2014

Return of the Ghosts of Dispersal Past: Historical Spread and Contemporary Gene Flow in the Blue Sea Star Linckia Laevigata

Eric D. Crandall
California State University, Monterey Bay, ecrandall@csumb.edu

Recommended Citation

Crandall, E.D., Treml, E.A., Liggins, L., Gleeson, L., Yasuda, N., Barber, P.H., Wörheide, G. and C. Riginos. 2014. Return of the ghosts of dispersal past: historical spread and contemporary gene flow in the blue seastar Linckia laevigata. Bulletin of Marine Science 90(1): 399-425.

Bull Mar Sci. 90(1):399–425. 2014
http://dx.doi.org/10.5343/bms.2013.1052

research paper

Return of the ghosts of dispersal past: historical spread and contemporary gene flow in the blue sea star *Linckia laevigata*

Eric D Crandall [1*]
Eric A Treml [2]
Libby Liggins [3]
Lachlan Gleeson [3]
Nina Yasuda [4]
Paul H Barber [5]
Gert Wörheide [6]
Cynthia Riginos [3]

[1] *Department of Ocean Sciences, University of California, and Southwest Fisheries Science Center, Fisheries Ecology Division, 110 Shaffer Rd., Santa Cruz, California 95060.*

[2] *Department of Zoology, University of Melbourne, Melbourne, Vic., 3010, Australia.*

[3] *School of Biological Sciences, The University of Queensland, St. Lucia, Queensland 4072, Australia.*

[4] *Department of Marine Biology and Environmental Sciences, University of Miyazaki, Faculty of Agriculture, Gakuen-kibanadai-nishi-1-1, Miyazaki, Japan 889-2192.*

[5] *Department of Ecology and Evolutionary Biology, University of California Los Angeles, 621 Charles E. Young Dr. South, Los Angeles, California 90095.*

[6] *Department of Earth and Environmental Sciences, Division of Paleontology & Geobiology & GeoBio-Center LMU, Ludwig-Maximilians-Universität München, and Bavarian State Collections of Palaeontology and Geology, Richard-Wagner-Straße 10, 80333 Munich, Germany.*

[*] *Corresponding author email: <eric.crandall@noaa.gov>.*

Date Submitted: 7 June, 2013.
Date Accepted: 30 October, 2013.
Available Online: 4 December, 2013.

ABSTRACT.—Marine animals inhabiting the Indian and Pacific oceans have some of the most extensive species ranges in the world, sometimes spanning over half the globe. These Indo-Pacific species present a challenge for study with both geographic scope and sampling density as limiting factors. Here, we augment and aggregate phylogeographic sampling of the iconic blue sea star, *Linckia laevigata* Linnaeus, 1758, and present one of the most geographically comprehensive genetic studies of any Indo-Pacific species to date, sequencing 392 base pairs of mitochondrial COI from 791 individuals from 38 locations spanning over 14,000 km. We first use a permutation based multiple-regression approach to simultaneously evaluate the relative influence of historical and contemporary gene flow together with putative barriers to dispersal. We then use a discrete diffusion model of phylogeography to infer the historical migration and colonization routes most likely used by *L. laevigata* across the Indo-Pacific. We show that estimates of genetic structure have a stronger correlation to geographic distances than to "oceanographic" distances from a biophysical model of larval dispersal, reminding us that population genetic estimates of gene flow and genetic structure are often shaped by historical processes. While the diffusion model was equivocal about the location of the mitochondrial most recent common ancestor (MRCA), we show that gene flow has generally proceeded in a step-wise manner across the Indian and Pacific oceans. We do not find support for previously described barriers at the Sunda Shelf and within Cenderwasih Bay. Rather, the strongest genetic disjunction is found to the east of Cenderwasih Bay along northern New Guinea. These results underscore the importance of comprehensive range-wide sampling in marine phylogeography.

Edited by Managing Editor Rafael J Araújo

Marine biodiversity is concentrated in the Indo-Pacific region, with species diversity reaching its highest values in the Coral Triangle, a region centered in the Indo-Malay-Philippines archipelago (Roberts et al. 2002, Carpenter and Springer 2005, Tittensor et al. 2010). The Coral Triangle has a complex geological history with much tectonic activity, including substantial reconfigurations of landmasses due to moving plates (Hall 2002). Fluctuating sea levels have also substantially restructured land and sea configurations, as the shallow continental shelf is extensive in this region (Voris 2000). For example, sea levels 18,000 yrs ago are thought to have been about 130 m lower than present day levels, resulting in greatly reduced area for most shallow marine habitats and thus severe reductions in local population sizes (Crandall et al. 2012a). At that time, the Makassar Strait was much narrower than it is today, almost completely blocking the marine connection between the Pacific and Indian oceans (Chappell and Shackleton 1986). As sea levels rose over the next 10,000 yrs, continental shelf habitat of an area slightly smaller than the land area of the country of India (approximately 3.16×10^6 km^2) would have resubmerged and become available again to local marine species (Voris 2000).

Perhaps unsurprisingly, phylogeographic studies have revealed the imprint of these sea level changes on population genetic patterns of Coral Triangle species (reviewed in Carpenter et al. 2011). Some taxa show nearly reciprocal monophyly in mtDNA over relatively short distances, probably as a consequence of isolation during the Pleistocene (e.g., Barber et al. 2000, 2002). For many taxa, interestingly, the locations of likely vicariance due to Pleistocene sea level change are also associated with regions of persistent population genetic structure, most notably the Sunda shelf (McMillan et al. 1999, DeBoer et al 2008, Ackiss et al. 2013), Torres Strait (Mirams et al. 2011), and Halmahera Eddy (Barber et al. 2006, 2011), although there is also evidence that this structure is being eroded by contemporary dispersal in some species (Gaither et al. 2011, Liu et al. 2012, DeBoer et al. 2014). Finally, nearly all loci and taxa that have been examined to date show signatures of Pleistocene-era population expansions onto newly submerged continental shelf habitats (e.g., Chenoweth et al. 1998, Lind et al. 2007, Crandall et al. 2008a,b, 2012a, Gaither et al. 2010).

Given the globally acknowledged value of the Coral Triangle in terms of marine biodiversity as well as the multitude of current threats to the region (Roberts et al. 2002, Burke et al. 2011), it is of practical importance to understand both the history of its marine communities (how and from where did genetic and species diversity arise? e.g., Renema et al. 2008, Williams and Duda 2008, reviewed in Bowen et al. 2013), as well as how the regional seascape is presently constructed (how are different parts of the region connected demographically by larval dispersal? reviewed in Riginos and Liggins 2013). Previous phylogeographic attempts to answer these questions have focused on the measurement of genetic structure (F_{ST} and its analogues, Wright 1950), and testing specific hypotheses of population structure primarily with analyses of molecular variance (AMOVA, Excoffier et al. 1992). However, these approaches based on allele frequencies are poorly suited to organisms with evolutionarily high levels of gene flow (10–100 migrants per generation) and large coalescent effective population sizes (partially arising from high gene flow), which depress traditional estimates of population structure (such as F_{ST}, Hedrick 2005). Furthermore, because F_{ST} and AMOVA summarize a combined model of gene flow and effective population size (Whitlock 2011), these frequency-based approaches allow estimation of marine population structure only at a very coarse resolution, often resulting

in substantial bias when parameters such as effective population size and timing of population divergence are not considered (Bird et al. 2011, Marko and Hart 2011, Faurby and Barber 2012). As an alternative, one might consider the inverse approach: testing explicit hypotheses of gene flow rather than genetic structure (Crandall et al. 2012b).

The substitution of gene flow for genetic structure as the parameter of interest in seascape genetics makes intuitive sense. Gene flow in most marine organisms is mediated by the planktonic larval stage, where millions of larvae disperse through a complex milieu of currents and environmental conditions. Thus, there are few impermeable barriers to larval dispersal in the ocean: for almost every physical or oceanographic entity that is thought to impede larval dispersal for some species (see examples in Rocha et al. 2007) one can find several species that show no evidence of isolation whatsoever (Lessios and Robertson 2006, Carpenter et al. 2011, Toonen et al. 2011). Although ocean currents and land masses may effectively act as barriers to gene flow, it is more appropriate to think in terms of probabilistic larval dispersal kernels for which the probability of a successful dispersal event (and therefore of gene flow) declines sharply with distance due to larval diffusion, behavior and mortality (Cowen et al. 2000, Gerlach et al. 2007, Buston et al. 2012). The most informed hypotheses of gene flow therefore come from modeling such dispersal through oceanographic current vectors while taking these additional factors into account (Kool et al. 2011, Treml et al. 2012). Empirical testing with genetic data has shown this to be the case using coalescent estimates of gene flow (Crandall et al. 2012b), assignment tests (Fievet et al. 2007), parentage-based tagging (Saenz-Agudelo 2012) or F_{ST} (Galindo and Palumbi 2006, White et al. 2010, Alberto et al. 2011, Foster et al. 2012). Until recently, gene flow estimates have been constrained to the unrealistic assumptions of Wright's island model by relying on the F_{ST} summary statistic (Whitlock and McCauley 1999), which assumes equal levels of gene flow throughout the sampled area and does not take historical factors (such as lineage sorting) into account. However, coalescent modeling approaches now allow flexible evaluation and selection of specific models of gene flow, which are estimated simultaneously with the genealogical history of genetic sequence data (Hey and Nielsen 2007, Lemey et al. 2009, Beaumont 2010, Beerli and Palczewski 2010).

A further challenge to understanding genetic diversity in the Coral Triangle arises from its location: it exists near the junction of the Indian and Pacific Oceans, at the center of the Indo-Pacific region, which is the largest biogeographic region on Earth (Spalding et al. 2007). Many species found in the Coral Triangle have vast ranges that may include large portions of the Indian and Pacific oceans. Although there have been considerable recent efforts to document genetic patterns within the Coral Triangle, there are few studies with dense sampling (many locations, many individuals per location) that include both the Coral Triangle and surrounding regions of the Indo-Pacific (see Keyse et al. 2014). Although logistically challenging, large-scale geographic coverage is necessary to determine the context of genetic variation of marine species. Without broad-scale sampling that includes both the Coral Triangle and other parts of the species' range, it is not possible to completely resolve the extent of divergent genetic lineages (Manel and Holdregger 2013).

The sea star *Linckia laevigata* Linnaeus, 1758, easily recognizable for its striking blue coloration, is one of the best-studied species in the Indo-Pacific region. Besides the well-known royal blue phenotype, several color variations are reported

from different geographic regions without apparent differentiation by morphotype (Williams 2000). Like many other marine benthic taxa, adult *L. laevigata* are sedentary, but the larvae have a moderate (at least 22 d before metamorphosis, Yamaguchi 1973) pelagic larval duration (PLD) after external fertilization.

