoga* coincided with diastrophism within its modern range, but largely or entirely preceded tectonic formation of major geographical features at sea level such as the Gulf of California, which initiated *c.* 6 Ma (Ingersoll & Rumelhart, 1999) and the California Channel Islands *c.* 3 Ma (Davis *et al.*, 1989; Ward & Valensise, 1994). Based on the same rates of molecular evolution, north-eastern and south-eastern lineages of *E. analoga* diverged *c.* 1.8 ± 0.22 Ma to 0.9 ± 0.11 Ma.

Inferred causes of divergence within *Emerita analoga*

It is clear from the estimated timing of molecular divergences that cladogenetic events in *Emerita* were not caused by vicariance across a 'hard' physical barrier to gene flow such as the emergent Central American Isthmus. A hard vicariant barrier also cannot have caused divergence of *E. analoga* from the north-eastern and south-eastern Pacific, which are separated only by open ocean. Rather, as is the case for preceding cladogenetic events in *Emerita*, the divergence within *E. analoga* more likely resulted from changing oceanographic conditions.

Emerita analoga is a temperate descendant in a primarily tropical clade, now ranging from the upwelling regions of southern Baja California to Alaska (Haig & Abbott, 1980) and with the disjunct lineage occurring from northern Peru to southern Chile (Tam *et al.*, 1996). In the north-eastern Pacific, gene flow within *E. analoga* is largely unhindered along *c.* 2000 km of coastline, across which mean annual sea-surface temperature (SST) varies between *c.* 12 °C and 23 °C (see http://aquarius.nasa.gov/images/global_sst_map.jpg) and mean maximum SST approaches 25 °C (Dawson *et al.*, 2010); the range in mean annual SST is similar across the distribution of *E. analoga* in the south-eastern Pacific. This indicates that oceanographic processes acting as regional filters to genetic exchange in many marine species along the Pacific coast of North America either do not affect *E. analoga*, or are sufficiently permeable as to prevent the accumulation of regional genetic structure within this species after recent range expansion. In contrast, there has been no recent gene flow between north-eastern and south-eastern populations of *E. analoga* separated by *c.* 3000 km of ocean that is on average *c.* 23 °C to *c.* 29 °C (see also http://coralreefwatch.noaa.gov/satellite/images/current/key_sst_50kmcurrent.gif).

Significant gene flow between north-eastern and south-eastern Pacific populations of *E. analoga* last occurred *c.* 1.5 Ma (Fig. 1). This estimate coincides with an *c.* 0.25-Myr period of predominantly low sea level and cold climate that subsequently gave way to an *c.* 0.5-Myr period of predominantly high sea level and warm climate (Fig. 1),