Early genetic surveys of *L. laevigata* were based on allozymes (Williams and Benzie 1993, Williams and Benzie 1996, 1998, Williams et al. 2002) or mtDNA (Williams and Benzie 1997, 1998, Williams et al. 2002, Crandall et al. 2008b, Kochzius et al. 2009) and have had broad coverage from the western Pacific to the Indian ocean without much sampling in the Coral Triangle (e.g., Williams and Benzie 1998) or have exclusively focused sampling within the Coral Triangle (Crandall et al. 2008b, Kochzius et al. 2009). At the broadest scale, support was found for differentiation between the Indian and Pacific oceans based on allozymes, albeit with individuals from Western Australia (Ningaloo) showing greater affinity to western Pacific populations than to western Thailand and South Africa (Williams and Benzie 1998, Williams et al. 2002, see also Vogler et al. 2013 for similar results in crown-of-thorns starfish). For mtDNA COI sequences, Williams et al. (2000) described two major clades: an "Indian Ocean" clade, which included both Pacific and Indian Ocean individuals and a "Pacific Ocean" clade that contains only Pacific Ocean individuals with the exception of a few western Australian individuals. Crandall et al. (2008b) as well as Kochzius et al. (2009) sampled comprehensively within the Coral Triangle and found that the Indian clade haplotypes dominate most Indonesian populations, declining in frequency from Aceh in the west to Jayapura in the east. Within the Coral Triangle, the greatest population structure was found by grouping Aceh with Krakatau against a cluster of remaining locations, whereas less support was found for a west-east delineation defined by the Sunda Shelf, and there was modest support for distinctiveness of Teluk Cenderawasih (Crandall et al. 2008b). Kochzius et al. (2009) suggested that *L. laevigata* has historically expanded populations into the western Pacific from eastern Indian Ocean origins, a route of colonization running counter to the Indonesian Throughflow (Kochzius et al. 2009).

In the present study we combine mtDNA data from some of the previous studies cited above with new sampling from 18 additional locations to evaluate the influence of putative historical and contemporary gene flow and spatial features on genetic patterns within *L. laevigata*. Although data from additional loci are desirable for future studies, the rich genealogical information in this mtDNA-only data set allows inference from the recent (approximately 100 kya) matrilineal history of this species and the cross compatibility of this DNA sequence data facilitates synergism across research groups (Bowen et al. 2014). The data set analyzed here represents the single most extensive population genetic survey both in geographic extent (diameter >14,000 km) and density for any species to date from the Indo-Pacific region (Keyse et al. 2014); its compilation was only possible by cooperation and collaboration among research groups.

We use this data set to consider the genealogical history of *L. laevigata* in terms of inferred mitochondrial gene flow in addition to estimates of genetic structure. We begin by parameterizing a biophysical model of contemporary larval dispersal to predict mean dispersal distances for *L. laevigata* larvae among sampling sites. We then use a multiple regression approach based on permutation (Legendre et al. 1994) to ask whether measurements of genetic structure in this species are better explained by the mean dispersal distances from the biophysical model, or by geographic

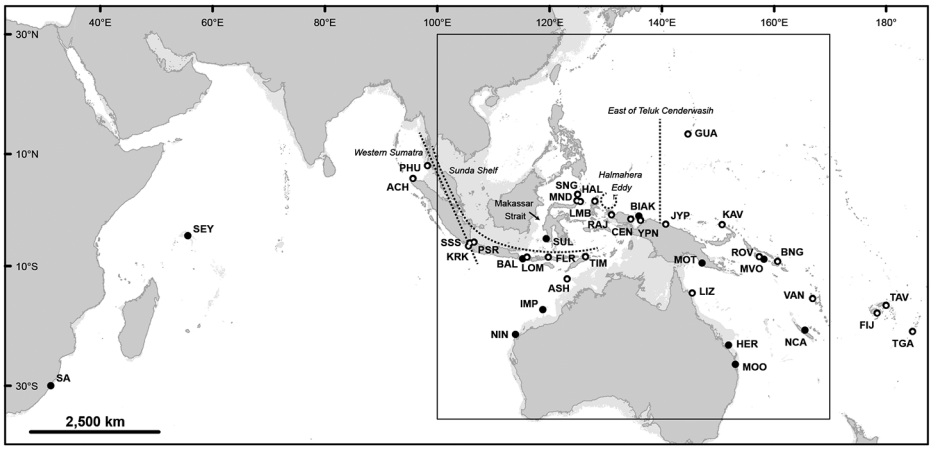

Figure 1. Sampling locations; see Table 1 for abbreviations. Open circles represent locations with sample size >10. The domain for the biophysical model is represented by a box. Four putative barriers to gene flow that were evaluated under the MRDM and AMOVA approaches are denoted as dotted lines.

distances. The biophysical model assesses whether mitochondrial patterns are best explained by equilibrium gene flow from contemporary larval dispersal whereas the geographic distance model assesses the relative importance of colonization processes and historical gene flow. We also use the biophysical model output to define modular geographic regions with greater larval connectivity within each region than between regions. These regions are then used in a phylogeographical model that reconstructs historical gene flow through time by treating each region as a character state that can be inferred for each node on the genealogy through ancestral state reconstruction (Lemey et al. 2009). This approach allows us to infer the vectors of gene flow that were important in the spread of this lineage throughout the Indo-Pacific region.

Methods

Study Sites and COI Sequencing.—Mitochondrial cytochrome oxidase I sequences from blue morph *L. laevigata* were obtained from both newly sampled individuals and from previous studies (Williams 2000, Crandall et al. 2008b). New samples were obtained from locations in western and eastern Australia, East Timor, Papua New Guinea, the Solomon Islands, Vanuatu, Fiji, and Tonga (see Fig. 1 and Table 1 for more details), using tube feet preserved in ethanol. New sequences were also obtained from locations in Thailand, Indonesia, Guam, and New Caledonia using old pyloric caecum samples preserved in DMSO, which were previously analyzed in an allozyme study (Williams et al. 2002). DNA preparation was undertaken at the University of Queensland (including NIN, ASH, TIM, KAV, MVO, MOT, LIZ, HER, MOO, TGA, ROV; see Table 1 for abbreviations), Boston University (including ROV, BNG, VAN, FIJ, and TAV) and Ludwig-Maximilians-Universität München (including LOM, PHU, NCA, and GUA). At the University of Queensland, genomic DNA was extracted using a modified salt extraction protocol (Aljanabi and Martinez 1997). The cytochrome oxidase I (COI) gene was amplified using polymerase chain reaction (PCR), with each reaction containing approximatley 1 µl of DNA and using Titanium

Table 1. Sampling locations and summary statistics. Number of samples (*n*), Haplotype diversity (H), percent segregating sites (S), nucleotide diversity (π), Tajima's D, and Fu's F_S.

Location	CODE	Latitude	Longitude	Source	*n*	H	S (%)	π (%)	D	F_S
South Africa	SA	−30.1	31.2	Williams 2000, 2002	5	1.00	1.8	0.9	0.50	−1.90
Seychelles	SEY	−4.6	55.6	Williams 2000, 2002	4	0.25	0.0	0.0	0.00	0.00
Ningaloo Reef, Australia	NIN	−21.7	114.0	Present study: Riginos	7	0.86	2.8	1.2	0.44	−1.19
Imperieuse Reef, Australia	IMP	−17.5	118.8	Williams 2000, 2002	3	1.00	2.0	1.4	0.00	0.46
Ashmore Reef, Australia	ASH	−12.2	123.1	Present study: Riginos	15	0.80	6.1	1.9	−0.04	−3.12
East Timor	TIM	−8.3	126.4	Present study: Riginos	19	0.68	4.8	1.1	−0.76	−4.90*
Aceh, Indonesia	ACH	5.6	95.7	Crandall et al. 2008	15	0.73	4.3	1.0	−0.96	−4.31**
Phuket, Thailand	PHU	7.9	98.3	Present study: Yasuda	28	0.75	6.9	1.0	−1.59*	−15.47
Krakatau	KRK	−6.1	105.5	Crandall et al. 2008	48	0.65	9.4	1.0	−1.79*	−25.36
Sebesi/Sebuku/Sangiang	SSS	−5.9	105.5	Crandall et al. 2008	51	0.51	9.4	1.4	−1.19	−11.13**
Pulau Seribu	PSR	−5.7	106.6	Crandall et al. 2008	79	0.48	10.7	1.4	−1.19	−22.60
South Sulawesi	SUL	−5.1	119.4	Crandall et al. 2008	7	0.86	3.8	1.5	−0.28	−0.84
Bali	BAL	−8.7	115.3	Crandall et al. 2008	5	0.80	3.6	1.7	0.09	0.98
Lombok	LOM	−8.4	116.0	Crandall et al. 2008, Present study: Yasuda;	23	0.70	4.3	1.2	0.05	−7.10***
Flores	FLR	−8.4	119.8	Crandall et al. 2008	14	0.79	4.8	1.4	−0.39	−3.53*
Manado	MND	1.6	124.9	Crandall et al. 2008	76	0.43	10.5	1.3	−1.21	−16.03***
Lembeh	LMB	1.5	125.2	Crandall et al. 2008	20	0.60	6.1	1.4	−0.70	−2.25
Sangihe	SNG	2.8	125.4	Crandall et al. 2008	17	0.47	3.1	0.8	−0.48	−1.20
Halmahera	HAL	1.5	128.0	Crandall et al. 2008	75	0.41	11.2	1.3	−1.37	−13.45***
Raja Ampat	RAJ	−0.9	131.1	Crandall et al. 2008	31	0.42	5.4	1.2	−0.39	−1.85
TelukCenderawasih	CEN	−1.7	134.5	Crandall et al. 2008	22	0.68	5.9	1.6	−0.10	−4.34*
Biak	BIAK	−1.1	136.0	Crandall et al. 2008	7	0.86	4.8	1.9	−0.24	−0.38
Yapan	YPN	−1.9	136.2	Crandall et al. 2008	19	0.42	3.8	0.9	−0.79	−0.62
Jayapura	JYP	−2.5	140.7	Crandall et al. 2008	19	0.58	4.1	1.1	−0.23	−2.54
Guam	GUA	13.5	144.7	Present study: Yasuda; Williams 2000, 2002	25	0.68	5.4	1.3	−0.21	−6.82**
Kavieng, PNG	KAV	−2.6	150.8	Present study: Riginos	16	0.69	4.1	1.3	0.06	−2.97

Table 1. Continued.