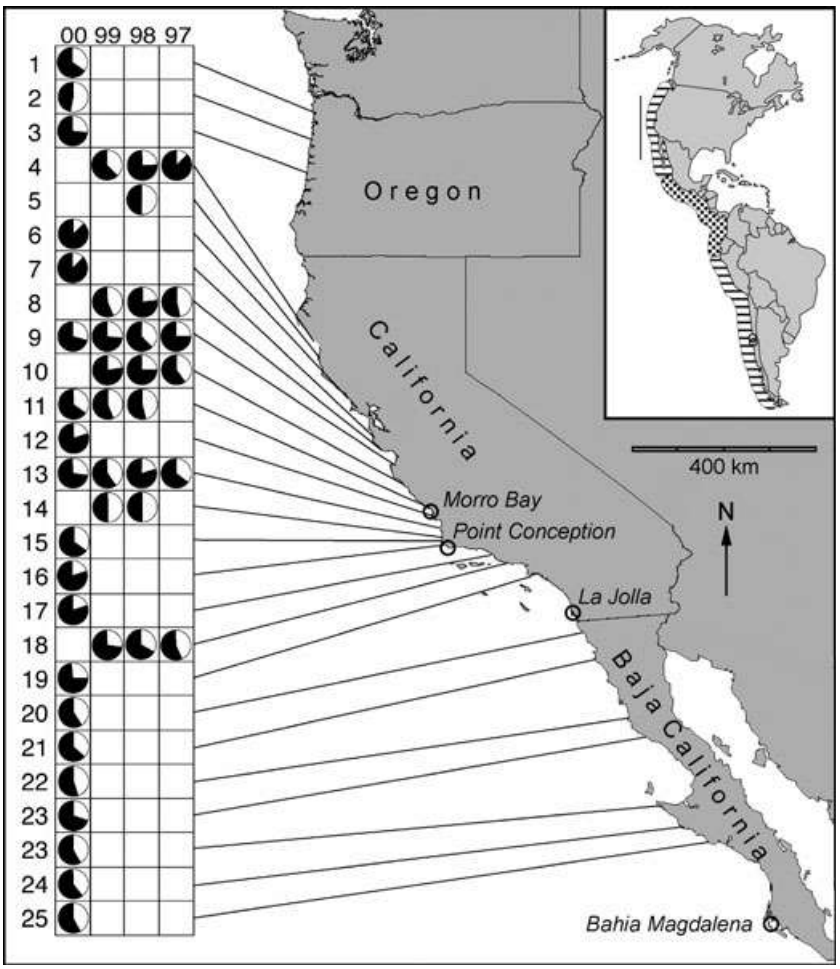

Figure 3 Geographic distribution of north-eastern Pacific *Emerita analoga* haplotypes composing the two major star-like haplotype clusters in Fig. 2. The first column corresponds to locality numbers in Table 1. Populations and/or sample years with fewer than eight individuals are excluded from this figure. Columns 2–5 indicate the frequencies of the two haplotype clusters in 2000, 1999, 1998 and 1997, respectively. Places mentioned in the text are shown. The California Transition Zone, between the cold temperate Oregonian and warm temperate Californian biogeographic provinces (Briggs, 1974), spans the region between Monterey Bay (locality 9) and Los Angeles (locality 19). Inset: the geographic distribution of *E. analoga* (lines) and *E. rathbunae* (dots); line indicates range of sampling in North America; open circles indicated the sampling locations in Chile.

during which the distribution of beach habitat must also have changed. Although sea level and climate oscillate at many frequencies, long-term median oceanographic conditions within these over-arching cold and warm periods approximated those during the Last Glacial Maximum and the pre-industrial period of the later Holocene (see Appendix S1). Using these analogues, we can infer that during the early Pleistocene cold period, the tropics probably presented a greatly reduced barrier to dispersal by *E. analoga*. The breadth of the ≥ 26 °C isotherm was *c.* 1000 km, with SST less than *c.* 28 °C across the entire tropical eastern Pacific, i.e. the proposed biogeographic filter was one-third the distance and 3 °C cooler than the biogeographic filters present during the preceding and following warm periods. Thus, successful dispersal of *E. analoga* between North America and South America was subsequently prevented during the predominantly warm period 1.5–1.0 Ma by increased geographic separation, altered current patterns, thermal tolerance of planktonic larvae, reduction of developmental time due to increased temperatures (O'Connor *et al.*, 2007) or some combination of these factors. During subsequent cold periods, assuming dispersal was again permitted by a narrower or less intense tropical barrier, introgression of North and South Pacific *E. analoga* may have been inhibited by drift, genetic incompatibilities or selection against allochthonous forms.

The Late Pleistocene demographic history of North American *E. analoga* provides further evidence that temperature change, or correlates of temperature change, causes evolution of the geographic range of *E. analoga*. The population of *E. analoga* in the north-eastern Pacific expanded in size beginning *c.* 250–200 ka, coinciding with an *c.* 100-kyr period of global cooling, glaciation and sea-level regression during Marine Isotope Stage 5 (Thompson & Goldstein, 2006). Peak population size was reached by the Last Glacial Maximum. Population expansion may therefore have occurred as habitat became available to the south of its prior range when cooling tropical regions led to contraction of the tropical sister taxon *E. rathbunae*. Reciprocally, effective population size may have decreased during the end of the Pleistocene and Holocene deglaciation (Appendix S1), presumably accompanying contraction of the southern range limit of *E. analoga* as ranges of temperate species shifted northwards (e.g. Fields *et al.*, 1993). If we instead applied a rate of molecular evolution of 2% Myr^{-1} (see Materials and Methods) to bracket plausible timing, we infer that the North American population of *E. analoga* expanded since the Sangamon Interglacial (*c.* 125 ka), during Marine Isotope Stage 6, stabilized at

approximately the Last Glacial Maximum (*c.* 20 ka) and declined slightly in size during the Holocene. Using either estimate of mutation rate, population growth was associated with SST cooling and population decline with SST warming.

Inferences from the evolution of *Emerita analoga*

Although variation in rates of molecular evolution, ambiguity in tectonic reconstructions and the many Pleistocene glacial–interglacial cycles notoriously complicate interpretation of phylogeographic signals, the events that best match mtCOI divergence of *Emerita* species and of *E. analoga* in particular yield a surprisingly consistent picture. Diversification of mole crabs has not been caused by hard vicariant events but rather by changes in the marine environment that led to fragmentation of a widespread ancestral species previously connected by long-distance gene flow, i.e. 'soft vicariance' *sensu* Hickerson & Meyer (2008). This interpretation from macroevolutionary patterns is consistent with intra-specific phylogeographic evidence that *Emerita* has high dispersal and colonization ability yet displays allele frequency differences that hint at a lack of complete panmixia. Indeed, our analyses, which are limited to the mitochondrial genome and modest sample sizes and imply a recent range expansion that would inflate estimated levels of modern gene flow due to shared ancestral polymorphisms, are likely to lack the power to distinguish very high gene flow, typically interpreted as panmixia, from extensive yet incomplete mixing, or 'eurymixis'. Eurymixis in north-eastern Pacific *E. analoga* is further supported by demographic patterns that indicate incomplete mixing of the larval pool, or chaotic genetic patchiness (Johnson & Black, 1982; Edmands *et al.*, 1996; Hellberg, 2009). For example, populations in Oregon are likely to be maintained by inconsistent larval immigration from California (Sorte *et al.*, 2001) and recruitment at sites in central California is spatially and temporally variable (Diehl *et al.*, 2007), arguing against panmixia. Eurymixis therefore may emerge over a wide variety of geographic and temporal scales, and although fluctuations in allele frequencies may often be viewed as short-lived 'aimless … evolution' or meaningful 'meandering' of micro-evolution (Thompson, 1999), eurymixis also provides potential for speciation when there is a heterogeneous environmental mosaic. For practical purposes, eurymixis might be defined as genetic heterogeneity despite estimated gene flow that is expected to homogenize all subpopulations [i.e. F_{ST}, D_{EST} etc. < 0.05; $Nm > 5$ (*Nm* being the average number of migrants per generation); Halliburton, 2004, p. 340].