Location	CODE	Latitude	Longitude	Source	*n*	H	S (%)	π (%)	D	F_S
Roviana, SOL	ROV	−8.3	157.4	Present study: Riginos	18	0.72	5.9	1.3	−0.91	−4.46*
Marovo, SOL	MVO	−8.8	158.3	Present study: Riginos	5	1.00	2.8	1.3	−0.38	−1.35
Boneagi, SOL	BNG	−9.2	160.7	Present study: Crandall	10	0.60	2.8	1.1	0.59	0.09
Motupore, PNG	MOT	−9.4	147.2	Present study: Riginos	8	0.88	3.6	1.3	−0.11	−1.78
Lizard Island, GBR	LIZ	−14.7	145.5	Present study: Riginos	16	0.81	6.6	1.5	−0.92	−4.75*
Heron Island, GBR	HER	−23.5	151.9	Present study: Riginos	9	0.89	5.1	1.6	−0.70	−2.10
Mooloolaba, QLD	MOO	−26.6	153.1	Present study: Riginos	3	1.00	0.8	0.5	0.00	−0.69
New Caledonia	NCA	−21.0	165.6	Present study: Yasuda	5	0.80	2.0	1.0	0.29	0.13
Vanuatu	VAN	−15.6	167.0	Present study: Crandall	16	0.63	5.4	1.1	−1.31	−2.35
Viti Levu, Fiji	FIJ	−18.1	178.4	Present study: Crandall	20	0.40	4.1	1.0	−0.39	0.08
Taveuni, Fiji	TAV	−16.8	180.0	Present study: Crandall	10	0.80	3.8	1.5	0.37	−1.62
Tonga	TGA	−21.2	−175.3	Present study: Riginos	12	0.75	2.6	0.8	−0.42	−4.06**

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Taq polymerase (Clontech Laboratories, Inc.) and the universal COI primers from Folmer et al. (1994). Amplicons were purified with an exo/sap procedure (New England Biolabs) and sent to Macrogen (Korea) for capillary sequencing. At Boston University, sequencing protocols followed those outlined in Crandall et al. 2008b. At Ludwig-Maximilians-Universität München, newly developed *L. laevigata*-specific primers (LL-F1, 5′–ACCACCGGCTGGGTCGAA–3′ and LL-R1, 5′–TAATCTTTGGGGCGTGAGC–3′) were used for PCR and sequencing for three populations (PHU, GUA, and NCA), to increase PCR efficiency. The amplifications were made in 10 µl reactions with a final concentration of 3mM of MCl2, 1× buffer, 0.3 µM of each primer, 0.8 mM of dNTP, 0.07 units of Kapa Taq DNA Polymerase (Kapa biosystems) and 1 µl of DNA template. Thermocycling consisted of denaturation of DNA at 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 50 °C for 30 s and 72 °C for 30 s, followed by final extension of DNA at 72 °C for 7 min and cooling at 4 °C. The amplification of PCR products was confirmed by gel electrophoresis in a 1% agarose gel. Sequencing reactions were conducted on an ABI 3730 capillary sequencer using BigDye v.3.1. Sequence was determined from both directions using LL-F1 and LL-R1 primers. Trace files were manually checked and edited using CodonCode v3.0.2. These newly generated sequences were aligned against published sequences from Williams (2000), Williams et al. (2002), and Crandall et al. (2008b). Individual haplotypes from Crandall et al. (2008b) were assigned in the same manner as the original paper.

In total, we aggregated mtDNA haplotypes from 38 localities (Fig. 1) and data from all locations were represented in the haplotype network (Fig. 2). To obtain an accurate and unbiased estimate of pairwise genetic relationships between locations (Φ_{ST}, F_{ST}, and D_{est} with respect to Euclidean distance and overwater distances), we included only those locations which had at least 10 individuals sampled, leaving 26 populations for these analyses (Fig. 1, sites with white centers). Those locations within the domain of the biophysical model (box in Fig. 1) were used for two purposes. First, those sample sites with more than 10 individuals sampled, that also contained substantial reef habitat were used for reanalyzing the pairwise genetic relationships between locations with respect to dispersal distances derived from the biophysical model (20 total sites; all white sites within box in Fig. 1 excluding Jayapura, which does not have substantial reefs, and is thus not included in the biophysical model). Second, all collection locations within the model domain were used, except Jayapura and Mooloolaba, due to reef representation to extract the geographic cluster membership of the remaining 29 sample locations.

Genetic Diversity and Differentiation.—To visualize the total COI diversity, a median joining haplotype network was constructed in Network 4.611 and edited in Network Publisher 2.0 (fluxus-engineering.com, Bandelt et al. 1999). To reduce complexity, non-parsimonious links were deleted using the maximum parsimony calculation option (MP, Polzin and Daneshmand 2003). Colors across the haplotype network represented different biophysically derived clusters (see below) and other locations/regions beyond the scope of the biophysical model, but assumed to be distinct demographically due to their geographic isolation, and significant genetic structure (Online Table S1).

Standard population summary statistics were calculated in Arlequin v3.5.1.3 (Excoffier and Lischer 2010). Statistics calculated included Watterson's θ (Watterson

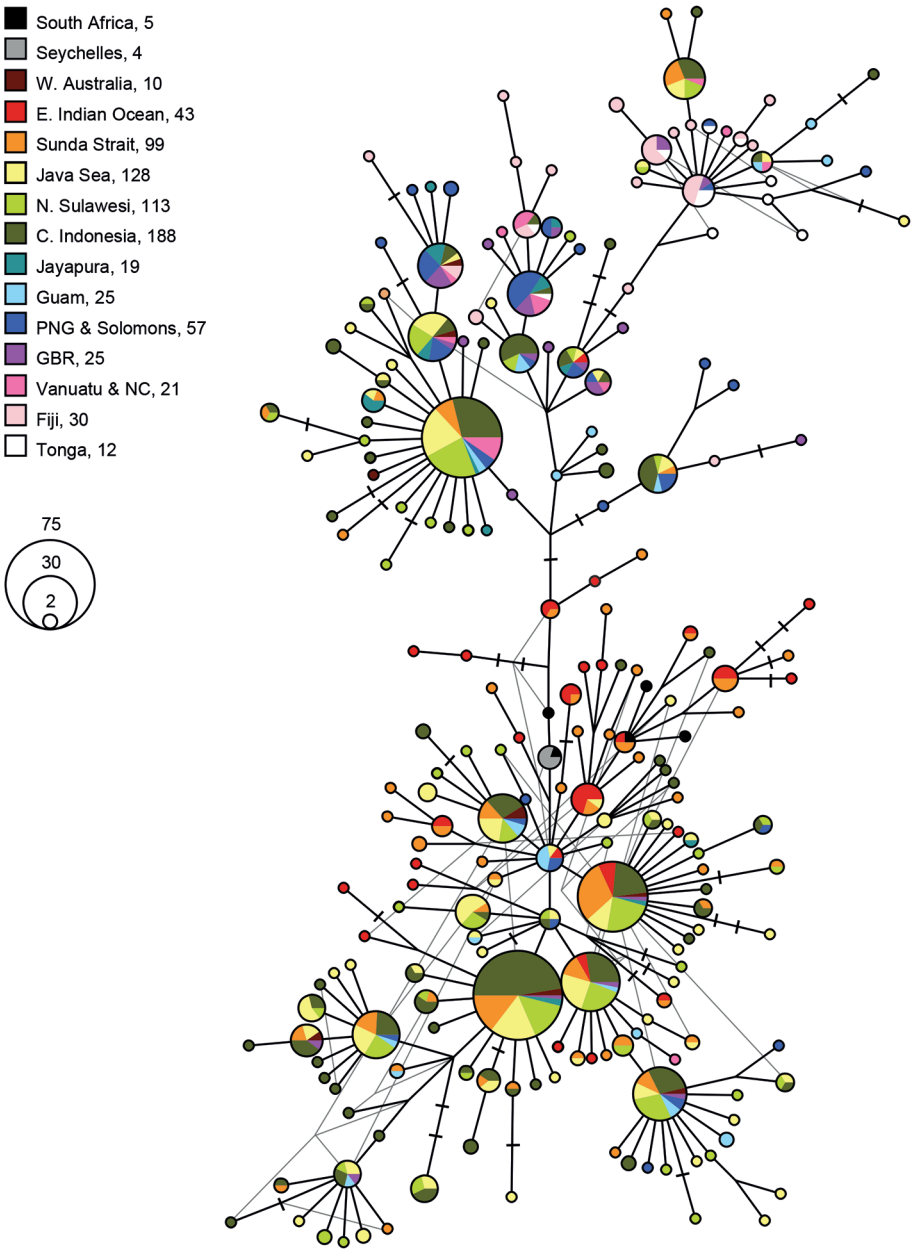

Figure 2. Median-joining haplotype network for *Linckia laevigata* colored according to modular clusters determined from the biophysical model. The number of individual sequences included in each cluster follow the name in the key. The frequency of each haplotype is indicated by size (see key, bottom right). Edges between haplotypes or small cross-bars indicate a mutational step. Black edges represent one of the maximum parsimony networks chosen at random; grey edges represent alternate relationships among haplotypes found in 29 other equally parsimonious networks.

1975), which estimates the average number of polymorphic sites, and π, which estimates the average number of differences between two random sequences from the same population. Tajima's *D* (Tajima 1989) and Fu's F_s (Fu 1997) were also measured to compare *L. laevigata* COI diversity against neutral, equilibrium expectations.

We used a variety of approaches to summarize genetic differentiation. Pairwise relationships among populations (for sampling locations including 10 individuals or more: 26 populations) were described by Φ_{ST} (based on the Tamura-Nei distance as selected by jModeltest; Posada 2009) as well as F_{ST} (based on haplotype identities) in Arlequin. To reduce the effects of high allelic variability, we also used D_{est} (Jost 2008) calculated in Genodive (Meirmans and Van Tienderen 2004) by reducing infiles to haplotype identity. Relationships among populations based on these pairwise statistics were visualized using non-metric multidimensional scaling (vegan package in R, Oksanen et al. 2012).

We used AMOVA (Excoffier et al. 1992) analysis in Arlequin to test a priori predictions about various spatial configurations of genetic differentiation arising from barriers to dispersal. To test the hypothesis that divergence between Indian and Pacific populations explains the most molecular variance in COI, we tried two different spatial configurations. We first separated locations in western Indonesia from central Indonesia (following Crandall et al. 2008b), and then delineated Indian and Pacific populations along the Sunda Shelf. The contribution of the Halmahera Eddy to population structure was assessed by configurations whereby central Indonesia locations were grouped separately from Cenderwasih locations, and these groupings were contrasted against configurations that combined central Indonesia, and Cenderwasih locations as a single group. Similarly the distinctiveness of Cenderwasih from the Pacific was evaluated by comparing separate and combined groupings. Furthermore, Guam was allowed to group by itself and with other Pacific locations due to preliminary analyses pointing to genetic distinctiveness of this location. These varying spatial configurations were evaluated using AMOVA based on Φ-statistics derived from Tamura-Nei distances between haplotypes and from haplotypes identified in Arlequin (as above). For AMOVAs based on haplotype identity, we also used *F′*-statistics (Meirmans and Van Tienderen 2004).

Geographic and Biophysically Informed Predictions of Gene Flow Among Sample Locations.—Three pair-wise distance metrics were calculated as proxies for gene flow (i.e., realized dispersal over many generations) among sampled populations. As a simple null model of gene flow, we calculated the Euclidean distance among all 38 sampled locations. Although this model of gene flow implies that dispersal occurs relative to straight-line routes only, it is a common dispersal distance proxy used in marine population genetic studies (Riginos and Liggins 2013). The second dispersal distance proxy used for all locations was the shortest overwater distance calculated with a least-cost path algorithm. The distance calculations are similar to the Euclidean distance, but the least-cost path was forced around all land boundaries. These two measures of geographic distance are expected to explain more of the variance in genetic structure when historical processes such as colonization dominate the signal from gene flow (Selkoe 2008).