The patterns and probable causes of diversification in *Emerita* are similar to those seen in other taxa. A majority of trans-isthmian sister taxa, including western Atlantic and eastern Pacific *Emerita* species, most probably diverged many millions of years before complete emergence of the Central American Isthmus (Lessios, 2008). Similarly, amphitropical divergence in *E. analoga* was contemporaneous with the divergence of many amphitropical eastern Pacific high-dispersal fishes (Burridge, 2002). Furthermore, like North American *E. analoga*, the grapsid crab *Pachygrapsus crassipes*, which inhabits predominantly rocky shores in the eastern North Pacific (Garth & Abbott, 1980) and also has a very long pelagic larval duration of at least 3–4 months (Cuesta & Rodríguez, 2000; DiBacco, 2000), shows no significant differentiation of populations from Baja California Norte to Vancouver Island ($\Phi_{ST} = 0.004$, $P = 0.158$); although, interestingly, *P. crassipes* shows high regional differentiation of populations from the eastern and western Pacific ($\Phi_{CT} = 0.843$, $P < 0.001$; Cassone & Boulding, 2006) analogous to the amphitropical divergence in *E. analoga*. Finally, the rocky intertidal volcano barnacle *Tetraclita rubescens*, which has a 3–4-week-long pelagic larval phase, similarly achieves very high gene flow ($Nm >> 1000$) across its entire north-eastern Pacific range on aggregate timescales, although samples near its recently expanded northern range limit share disproportionately many alleles with the northern portion of its ancestral range, demonstrating that the species is not truly panmictic (Dawson *et al.*, 2010). Thus, while high dispersal via ocean currents may link locations genetically, and even demographically, the spatially and temporally variable dynamics of oceanographic settings may ultimately cause and maintain population and regional genetic differentiation (Norris, 2000). Eurymixis and soft vicariance may therefore be relatively common mechanisms of diversification in marine taxa.

ACKNOWLEDGEMENTS

P.H.B. gratefully acknowledges the efforts of S. L. Boyce and M. E. Jones in collection of DNA sequence data, and Boston University and NSF OCE-0349177 for funding. Y. E. Stuart, B. Cameron, and the College of Biological Sciences DNA Sequencing Facility provided laboratory assistance at UC Davis. Jennifer Diehl collected the time series of samples at many sites. Funding from the National Science Foundation (OCE 0929057) and the Mellon Foundation to R.K.G. supported this research. Funding for sequencing performed in the EPSCoR Core Facility at HIMB came from NSF OCE-0623678 to R.J.T. L.I.G.-G. gratefully acknowledges the postdoctoral advice provided by S. Gaines and financial support through the Partnership for Interdiciplinary Studies of Coastal Oceans and a UC-Mexus grant during the field component of the present study. California field collections were supported in part by a grant to J.E.D. and A. Wenner from the Minerals Management Service, US Department of the Interior. J.E.D. was supported in part by the Santa Barbara Coastal LTER (NSF OCE no. 99-82105). C. Eckert and two anonymous referees provided comments that improved the manuscript.

REFERENCES

Beckwitt, R. (1985) Population genetics of the sand crab, *Emerita analoga* Stimpson, in southern California. *Journal of Experimental Marine Biology and Ecology*, **91,** 45–52.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22,** 148–156.

Bohonak, A.J. (1999) Dispersal, gene flow, and population structure. *Quarterly Review of Biology*, **74,** 21–45.

Briggs, J.C. (1974) *Marine zoogeography*. McGraw-Hill, New York.

Burridge, C.P. (2002) Antitropicality of Pacific fishes: molecular insights. *Biology of Fishes*, **65,** 151–164.

Byers, J.E. & Pringle, J.M. (2006) Going against the flow: retention, range limits and invasions in advective environments. *Marine Ecology Progress Series*, **313,** 27–41.

Cassone, B.J. & Boulding, E.G. (2006) Genetic structure and phylogeography of the lined shore crab, *Pachygrapsus crassipes*, along the northeastern and western Pacific coasts. *Marine Biology*, **149,** 213–226.

Chao, A. & Shen, T.-J. (2009) *Program SPADE (Species Prediction and Diversity Estimation)*. Available at: http://chao.stat.nthu.edu.tw.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9,** 1657–1660.

Coates, A.G., McNeill, D.F., Aubry, M.-P., Berggren, W.A. & Collins, L.S. (2005) An introduction to the geology of the Bocas del Toro Archipelago, Panama. *Caribbean Journal of Science*, **41,** 374–391.

Cowen, R.K., Lwiza, K.M.M., Sponaugle, S., Paris, C.B. & Olson, D.B. (2000) Connectivity of marine populations: open or closed? *Science*, **287,** 857–859.

Cuesta, J.A. & Rodríguez, A. (2000) Zoeal stages of the intertidal crab *Pachygrapsus marmoratus* (Fabricius, 1787) (Brachyura, Grapsidae) reared in the laboratory. *Hydrobiologia*, **436,** 119–130.

Davis, T.L., Namson, J. & Yerkes, R.F. (1989) A cross section of the Los Angeles area: seismically active fold and thrust belt, the 1987 Whittier Narrows earthquake, and earthquake hazard. *Journal of Geophysical Research*, **94,** 9644–9664.

Dawson, M.N (2001) Phylogeography in coastal marine animals: a solution from California? *Journal of Biogeography*, **28,** 723–736.

Dawson, M.N & Hamner, W.M. (2005) Rapid evolutionary radiation of marine zooplankton in peripheral environments. *Proceedings of the National Academy of Sciences USA*, **102,** 9235–9240.

Dawson, M.N, Grosberg, R.K., Sanford, E. & Stuart, Y.E. (2010) Population genetic analysis of a recent range expansion: mechanisms regulating the poleward range limit in the volcano barnacle *Tetraclita rubescens*. *Molecular Ecology*, **19,** 1585–1605.

DiBacco, C. (2000) The influence of bay versus coastal habitats on development and survival of striped shore crab larvae. *Bulletin Southern California Academy of Sciences*, **99,** 161–170.

Diehl, J.M., Toonen, R.J. & Botsford, L.W. (2007) Spatial variability of recruitment in the sand crab *Emerita analoga* throughout California in relation to wind-driven currents. *Marine Ecology Progress Series*, **350,** 1–17.

Doherty, P.J., Planes, S. & Mather, P. (1995) Gene flow and larval duration in seven species of fish from the Great Barrier Reef. *Ecology*, **76,** 2373–2391.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7,** 214.

Dugan, J.E., Hubbard, D.M. & Wenner, A.M. (1994) Geographic variation in life history of the sand crab, *Emerita analoga* (Stimpson) on the California coast: relationships to environmental variables. *Journal of Experimental Marine Biology and Ecology*, **181,** 255–278.

Dugan, J.E., Hubbard, D.M. & Lastra, M. (2000) Burrowing abilities and swash behavior of three crabs, *Emerita analoga* Stimpson, *Blepharipoda occidentalis* Randall and *Lepidopa californica* Efford (Anomura, Hippoidea), of exposed sandy beaches. *Journal of Experimental Marine Biology and Ecology*, **255,** 229–245.

Dugan, J.E., Hubbard, D.M., McCrary, M.D. & Pierson, M.O. (2003) The response of macrofauna communities and shorebirds to macrophyte wrack subsidies on exposed sandy beaches of southern California. *Estuarine Coastal and Shelf Science*, **58,** 25–40.

Dugan, J.E., Defeo, O., Jaramillo, E., Jones, A.R., Lastra, M., Nel, R., Peterson, C.H., Scapini, F., Schlacher, T. & Schoeman, D.S. (2010) Give beach ecosystems their day in the sun. *Science*, **329,** 1146.

Edmands, S., Moberg, P.E. & Burton, R.S. (1996) Allozyme and mitochondrial DNA evidence of population subdivision in the purple sea urchin *Strongylocentrotus purpuratus*. *Marine Biology*, **126,** 443–450.

Efford, I.E. (1970) Recruitment to sedentary marine populations as exemplified by the sand crab, *Emerita analoga* (Decapoda, Hippidae). *Crustaceana*, **18,** 293–308.

Efford, I.E. (1976) Distribution of the sand crabs in the genus *Emerita* (Decapoda, Hippidae). *Crustaceana*, **30,** 169–183.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1,** 47–50.

Faucci, A., Toonen, R.J. & Hadfield, M.G. (2007) Host shift and speciation in a coral-feeding nudibranch. *Proceedings of the Royal Society B: Biological Sciences*, **274,** 111–119.