Finally, we used a biophysical model of larval dispersal (Treml et al. 2012) to quantify the relative dispersal strength among sampled populations within the model domain (Fig. 1). This dispersal model includes coral reef habitat (Spalding et al. 2001),

oceanographic data describing sea surface currents for three years (ROMS, Wang et al. 2005), and several biological parameters describing the dispersal characteristics of *L. laevigata*: seasonal spawning periodicity, a 25-d maximum pelagic larval duration, a 1- to 2-d precompetency period, strong late-stage swimming/homing behavior, and a 30% d^{-1} larval mortality (Yamaguchi 1973). The model outputs the probability that larvae released in one location survive and settle in every other recipient location, summarized as a 1002 × 1002 source-reef by destination-reef matrix. This dispersal probability matrix was converted to a migration matrix representing the proportion of settlers to every reef patch that came from all upstream larval sources. See Treml et al. (2012) for model details and sensitivity analysis. The migration matrix, M, was converted to "oceanographic dispersal distance" using $\log(M^{-1})$ to transform the values to be the same rank-order as geographic distance (high proportion of settlers then have a short distance) required for many network-based algorithms. This inverse dispersal strength matrix was used as a proxy for dispersal distance, and is referred to as such throughout this paper. This oceanographic dispersal distance is expected to explain more of the variance in genetic structure when contemporary dispersal events dominate the signal relative to historical connections (White et al. 2010).

To identify the emergent geographic clustering of reef habitat (and sample locations) determined by the dispersal strengths represented in the migration matrix, we used a network-based leading eigenvector community detection algorithm (Newman 2006). This algorithm identifies the optimal clustering within a network by optimizing the network's modularity, or simply maximizing the density of within-cluster connections while minimizing between-cluster connections. The original asymmetric migration matrix was converted to a symmetric matrix by taking the maximum dispersal strength between all pairs of reefs. *Linckia laevigata* sample sites were overlaid with the network clustering results thereby revealing the potential clustering of sample sites based on dispersal potential among all reefs. These spatial clusters were then used to aggregate sampling sites in the phylogeographic diffusion model described below.

Evaluating Genetic Differentiation Using Dispersal Proxies.—We used a multiple regression on distance matrices (MRDM, Legendre et al. 1994) to evaluate the relative influence geographic distance and multiple discrete landscape factors that might contribute to population genetic structure; this methodology performs favorably compared to many other methods (Balkenhol et al. 2009). In a simple model where the linear relationship between geographic distance and genetic distance are evaluated, MRDM is equivalent to a Mantel test. Dispersal distance proxies included Euclidean distances and overwater distance, as well as the biophysical dispersal distances between populations. The biophysical distances represent a hypothesis of contemporary dispersal. Euclidean and overwater distances could represent simpler (null) models of contemporary dispersal, but they also might capture historical averages of dispersal (that is both recurrent gene flow and colonization). The predictive contributions of four putative barriers/divisions were also evaluated including (1) Western Sumatra, as found in Crandall et al.'s (2008) earlier survey of *L. laevigata;* (2) Sunda Shelf, a focal point of Pleistocene marine disjunctions due to the expansion of the Sunda Shelf land mass at low sea level stands; (3) Halmahera Eddy, a contemporary hydrodynamic barrier; and (4) Pacific east of Cenderwasih, which might

represent a location of habitat limitation. For barriers, dummy variables (0 vs 1) were coded with 0 for population pairs found on the same side of the putative barrier. Φ_{ST}, F_{ST}, and D_{est} values between pairs of populations were used as response variables. All variables were normalized and both forward and backward model selection were implemented as in Legendre et al. (1994). Models including Euclidean and overwater distances included all populations with sample size ≥10 (26 populations). Because biophysical distances were only available for some population pairs, the analyses involving these predictors was reduced to a more restricted subset of populations (20 populations). In addition, the decomposed matrices involving biophysical predictors were twice as large because the distances are asymmetric (distance from *X* to *Y* does not equal *Y* to *X*). Permutated probabilities for all matrix regression models were evaluated with lmPerm (Wheeler 2010).

Bayesian Selection of Important Migration Parameters.—To estimate the historical gene flow required to explain the current distribution of mitochondrial genetic variation in *L. laevigata*, we modeled the phylogeographic history of COI as a discrete diffusion process following methods developed by Lemey et al. (2009). Using a coalescent perspective, we assigned a geographic location as a discrete character trait having one of 14 possible states (based on spatial clusters delineated by the biophysical model, see below) to each COI sequence, and reconstructed the most probable location of each ancestor back to the most recent common ancestor (MRCA) of the entire sample using BEAST 1.6.2 (Drummond and Rambaut 2007). Changes in location between ancestor and descendent nodes were modeled as a migration event, the probability of which was governed by a time-reversible matrix of migration rates among locations (similar to the GTR model used for models of molecular evolution). To reduce the degrees of freedom in the matrix, this method uses Bayesian stochastic search variable selection to only allow *W* of the rates to be non-zero, where the prior on *W* is a truncated Poisson distribution with mean η (see Lemey et al. 2009 for full details). In a Bayesian framework, the geographic location of the MRCA, as well the migration events required to explain most topologies can be estimated simultaneously with models of nuisance parameters such as those for molecular evolution, demography and tree topology.

Because our sequences came from 38 different localities, many with relatively small sample sizes, we assigned their geographic location as one of the k = 9 spatial clusters resulting from the clustering algorithm on the migration matrix plus the following groups of sites based on geographic isolation: Fiji/Taveuni, Tonga, Phuket/Aceh, Seychelles, and South Africa, for a total of 14 clusters (Fig. 3). We used a TN93 model of molecular evolution and an uncorrelated relaxed clock (Drummond et al. 2006) to model COI sequence evolution. Because this species and most other Indo-Pacific species appear to have a history of demographic fluctuation over evolutionary time, we implemented a Bayesian skyline model of demography as well (Drummond et al. 2005), which relaxes assumptions of any particular demographic history. We set a fairly uninformative truncated Poisson prior for the number of allowable migration rates within the matrix with an offset of k – 1 = 13 (the minimum number of rates required to connect all populations) and a mean of 10 (95% of the prior probability mass lies between 5 and 21 rates). All other priors were set to their defaults, and we elected not to use distance-informed priors on gene flow parameters, so that the model had no a priori information about the underlying geography. The

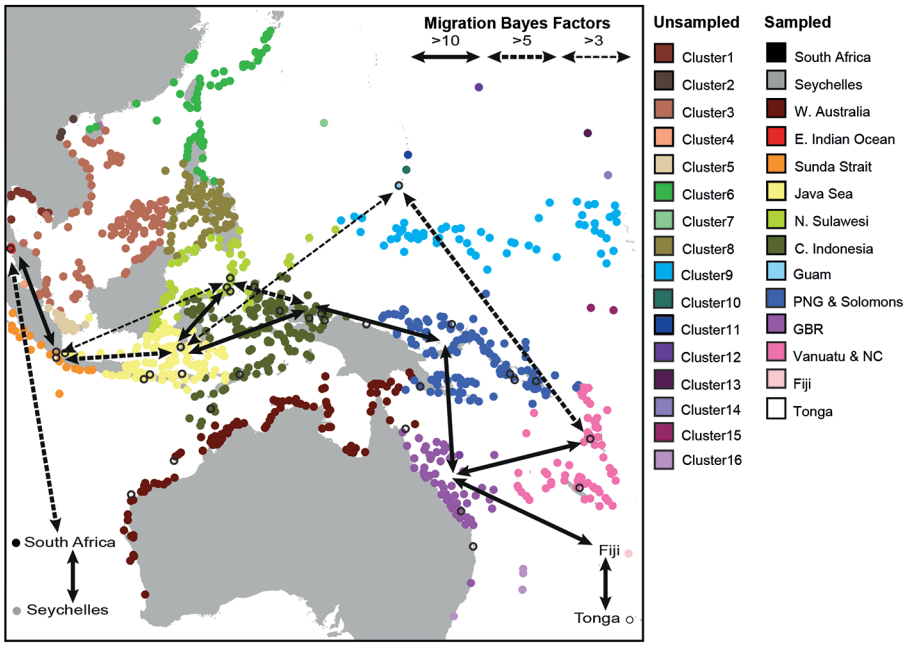

Figure 3. Best supported migration routes (Bayes Factor > 3) among modular population clusters delineated by the biophysical model. Optimal clusters were identified using the network modularity algorithm, and are shown in different colors. Each colored point represents the geographic centroid of reef patches used in the biophysical model. Open circles show genetic sample locations, as in Figure 1.

model was run four times for at least 50 million steps of Markov chain Monte Carlo (MCMC), and convergence was checked in Tracer 1.5. We then combined the logfiles and treefiles after trimming off an appropriate amount of burn in, and constructed a maximum clade credibility tree, from which we extracted the ancestral location probabilities. We established the significance of migration rate parameters if they were supported by a Bayes Factor of 3, which also corresponded with the migration rate parameter being required to explain >50% of sampled topologies.

Because MRCA location probability might be biased toward spatial clusters with a large sample size, we evaluated the prior distribution for root location by randomly swapping the location state among sequences during the MCMC chain such that sample size from each spatial cluster remained the same, but the sequences assigned to each location becomes random. This randomization, performed with the tip state operator in BEAST 1.7.5, as described by Edwards et al. (2011), removed location information from the dataset, allowing us to observe the prior expectation for root location.

Given that single-locus inference is still common for Indo-Pacific species (see discussion in Bowen et al. 2014), the genealogical approaches used here make much more effective use of the high information content in the mitochondrial locus than do estimates of genetic structure. Moreover, F_{ST} methods as well as more explicit population genetic models such as Migrate (Beerli and Felsenstein 2001) make an implicit assumption of genetic equilibrium (i.e., haplotypes have maintained the observed distribution for a long time about ½ N_E generations) that is violated by the dynamic changes in the marine habitats of the Coral Triangle over the past hundreds

of millennia. The spatial diffusion approach is not completely parameterized as a population genetic model (although Θ is still estimated as part of the skyline model) so it cannot make estimates for the amount of gene flow among populations as is done in Migrate and IMa. However, the absence of population genetic parameters allows a more flexible model that can reconstruct the most likely location of the mitochondrial common ancestor, together with the avenues of colonization and gene flow among multiple populations without prior knowledge of population history (unlike IMa2, Hey and Nielsen 2007; see also Bloomquist et al. 2010). As with all coalescent methods, the addition of sequence data from nuclear loci will eventually allow us broaden our inference to the demographic history of the species, rather than just that of the maternal lineage.

Results

Study Locations and COI Sequencing.—We obtained mtDNA sequences from a total of 791 *L. laevigata* individuals including 274 new sequences that greatly expanded the geographic scope of population sampling, especially from the western Pacific (Table 1, Fig. 1). To avoid using too much missing data by nucleotide position (<5%), we trimmed our alignment to 392 bp in length, containing a total of 110 substitution sites, and this portion of COI was used for subsequent analysis. This resulted in 209 unique haplotypes. New sequences have been deposited in Genbank (Accession Numbers KF834572–KF834833) and a complete Fasta-formatted file of the 791 individuals is deposited in http://www.datadryad.org.

Genetic Diversity and Differentiation.—Thirty equally parsimonious haplotype networks were recovered (one network chosen at random is shown in Fig. 2). The haplotype network highlights the high diversity of haplotypes and presence of many unique or private haplotypes. Consistent with previous mtDNA sequencing surveys (i.e., Williams 2000, Crandall et al. 2008b, Kochzius et al. 2009), there were two large emergent clusters. The haplotypes of sampled locations in the West and Central Pacific were largely restricted to one cluster (i.e., Vanuatu, New Caledonia, Tonga and parts of the Solomons, Fiji, and Papua New Guinea; top cluster, Fig. 2.) but not entirely (e.g., the sample from Taveuni in Fiji comprised individuals from both clusters). In contrast, locations sampled in the Indian Ocean had haplotypes that fell within the other cluster (bottom cluster, Fig. 2) or were intermediary to both clusters. Many haplotypes were shared across many regions, and especially within the Central Indo-Pacific locations, haplotypes from both clusters were common.

Individual populations varied in observed mtDNA diversity (Table 1) with haplotype diversity ranging from 0.42 to 0.80 in populations with $n > 10$. Several populations showed deviations from neutral equilibrium conditions especially as evaluated by Fu's (1997) F_S statistic. Among the 26 populations where 10 or more individuals were sampled, there was significant genetic differentiation among many population pairs regardless of the statistic used (Online Table S1). Non-metric multidimensional scaling (NMDS) based on F_{ST} values with Tamura-Nei distances among haplotypes recovered relationships among populations that roughly approximated geography (Fig. 3; non-metric stress = 0.045). Dimension 1 of the NMDS was very strongly correlated with longitude ($R^2 = 0.81$, $P < 1 \times 10^{-9}$), while there was no correlation of NMDS dimension 2 to latitude ($R^2 = 0.01$, $P = 0.58$).

All population groupings in hierarchical AMOVAs resulted in significant Φ_{CT} and $F\text{¢}_{CT}$ values, indicative of substantial regional population structure. The population grouping that consistently returned the highest Φ_{CT} and F'_{CT} values was a two-regional grouping whereby Pacific populations (all populations east of Cenderwasih Bay) were delineated from the remaining populations including those from Cenderwasih, the central Indo-Pacific, and the Indian Ocean populations ($\Phi_{CT} \leq 0.282$, $P < 0.001$; $F'_{CT} \leq 0.637$, $P < 0.038$). For all groupings, Φ_{SC} values were also significantly greater than zero ($P < 0.001$) and F'_{SC} were marginally significant ($P \leq 0.06$), indicative of genetic differentiation among populations within regions (Online Table S2).

Dispersal Distance Proxies.—The bivariate correlation between dispersal distance matrices, evaluated with the simple Mantel Test, revealed significant relationships for 20 localities within the Coral Triangle model domain. For the 190 site-pairs, the correlations between all distance-based dispersal proxies were high. The correlation coefficient for Euclidean distance and overwater distance was $r_M = 0.996$ (10,000 permutations, $P < 0.0001$). For Euclidean distance and mean dispersal distance, $r_M = 0.896$ (10,000 permutations, $P < 0.0001$). For overwater distance and mean dispersal distance, $r_M = 0.902$ (10,000 permutations, $P < 0.0001$).

The leading eigenvector community structure algorithm revealed 25 groups across the Coral Triangle (highlighted by different colors in Fig. 3), with strong consistency among dispersal distance metrics, thresholds, and community detection algorithms (not shown). Nine of these groups contained sampling sites for which we had more than 10 samples.

Distance as a Predictor of Gene Flow and Differentiation.—Using the multiple regression on distance matrices (MRDM) approach to evaluate predictors of genetic structure showed strong positive relationships between both Euclidean and overwater distances and either Φ_{ST} or D_{est} (R_M^2 ranged from 0.46 to 0.68, for the 26 population comparisons), consistent with an isolation by distance pattern. The relationship between F_{ST} and distance was substantively lower ($R_M^2 < 0.03$). Due to the collinearity of Euclidean distance and overwater distance ($R_M^2 = 0.99$ for 26 localities) and overwater distance and dispersal distance ($R_M^2 = 0.90$ for the 20 localities) optimal full models retained only one of these distances. Both forward and backward model selection converged on the same linear model for Φ_{ST} or D_{est}. For Φ_{ST}, the best model contained both Euclidean distance and a barrier to the east of Cenderwasih Bay (Fig. 4: $R^2 = 0.69$, $P < 0.001$), and for D_{est}, the best model contained overwater distance, the barrier to the east of Cenderwasih Bay, and the West Sumatra delineation ($R^2 = 0.50$, $P < 0.001$). For F_{ST}, the best model contained only Euclidean distance but did not explain much variance ($R^2 = 0.03$, $P = 0.014$). With the more restricted set of 20 populations for which we had mean dispersal distance predictions, overwater distance was better predictor of genetic structure than any of the other predictors for Φ_{ST} and D_{est} (Φ_{ST}: $R_{OW}^2 = 0.48$, $R_{Euc}^2 = 0.46$, $R_{BP}^2 = 0.39$, D_{est}: $R_{OW}^2 = 0.30$, $R_{Euc}^2 = 0.30$, $R_{BP}^2 = 0.21$) and F_{ST} was not well predicted regardless of distance metric ($R^2 < 0.12$).

Bayesian Selection of Important Migration Parameters.—Four independent replicate BEAST runs converged to the same likelihood distribution after the removal of approximately 10–30 million burn-in steps from each run. The combined logfiles each contained about 107 million total steps, with high effective sample size (ESS > 200) values further indicating convergence. The analysis highlighted 14

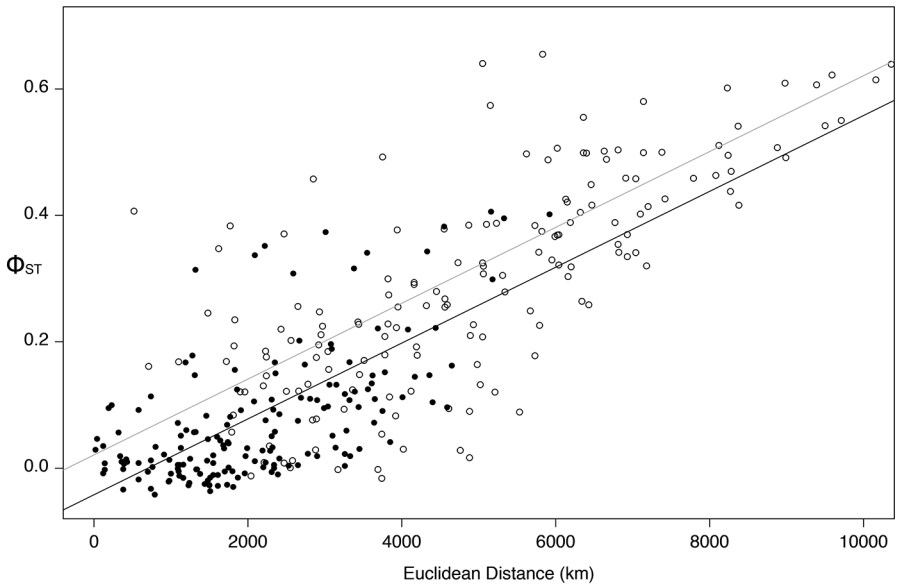

Figure 4. Best linear model of genetic differentiation. Pairwise Φ_{ST} values by Euclidean distance and showing the effect of the division between Indian and Pacific oceans east of Cenderwasih ($R^2 = 0.69$, $P < 0.001$). Grey points represent population pairs including both Indian and Pacific ocean populations whereas black points represent population pairs within either ocean. Gray and black lines represent the regression lines for between and within ocean comparisons.

migration rates as having a Bayes Factor (BF) of >3 (posterior odds of > 3:1). Most of these (11/14) were between adjacent spatial clusters, even though no prior information on location was given in the analysis (Fig. 3). Regions in the central portion of the Coral Triangle had multiple connections among them. All population clusters were connected to the network by at least one significant migration parameter with the exception of Western Australia (here comprising only 7 samples from Ningaloo Reef). The best-supported migration parameter to this region came from the GBR with a BF of 2.54. Randomization of tip locations yielded a migration matrix that was also apparently random, with only five of the 22 well-supported migration parameters occurring between neighboring spatial clusters.

The probability distribution for the location of the most common recent ancestor shows a fairly flat surface across the Coral Triangle, but these probabilities were generally higher than for peripheral populations (Online Fig. 2A). However, randomization of tip locations showed that the prior expectation for the location of the root was highly correlated with sample size (Online Fig. 2A,B; $R^2 = 0.995$).

Discussion

Most phylogeographic studies of Indo-Pacific species to date have relied on estimates of mitochondrial genetic structure to then make inferences about gene flow (e.g., Crandall et al. 2008b, Ackiss et al. 2013, Raynal et al. 2014, and see Keyse et al. 2014 for a review of the geographic and genetic scope of 108 such studies). Because

most such studies assume that patterns of gene flow have been held at a static equilibrium over a long period of time, it is impossible to determine from them whether this gene flow (or lack thereof) is historical or contemporary (but inferences are often made about the latter). However, when we invert our view to consider explicit models of gene flow (IBD and spatial diffusion) rather than genetic structure, we are able to see that the data contain a good deal of spatial information, even in a species with relatively good planktonic dispersal potential and low genetic structure. The fact that spatial distances explain the data better than do oceanographic distances from a biophysical model indicates that historical colonizations and subsequent gene flow events are more strongly reflected in the mitochondrial genome of *L. laevigata* than is contemporary gene flow. These results complement and extend upon earlier observations that genetic data sets from this diverse region are often haunted by the "ghosts of dispersal past" (Benzie 1999).

These novel insights into the population structure and evolution of *L. laevigata* result directly from population sampling that includes localities from the periphery of the Indian and Pacific oceans as well as from the Coral Triangle. The vast species ranges of many Indo-Pacific marine animals, such as that of *L. laevigata,* make comprehensive population genetic and phylogeographic studies difficult. Previous studies of Indo-Pacific taxa have either sampled in the Pacific and Indian oceans with limited sampling in the Coral Triangle (e.g., Lavery et al 1996, Benzie 1999, Williams and Benzie 1998), or extensively within the Coral Triangle without a broader context (e.g., Barber et al 2006, DeBoer et al. 2008, Crandall et al 2008b), albeit with some notable exceptions (e.g., Crandall 2008a, Vogler et al. 2012, 2013). Here, we are able to bring together one of the most spatially comprehensive surveys of genetic variation for any single Indo-Pacific species to date, resulting in the broad geographic context necessary for strong inference. Although our inference is confined to the history of a single locus, the extensive geographic scope of sampled populations combined with state-of-the-field analyses (reviewed in Liggins et al. 2013) as well as results from a biophysical model allow us to understand the phylogeography of this species at greater resolution than ever before.