Fields, P.A., Graham, J.B., Rosenblatt, R.H. & Somero, G.N. (1993) Effects of expected global climate change on marine faunas. *Trends in Ecology and Evolution*, **8,** 361–367.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3,** 294–299.

Galindo, H.M., Olson, D.B. & Palumbi, S.R. (2006) Seascape genetics: a coupled oceanographic-genetic model predicts population structure of Caribbean corals. *Current Biology*, **16,** 1622–1626.

Garth, J.S. & Abbott, D.P. (1980) Brachyura: the true crabs. *Intertidal invertebrates of California* (ed. by R.H. Morris, D.P. Abbott and E.C. Haderlie), pp. 594–630. Stanford University Press, Stanford, CA.

Gilg, M.R. & Hilbish, T.J. (2003) Patterns of larval dispersal and their effect on the maintenance of a blue mussel hybrid zone in southwestern England. *Evolution*, **57,** 1061–1077.

Haig, J. & Abbott, D.P. (1980) Macrura and Anomura: the ghost shrimps, hermit crabs, and allies. *Intertidal invertebrates of California* (ed. by R.H. Morris, D.P. Abbott and E.C. Haderlie), pp. 577–593. Stanford University Press, Stanford, CA.

Halliburton, R. (2004) *Introduction to population genetics*. Pearson Prentice Hall, Upper Saddle River, NJ.

Haq, B.U., Hardenbol, J. & Vail, P.R. (1987) Chronology of fluctuating sea levels since the Triassic. *Science*, **235,** 1156–1167.

Hardy, A. (1962) Studying the ever-changing sea. *Nature*, **196,** 207–210.

Harmon, R.S. (2005) Geological development of Panama. *The Rio Chagres, Panama* (ed. by R.S. Harmon), pp. 45–62. Springer, Amsterdam.

Haye, P.A., Tam, Y.K. & Kornfield, I. (2002) Molecular phylogenetics of mole crabs (Hippidae: *Emerita*). *Journal of Crustacean Biology*, **22,** 903–915.

Heled, J. & Drummond, A.J. (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8,** 289.

Hellberg, M.E. (1998) Sympatric sea shells along the sea's shore: the geography of speciation in the marine gastropod *Tegula*. *Evolution*, **52,** 1311–1324.

Hellberg, M.E. (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics*, **40,** 291–310.

Hickerson, M.J. & Meyer, C.P. (2008) Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical Bayesian approach. *BMC Evolutionary Biology*, **8,** 322.

Hilbish, T.J. & Koehn, R.K. (1985) The physiological basis of natural selection at the *Lap* locus. *Evolution*, **39,** 1302–1317.

Ingersoll, R.V. & Rumelhart, P.E. (1999) Three-stage evolution of the Los Angeles basin, southern California. *Geology*, **27,** 593–596.

Johnson, M.S. & Black, R. (1982) Chaotic genetic patchiness in an intertidal limpet. *Marine Biology*, **70,** 157–164.

Johnson, M.W. (1939) The correlation of water movements and dispersal of pelagic larval stages of certain littoral animals, especially the sand crab, *Emerita*. *Journal of Marine Research*, **11,** 236–245.

Jost, L. (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17,** 4015–4026.

Kelly, R.P. & Palumbi, S.R. (2009) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE*, **5,** e8594, doi:10.1371/journal.pone.0008594.

Kinlan, B.P. & Gaines, S.D. (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology*, **84,** 2007–2020.

Knowlton, N. (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia*, **420,** 73–90.

Knowlton, N. & Weigt, L.A. (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences*, **265,** 2257–2263.

Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the rise of the central American isthmus. *Annual Review of Ecology, Evolution, and Systematics*, **39,** 63–91.

Lessios, H.A. & Robertson, D.R. (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences*, **273,** 2201–2208.