Signal from the Seascape.—The first line of evidence for a strong geographic signal in this mitochondrial data set is provided by the result from non-metric multidimensional scaling of Φ_{ST}, which shows an extremely strong correlation with longitude ($R^2 = 0.81$, Fig. 5). This NMDS result is reminiscent of classic results from human population genetics (Cavalli-Sforza et al. 1994, Novembre et al. 2008), in which the first two dimensions of variation in the genetic data show a good fit to geography. However, this analysis is based on population genetic distances from a single information-rich locus, rather than PCA scores among individual genotypes (although the latter can be a special case of the former, Wang et al. 2010). Those studies demonstrated that a species (e.g., humans) that has experienced recent expansions and/or high equilibrium gene flow among neighboring populations (and thus has relatively few truly genetically distinct populations) will contain a good deal of spatial information in the two best explanatory dimensions of a multivariate analysis. Although it may be difficult to differentiate between historical and contemporary processes with this non-parametric method, our NMDS approach demonstrates that it is more fruitful to think about Indo-Pacific species in terms of gradational differentiation reflecting historical or contemporary gene flow rather than in terms of

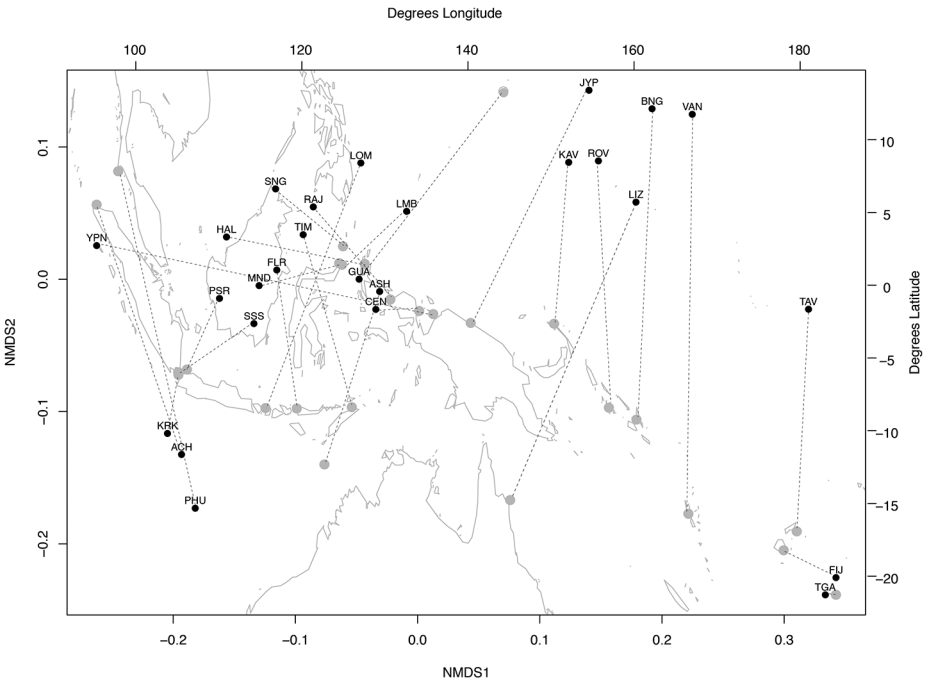

Figure 5. NMDS plots for Φ_{ST} values from populations with sample size > 10. See Table 1 for abbreviations. The first NMDS dimension has a strong correlation with longitude ($R^2 = 0.81$, $P < 1 \times 10^{-9}$), so a map of geography has been underlaid for reference, with the geographical positions of each sample noted.

distinctly structured populations (see also Patterson et al. 2006; and Novembre and Stephens 2008 for in-depth discussion of underlying theory).

A second line of evidence for geographic signal in this data set is found in strong and significant MRDM correlations of population genetic distances (Φ_{ST} or D_{est}) with three different proxies for dispersal distance (Fig. 4). This isolation-by-distance result is consistent with a stepping-stone model of gene flow or colonization among neighboring populations. Interestingly, the two geographic distance proxies were consistently better than the modeled larval dispersal distance for explaining the contemporary pattern of genetic distances in mtDNA. This result appears to run counter to that of White et al. (2010) who found that larval dispersal distance from a biophysical model provided a much stronger explanation than geographic distance for patterns of genetic structure in microsatellites. The stronger correlation of genetic distance to geographic distance as compared to oceanographic distance in our dataset likely arises from a mismatch between the timescale over which mtDNA integrates (1000–100,000 yrs) and the timescale of contemporary larval dispersal among the geographic clusters (approximately 1–100 yrs). The genetic distances calculated from our mitochondrial data set are therefore probably capturing information about historical gene flow and colonization events rather than contemporary gene flow. Thus, mtDNA provides a valuable historical contrast to multi-locus genotyping (notably microsatellites and SNPs), which probably provide a closer fit to contemporary processes (Selkoe and Toonen 2006).

A final line of evidence comes from our phylogeographic diffusion model. By directly reconstructing the mitochondrial history in a spatial context, we are able to consider the role of colonization history and gene flow in the phylogeographic distribution of haplotypes, while simultaneously integrating over the uncertainty in the genealogy. Although we provided no prior information to the diffusion model about the relative spatial locations of each geographic cluster, it selected gene flow parameters that make intuitive sense: gene flow occurs for the most part among clusters that are geographic neighbors (Fig. 3). For example, our model found the Great Barrier Reef to be the most likely source for colonization of Western Australia (although this result did not rise above our threshold for significance), consistent with previous findings in another sea star (Vogler et al. 2013).

By starting with known sample locations and using ancestral state reconstruction to trace these locations back to the root over all possible genealogies, we were also able to arrive at a posterior probability distribution for the location of the MRCA (Online Fig. 2). While it is tempting to interpret from this distribution that the birthplace of the *L. laevigata* mitochondrial matriline was somewhere in the Coral Triangle, it does not depart significantly from a prior expectation based solely on sample size. We therefore can conclude that there is not much information about root location in this data set.

Population Structure of *L. laevigata* in a Broader Geographic Context.—Extensive sampling of populations of *L. laevigata* across the majority of its Indo-Pacific range revealed the presence of pronounced genetic structure, but in ways that are materially different from previous studies. Early studies of *L. laevigata* (Williams and Benzie 1996, 1997, 1998) are frequently cited as classic examples of divergence among populations of Pacific and Indian ocean marine species across the Indo-Malay-Philippine Archipelago. However, the present study shows a more nuanced picture. First, although the haplotype network delineates two large clusters of haplotypes, these are not highly divergent, regionally distinct clades. Enhanced sample size in our study has filled in previously missing haplotypes. The frequency of each cluster follows a longitudinal cline, resulting in a distinctive pattern of isolation-by-distance (Fig. 4). The multiple-regression approach implemented in MRDM considers the effects of putative barriers to dispersal simultaneously with those of isolation by distance (i.e., historical or contemporary stepping-stone gene flow). Out of four possible barriers considered in our MRDM analyses, we found that the putative barrier to the east of Cenderwasih Bay was the only barrier that, when considered together with the effects of geographic distance, provided a consistently good fit to the genetic distances. This result was also captured by our AMOVA analyses, which showed that a simple partition segregating populations to the west (Cenderwasih Bay plus Central Indonesia and Indian Ocean localities) and east (western and central Pacific localities) was a better descriptor of geographic differentiation than were partitions based on lines of disjunction associated with western Sumatra, the Sunda Shelf, the Halmahera current, or any combination of the four (Online Table 2).

Our recovery of only a single potential barrier to gene flow is distinctly different from previous work, which suggested barriers in the Sunda Strait and to the west of Cenderwasih Bay (Crandall et al. 2008b, Kochzius et al. 2009). It reflects the change in perspective provided by a larger study area and explicit consideration of the effects of stepping-stone gene flow (Meirmans 2012). While our other AMOVA partitions

were significant, and our NMDS plot shows some clustering of Indian Ocean, Central Indonesian, Papua + Great Barrier Reef, and South Pacific localities, these distinctions may be more an artifact of sampling design than of any discrete barrier to gene flow.

The proximal explanation for the significance of the genetic disjunction to the east of Cenderwasih Bay is a change in relative frequency of the Indian and Pacific clusters, with the Pacific cluster becoming dominant to the east of this barrier (see fig. 4a in Crandall et al. 2008). In a way then, the disjunction is an artifact of phylogenetic distance between the two major clusters. However, it has been shown that phylogeographic breaks such as this often come to rest in regions of low contemporary migration (Barton and Hewitt 1985). Therefore we suggest that this particular disjunction could ultimately be due to the lack of reef habitat – moving east from Cenderwasih Bay there is almost 700 km of coastline with very sparse and minimal reef habitat (Spalding et al. 2001). Whereas gene flow likely occurs occasionally across Northern Papua, there are few stepping-stone populations to facilitate the exchange of migrants across generations. The stomatopod *Haptosquilla pulchella* Miers, 1880 shows a sharp genetic discontinuity between Cenderwasih Bay and populations in Papua New Guinea, suggesting that this pattern occurs in other Indo-Pacific taxa, but in general this region is surprisingly unknown as few other studies have included samples from Cenderwasih Bay and the western Pacific (Keyse et al. 2014, and see Liu et al. 2012).

Conclusions

The high dispersal capacity and wide range of many shallow reef Indo-Pacific marine organisms necessitate both dense and comprehensive sampling to provide the fullest phylogeographic context for each species. The present mitochondrial data set provides an example of the benefits of a spatially broadened perspective. Previously inferred barriers to gene flow turn out to be artifacts of an isolation-by-distance signature that was invisible at a smaller scale. The fact that this spatial signal is better explained by geographic distance than by oceanographic distances travelled by larvae indicates that the temporal resolution in the current dataset is low: our view of the present is obscured by the ghosts of dispersal past (sensu Benzie 1999). Future phylogeographic studies should strive to broaden their genetic perspective as well as their geographic perspective to increase temporal resolution. Nevertheless, when we consider models of gene flow rather than genetic structure we realize the importance of genetic exchange among neighboring demes as the primary reason for connectivity across a marine species' range (Crandall et al. 2012).

Detailed and broad genetic surveys such as this one usually have been beyond the scope of an individual study, or any single research group. As with most fields of science, increased collaboration and equitable sharing of data and expertise can provide a way forward (see Barber et al. 2014). We hope that the present study, which brings together data from numerous different laboratories and research efforts, will herald even greater collaborative endeavours for this region.

Acknowledgments

Funding for this work was provided by the Australian Research Council (DP0878306, to CR), the US National Science Foundation (OCE-0349177 and DEB-0338566 to PB, and DEB-0508788 to EDC), the World Wildlife Fund (Kathryn Fuller Post-doctoral Research Fellowship, to EAT) and an Explorer's Club Exploration Fund (to LL). Sampling in Timor-Leste was supported by the Coral Triangle Support Partnership and the Ministério da Aquicultura e Pescas, Direcção Nacional de Pescas e Aquicultura (authorised by A Fernandes, L Fontes, J Freitas; guia de marssa: 502/DNPA/VIII/10 and 452/DNPA/VII/11). Export of samples was authorised by the Departmento de Quarentena das Pescas (export permit: 162/FQ006/EXP./DNQB/VII/2011). Sampling in the Solomon Islands was via the Australian Government's Pacific Strategy Assistance Program and with the assistance of the Roviana Conservation Foundation (Solomon Islands Government Ministry of Education & Human Resource Development and Ministry of Fisheries & Marine Resources research permit to S Albert). Sampling in Papua New Guinea was in coordination with the National Research Institute, the Department of Foreign Affairs and Immigration (Research Visa: 10350008304), and the Department of Environment and Conservation (Permit to Export Wildlife: 011318). Sampling in Indonesia was permitted by the Indonesian Institute of Sciences (#1187/SU/KS/2006 and #04239/SU.3/KS/2006). Sampling in Fiji was permitted to EC by the Ministry of Fisheries and Forests, as was export permission (Permit #C300/3006). The Vanuatu Environment Unit provided sampling and export permission to EC in Vanuatu. Authority to sample in Tonga was provided by the Ministry of Agriculture and Food, Forests and Fisheries (by U Fa'anunu to J Drew). Authority to sample at Ashmore Reef was provided by the Department of the Environment, Water, Heritage and the Arts (Access to Biological Resources in a Commonwealth Area for Non-Commercial Purposes permit number: AU-COM2010068) and with logistic support from Australian Customs and Border Control. Sampling in the Coral Sea was supported by the Marine Division of the Department of Sustainability, Environment, Water, Population & Communities (Access to Biological Resources in a Commonwealth Area for Non-Commercial Purposes permit number: AU-COM2008042). Sampling at Ningaloo Reef was under the authority of the Western Australia Department of Environment & Conservation (Licence to take Fauna for Scientific Purposes: SF007126, SF006619; Authority to enter calm land/or waters: CE002227, CE002627). We are grateful to the staff of the Australian Museum Lizard Island Research Station and Heron Island Research Station for their facilities and support (Great Barrier Reef Marine Park Authority and Queensland Parks and Wildlife Marine Parks Permit: G08/28114.1, G09/31678.1, G10/33597.1, G11/34640.1; Queensland Government Department of Primary Industries General Fisheries Permit: 118636, 150981; Australian Quarantine Inspection Service Permit to Import Quarantine Material: IP10017966). Collection of older samples from Thailand, Indonesia, Guam, and New Caledonia used in this study were permitted to J Benzie, and are now in the collections of G Wörheide. J Benzie is gratefully acknowledged for ceding his collection to G Wörheide. We especially thank the crew of the Ashmore Guardian, JD Aguirre, J Aini (Ailan Awareness), S Albert, J Barber-Choi, JF Bertrand, K Davis, J Drew, M Erdmann, A Jackson, M Jimuru, J Keyse, J Kinch (National Fisheries College, Papua New Guinea), E Lewis, W Lovell (Freeflow Dive, Dili), I McLeod, A Mirams, M Muñoz, S Pardede, S Penny, R Pinto (and staff of the Coral Triangle Support Partnership), A Smith (Tiki2 Adventure Tours), T Sinclair-Taylor, C Starger, M Subia, C Taylor, A Turner, P Waldie, C Yusuf, Ambariyanto, Stephen, Lavud, and Takenda for logistical support, field assistance and help in the laboratory. Finally, we thank C Bird, S Williams, J Pearse, and an anonymous reviewer for insightful critiques of a previous draft of this paper.

Literature Cited

Ackiss AS, Pardede S, Crandall ED, Ablan-Lagman M, Ambariyanto, Romena N, Barber PH, Carpenter KE. 2013. Pronounced genetic structure in a highly mobile coral reef fish, *Caesio*

cuning, in the Coral Triangle. Mar Ecol Prog Ser. 480:185–197. http://dx.doi.org/10.3354/meps10199

Alberto F, Raimondi P, Reed DC, Watson JR, Siegel DA, Mitarai S, Coelho N, Serrao EA. 2011. Isolation by oceanographic distance explains genetic structure for *Macrocystis pyrifera* in the Santa Barbara Channel. Mol Ecol. 20:2543–2554. PMid:21535280. http://dx.doi.org/10.1111/j.1365-294X.2011.05117.x

Aljanabi SM, Martinez I. 1997. Universal and rapid salt-extraction of high quality genomic DNA from PCR-based techniques. Nucleic Acids Res. 25:4692–4693. http://dx.doi.org/10.1093/nar/25.22.4692

Balkenhol N, Waits LP, Dezzani RJ. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography. 32:818–830. http://dx.doi.org/10.1111/j.1600-0587.2009.05807.x

Bandelt H-J, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16:37–48. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026036

Barber PH, Cheng SH, Erdmann MV, Tenggardjaja K, Ambariyanto. 2011. Evolution and conservation of marine biodiversity in the Coral Triangle: insights from stomatopod Crustacea. In: Held C, Koenemann S, Schubart C, editors. Phylogeography and population genetics in Crustacea. Vol. 19. Crustacean Issues. CRC Press. p 1–31.

Barber PH, Erdmann MV, Palumbi SR. 2006. Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the Coral Triangle. Evolution. 60:1825–1839. PMid:17089967.

Barber PH, Bellwood DR. 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. Mol Phylogenet Evol. 35:235–253. PMid:15737594. http://dx.doi.org/10.1016/j.ympev.2004.10.004

Barber PH, Palumbi SR, Erdmann MV, Kasim Moosa M. 2000. A marine Wallace's line? Nature. 406:692–693. PMid:10963585. http://dx.doi.org/10.1038/35021135

Barber PH, Palumbi SR, Erdmann MV, Moosa MK. 2002. Sharp genetic breaks among populations of *Haptosquilla pulchella* (Stomatopoda) indicate limits to larval transport: patterns, causes, and consequences. Mol Ecol. 11:659–674. PMid:11972755. http://dx.doi.org/10.1046/j.1365-294X.2002.01468.x

Barton NH, Hewitt G. 1985. Analysis of hybrid zones. Annu Rev Ecol Syst. 16:113–148. http://dx.doi.org/10.1146/annurev.es.16.110185.000553

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. Annu Rev Ecol Evol Syst. 41(1):379–406. http://dx.doi.org/10.1146/annurev-ecolsys-102209-144621

Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc Natl Acad Sci USA. 98(8):4563–4568. PMid:11287657. PMCid:PMC31874. http://dx.doi.org/10.1073/pnas.081068098

Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics. 185(1):313–326. PMid:20176979. PMCid:PMC2870966. http://dx.doi.org/10.1534/genetics.109.112532

Bellwood DR, Hughes T. 2001. Regional-scale assembly rules and biodiversity of coral reefs. Science. 292:1532–1534. PMid:11375488. http://dx.doi.org/10.1126/science.1058635

Benzie J. 1999. Genetic structure of coral reef organisms: ghosts of dispersal past. Am Zool. 39:131–145.

Bird CE, Karl SA, Smouse PE, Toonen RJ. 2011. Detecting and measuring genetic differentiation. *In:* Held C, Koenemann S, Schubart C, editors. Phylogeography and population genetics in Crustacea. Vol. 19. Crustacean Issues. CRC Press. p. 31–73.

Bloomquist EW, Lemey P, Suchard MA. 2010. Three roads diverged? Routes to phylogeographic inference. Trends Ecol Evol. 25(11):626–632. PMid:20863591. PMCid:PMC2956787. http://dx.doi.org/10.1016/j.tree.2010.08.010

Bowen BW, Rocha LA, Toonen RJ, Karl SA. 2013. The origins of tropical marine biodiversity. Trends Ecol Evol. 28(6):359–366. PMid:23453048. http://dx.doi.org/10.1016/j.tree.2013.01.018

Buston PM, Jones GP, Planes S, Thorrold SR. 2012. Probability of successful larval dispersal declines fivefold over 1 km in a coral reef fish. Proc R Soc Lond Ser B-Biol Sci. 279(1735):1883–1888. PMid:22158958. PMCid:PMC3311883. http://dx.doi.org/10.1098/rspb.2011.2041

Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. Nature. 325(6099):31–36. PMid:3025745. http://dx.doi.org/10.1038/325031a0

Carpenter KE, Barber PH, Crandall ED, Ablan-Lagman MCA, Ambariyanto, Mahardika GN, Manjaji-Matsumoto BM, Juinio-Meñez MA, Santos MD, Starger CJ, et al. 2011. Comparative phylogeography of the Coral Triangle and implications for marine management. J Mar Biol. 2011:1–14. http://dx.doi.org/10.1155/2011/396982

Carpenter KE, Springer VG. 2005. The center of the center of marine shore fish biodiversity: the Philippine Islands. Environ Biol Fish. 72:467–480. http://dx.doi.org/10.1007/s10641-004-3154-4

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton, New Jersey, USA: Princeton University Press.

Chappell J, Shackleton NJ. 1986. Oxygen isotopes and sea level. Nature. 324:137–140. http://dx.doi.org/10.1038/324137a0

Chenoweth SF, Hughes JM, Keenan CP, Lavery S. 1998. When oceans meet: a teleost shows secondary intergradation at an Indian-Pacific interface. Proc R Soc Lond Ser B-Biol Sci. 265(1394):415–420. PMCid:PMC1688902. http://dx.doi.org/10.1098/rspb.1998.0310

Cowen RK, Lwiza KM, Sponaugle S, Paris CB, Olson DB. 2000. Connectivity of marine populations: open or closed? Science. 287(5454):857–859. PMid:10657300. http://dx.doi.org/10.1126/science.287.5454.857

Crandall ED, Frey MA, Grosberg RK, Barber PH. 2008a. Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. Mol Ecol. 17(2):611–626. PMid:18179436. http://dx.doi.org/10.1111/j.1365-294X.2007.03600.x

Crandall ED, Jones ME, Muñoz MM, Akinbronbi B, Erdmann MV, Barber PH. 2008b. Comparative phylogeography of two sea stars and their ectosymbionts within the Coral Triangle. Mol Ecol. 17:5276–5290. PMid:19067797. http://dx.doi.org/10.1111/j.1365-294X.2008.03995.x

Crandall ED, Sbrocco EJ, DeBoer TS, Barber PH, Carpenter KE. 2012a. Expansion dating: calibrating molecular clocks in marine species from expansions onto the Sunda Shelf following the last glacial maximum. Mol Biol Evol. 29:707–719. PMid:21926069. http://dx.doi.org/10.1093/molbev/msr227

Crandall ED, Treml EA, Barber PH. 2012b. Coalescent and biophysical models of stepping-stone gene flow in neritid snails. Mol Ecol. 21(22):5579–5598. PMid:23050562. http://dx.doi.org/10.1111/mec.12031

DeBoer TS, Subia MD, Ambariyanto, Erdmann MV, Kovitvongsa K, Barber PH. 2008. Phylogeography and limited genetic connectivity in the endangered boring giant clam across the Coral Triangle. Conserv Biol. 22(5):1255–1266. PMid:18637905. http://dx.doi.org/10.1111/j.1523-1739.2008.00983.x

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88. PMid:16683862. PMCid:PMC1395354. http://dx.doi.org/10.1371/journal.pbio.0040088

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 7:214. PMid:17996036. PMCid:PMC2247476. http://dx.doi.org/10.1186/1471-2148-7-214

Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol. 22:1185–1192. PMid:15703244. http://dx.doi.org/10.1093/molbev/msi103

Excoffier L, Lischer HEL. 2010. Arlequin suite v3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res. 10(3):564–567. PMid:21565059. http://dx.doi.org/10.1111/j.1755-0998.2010.02847.x

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 131:479–491. PMid:1644282. PMCid:PMC1205020.

Faurby S, Barber PH. 2012. Theoretical limits to the correlation between pelagic larval duration and population genetic structure. Mol Ecol. 21:3419–3432. PMid:22574811. http://dx.doi.org/10.1111/j.1365-294X.2012.05609.x

Fievet V, Touzet P, Arnaud J-F, Cuguen J. 2007. Spatial analysis of nuclear and cytoplasmic DNA diversity in wild sea beet (*Beta vulgaris* ssp. *maritima*) populations: do marine currents shape the genetic structure? Mol Ecol. 16:1847–1864. PMid:17444897. http://dx.doi.org/10.1111/j.1365-294X.2006.03208.x

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. Mol Mar Biol Biotechnol. 3(5):294–299. PMid:7881515.