Lester, S.E., Ruttenberg, B.I., Gaines, S.D. & Kinlan, B.P. (2007) The relationship between dispersal ability and geographic range size. *Ecology Letters*, **10,** 745–758.

Marko, P.B. (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution*, **19,** 2005–2021.

Minin, V.N., Bloomquist, E.W. & Suchard, M.A. (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular Biology and Evolution*, **25,** 1459–1471.

Norris, R.D. (2000) Pelagic species diversity, biogeography, and evolution. *Paleobiology*, **26,** 236–258.

O'Connor, M.I., Bruno, J.F., Gaines, S.D., Halpern, B.S., Lester, S.E., Kinlan, B.P. & Weiss, J.M. (2007) Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences USA*, **104,** 1266–1271.

Osorio, C., Bahamonde, N. & López, M.T. (1967) El limanche [*Emerita analoga* (Stimpson)] en Chile. *Boletin Museo Nacional de Historia Natural, Santiago*, **29,** 61–116.

Palumbi, S.R. (1994) Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology and Systematics*, **25,** 547–572.

Pelc, R.A., Warner, R.R. & Gaines, S.D. (2009) Geographical patterns of genetic structure in marine species with contrasting life histories. *Journal of Biogeography*, **36,** 1881–1890.

Posada, D. & Buckley, T.R. (2004) Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53,** 793–808.

Rambaut, A. & Drummond, A.J. (2007) *Tracer v1.4*. Available at: http://beast.bio.ed.ac.uk/Tracer

Rocha, L.A. & Bowen, B.W. (2008) Speciation in coral-reef fishes. *Journal of Fish Biology*, **72,** 1101–1121.

Scheltema, R.S. (1988) Initial evidence for the transport of teleplanic larvae of benthic invertebrates across the East Pacific Barrier. *Biological Bulletin*, **174,** 145–152.

Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin ver. 2.000: a software for population genetics data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Shapiro, B., Rambaut, A. & Drummond, A.J. (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Molecular Biology and Evolution*, **23,** 7–9.

Sorte, C.J., Peterson, W.T., Morgan, C.A. & Emmett, R.L. (2001) Larval dynamics of the sand crab, *Emerita analoga*, off the central Oregon coast during a strong El Niño period. *Journal of Plankton Research*, **23,** 939–944.

Tam, Y.K., Kornfield, I. & Ojeda, F.P. (1996) Divergence and zoogeography of mole crabs, *Emerita* spp (Decapoda: Hippidae), in the Americas. *Marine Biology*, **125,** 489–497.

Thompson, J.N. (1999) Coevolution and escalation: are ongoing coevolutionary meanderings important? *The American Naturalist*, **153,** S92–S93.

Thompson, W.G. & Goldstein, S.L. (2006) A radiometric calibration of the SPECMAP timescale. *Quaternary Science Reviews*, **25,** 3207–3215.

Walsh, P.S., Metzger, D.A. & Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, **10,** 506–513.

Ward, S.N. & Valensise, G. (1994) The Palos Verdes terraces, California: bathtub rings from a buried reverse fault. *Journal of Geophysical Research*, **99,** 4485–4494.

Zwickl, D.J. (2006) *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. PhD Thesis, University of Texas at Austin, Austin, TX.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Supplementary materials and methods, results and figure.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

The authors share research interests in the ecology and evolution of marine invertebrates. They study the geographic and temporal distributions of marine biodiversity from the perspectives of population genetics, systematics and population and community ecology. They seek to explain the relationships between life histories, geography, genetic structure and the scale of adaptation and diversification in the sea.

Author contributions: all authors conceived the study; P.H.B., M.N.D. and R.J.T. collected the data; P.H.B., M.N.D. and L.I.G.-G. analysed the data; and P.H.B. and M.N.D. led the writing with R.J.T., R.K.G., L.I.G.-G. and J.E.D.

Editor: Craig McClain