Foster NL, Paris CB, Kool JT, Baums IB, Stevens JR, Sanchez JA, Bastidas C, Agudelo C, Bush P, Day O, Ferrari R, et al. 2012. Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. Mol Ecol. 21:1143–1157. PMid:22276913. http://dx.doi.org/10.1111/j.1365-294X.2012.05455.x

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchiking and background selection. Genetics. 147:915–925. PMid:9335623. PMCid:PMC1208208.

Gaither MR, Bowen BW, Bordenave T-R, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT. 2011. Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the indo-pacific barrier with contemporary overlap in the coral triangle. BMC Evol Biol. 11:189. PMid:21722383. PMCid:PMC3145601. http://dx.doi.org/10.1186/1471-2148-11-189

Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW. 2010. Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). J Biogeogr. 37:133–147. http://dx.doi.org/10.1111/j.1365-2699.2009.02188.x

Galindo HM, Palumbi SR. 2006. Seascape genetics: a coupled oceanographic-genetic model predicts population structure of caribbean corals. Curr Biol. 16:1622–1626. PMid:16920623. http://dx.doi.org/10.1016/j.cub.2006.06.052

Gerlach G, Atema J, Kingsford MJ, Black KP, Miller-Sims V. 2007. Smelling home can prevent dispersal of reef fish larvae. Proc Natl Acad Sci USA. 104(3):858–863. PMid:17213323. PMCid:PMC1783404. http://dx.doi.org/10.1073/pnas.0606777104

Hall R. 2002. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. J Asian Earth Sci. 20:353–431. http://dx.doi.org/10.1016/S1367-9120(01)00069-4

Hedrick PW. 2005. A standardized genetic differentiation measure. Evolution. 59:1633–1638. PMid:16329237.

Hey J, Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc Natl Acad Sci USA. 104(8):2785–2790. PMid:17301231. PMCid:PMC1815259. http://dx.doi.org/10.1073/pnas.0611164104

Jost L. 2008. G_{ST} and its relatives do not measure differentiation. Mol Ecol. 17:4015–4026. PMid:19238703. http://dx.doi.org/10.1111/j.1365-294X.2008.03887.x

Keyse J, Crandall ED, Toonen RJ, Meyer CP, Treml EA, Riginos CR. 2014. The scope of published population genetic data for Indo-Pacific marine fauna, and future research opportunities. Bull Mar Sci. 90:47–78. http://dx.doi.org/10.5343/bms.2012.1107

Kochzius M, Seidel C, Hauschild J, Kirchhoff S, Mester P, Meyer-Wachsmuth I, Nuryanto A, Timm J. 2009. Genetic population structures of the blue starfish Linckia laevigata and its

gastropod ectoparasite *Thyca crystallina*. Mar Ecol Progs Ser. 396:211–219. http://dx.doi.org/10.3354/meps08281

Kool JT, Paris CB, Barber PH, Cowen RK. 2011. Connectivity and the development of population genetic structure in Indo-West Pacific coral reef communities. Global Ecol Biogeogr. 20(5):695–706. http://dx.doi.org/10.1111/j.1466-8238.2010.00637.x

Legendre P, Lapointe F-J, Casgrain P. 1994. Modeling brain evolution from behavior: a permutational regression approach. Evolution. 48:1487–1499. http://dx.doi.org/10.2307/2410243

Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. PLoS Comput Biol. 5:e1000520. PMid:19779555. PMCid:PMC2740835. http://dx.doi.org/10.1371/journal.pcbi.1000520

Lessios H, Robertson D. 2006. Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. Proc R Soc Lond Ser B-Biol Sci. 273:2201–2208. PMid:16901840. PMCid:PMC1635520. http://dx.doi.org/10.1098/rspb.2006.3543

Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, et al. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 319(5866):1100–1104. PMid:18292342. http://dx.doi.org/10.1126/science.1153717

Liggins L, Treml EA, Riginos C. 2013. Taking the plunge: an introduction to undertaking seascape genetic studies and using biophysical models. Geogr Compass. 7(3):173–196. http://dx.doi.org/10.1111/gec3.12031

Lind CE, Evans BS, Taylor JJU, Jerry DR. 2007. Population genetics of a marine bivalve, *Pinctada maxima,* throughout the Indo-Australian Archipelago shows differentiation and decreased diversity at range limits. Mol Ecol. 16(24):5193–5203. PMid:18028303. http://dx.doi.org/10.1111/j.1365-294X.2007.03598.x

Liu S, Dai CF, Allen GR, Erdmann MV. 2012. Phylogeography of the neon damselfish *Pomacentrus coelestis* indicates a cryptic species and different species origins in the west Pacific Ocean. Mar Ecol Prog Ser. 458:155–167. http://dx.doi.org/10.3354/meps09648

Manel S, Holderegger R. 2013. Ten years of landscape genetics. Trends Ecol Evol. 28(10):614–621. http://dx.doi.org/10.1016/j.tree.2013.05.012

McMillan WO, Weigt LA, Palumbi SR. 1999. Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). Evolution. 53:247–260. http://dx.doi.org/10.2307/2640937

Meirmans PG. 2012. The trouble with isolation by distance. Mol Ecol. 21(12):2839–2846. PMid:22574758. http://dx.doi.org/10.1111/j.1365-294X.2012.05578.x

Meirmans PG, Van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes. 4:792–794. http://dx.doi.org/10.1111/j.1471-8286.2004.00770.x

Mirams AGK, Treml EA, Shields JL, Liggins L, Riginos C. 2011. Vicariance and dispersal across an intermittent barrier: population genetic structure of marine animals across the Torres Strait land bridge. Coral Reefs. 30:937–949. http://dx.doi.org/10.1007/s00338-011-0767-x

Newman MEJ. 2006. Finding community structure in networks using the eigenvectors of matrices. Phys Rev E. 74:036104. PMid:17025705. http://dx.doi.org/10.1103/PhysRevE.74.036104

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, et al. 2008. Genes mirror geography within Europe. Nature. 456(7218):98–101. PMid:18758442. PMCid:PMC2735096. http://dx.doi.org/10.1038/nature07331

Novembre J, Stephens M. 2008. Interpreting principal component analyses of spatial population genetic variation. Nature Genet. 40(5):646–649. PMid:18425127. http://dx.doi.org/10.1038/ng.139

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2012. vegan*:* community ecology package, R package version 2.0-5.

Oppenheimer S. 2012. Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. Philos Trans R Soc Lond Ser B-Biol Sci. 367(1590):770–784. PMid:22312044. PMCid:PMC3267120. http://dx.doi.org/10.1098/rstb.2011.0306

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. PLoS Genet. 2(12):e190. PMid:17194218. PMCid:PMC1713260. http://dx.doi.org/10.1371/journal.pgen.0020190

Paulay G. 1990. Effects of late Cenozoic sea-level fluctuations on the bivalve faunas of tropical oceanic islands. Paleobiology. 16:415–434.

Polzin T, Daneshmand SV. 2003. On Steiner trees and minimum spanning trees in hypergraphs. Operations Res Lett. 31:12–20. http://dx.doi.org/10.1016/S0167-6377(02)00185-2

Renema W, Bellwood DR, Braga JC, Bromfield K, Hall R, Johnson K, Lunt P, Meyer C, McMonagle L, Morley R. 2008. Hopping hotspots: global shifts in marine biodiversity. Science. 321(5889):654. PMid:18669854. http://dx.doi.org/10.1126/science.1155674

Riginos C, Liggins L. 2013. Seascape genetics: populations, individuals, and genes marooned and adrift. Geogr Compass. 7(3):197–216. http://dx.doi.org/10.1111/gec3.12032

Roberts CM, McClean CJ, Veron JEN, Hawkins JP, Allen GR, McAllister DE, Mittermeier CG, Schueler FW, Spalding M, Wells F, et al. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. Science. 295:1280–1284. PMid:11847338. http://dx.doi.org/10.1126/science.1067728

Rocha LA, Craig MT, Bowen BW. 2007. Phylogeography and the conservation of coral reef fishes. Coral Reefs. 26(3):501–512. http://dx.doi.org/10.1007/s00338-007-0261-7

Saenz-Agudelo P, Jones GP, Thorrold SR, Planes S. 2012. Patterns and persistence of larval retention and connectivity in a marine fish metapopulation. Mol Ecol. 21:4695–4705. http://dx.doi.org/10.1111/j.1365-294X.2012.05726.x

Selkoe KA, Henzler CM, Gaines SD. 2008. Seascape genetics and the spatial ecology of marine populations. Fish Fish. 9:363–377. http://dx.doi.org/10.1111/j.1467-2979.2008.00300.x

Spalding M, Ravilious C, Green EP. 2001. World atlas of coral reefs. Berkeley: University of California Press.

Spalding MD, Fox HE, Halpern BS, McManus MA, Molnar J, Allen GR, Davidson N, Jorge ZA, Lombana AL, Lourie SA, et al. 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. Bioscience. 57(7):573–583. http://dx.doi.org/10.1641/B570707

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123:585–595. PMid:2513255. PMCid:PMC1203831.

Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, Vanden Berghe E, Worm B. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature. 466(7310):1098–1101. http://dx.doi.org/10.1038/nature09329

Toonen RJ, Andrews KR, Baums IB, Bird CE, Concepcion GT, Daly-Engel TS, Eble JA, Faucci A, Gaither MR, Iacchei M, et al. 2011. Defining boundaries for ecosystem-based management: a multispecies case study of marine connectivity across the Hawaiian Archipelago. J Mar Biol. 2011:1–13. http://dx.doi.org/10.1155/2011/460173

Treml EA, Halpin PN. 2012. Marine population connectivity identifies ecological neighbors for conservation planning in the Coral Triangle. Conserv Lett. 5:441–449. http://dx.doi.org/10.1111/j.1755-263X.2012.00260.x

Treml EA, Roberts JJ, Chao Y, Halpin PN, Possingham HP, Riginos C. 2012. Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. Integrative Comp Biol. 52:525–537. PMid:22821585. http://dx.doi.org/10.1093/icb/ics101

Vogler C, Benzie J, Barber PH, Erdmann MV, Ambariyanto, Sheppard C, Tenggardjaja K, Gérard K, Wörheide G. 2012. Phylogeography of the crown-of-thorns starfish in the Indian Ocean. PLoS ONE. 7:e43499. PMid:22927975. PMCid:PMC3424128. http://dx.doi.org/10.1371/journal.pone.0043499

Vogler C, Benzie J, Tenggardjaja K, Barber PH, Wörheide G. 2013. Phylogeography of the crown-of-thorns starfish: genetic structure within the Pacific species. Coral Reefs. 32:515–525. http://dx.doi.org/10.1007/s00338-012-1003-z

Voris HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J Biogeogr. 27:1153–1167. http://dx.doi.org/10.1046/j.1365-2699.2000.00489.x

Wainscoat JS, Hill AV, Boyce AL, Flint J, Hernandez M, Thein SL, Old JM, Lynch JR, Falusi AG, Weatherall DJ. 1986. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. Nature. 319(6053):491–493. PMid:3003580. http://dx.doi.org/10.1038/319491a0

Wang C, Szpiech ZA, Degnan JH, Jakobsson M, Pemberton TJ, Hardy JA, Singleton AB, Rosenberg NA. 2010. Comparing spatial maps of human population-genetic variation using Procrustes analysis. Stat Appl Genet Mol Biol. 9:Article 13. http://dx.doi.org/10.2202/1544-6115.1493

Wang P, Song YT, Chao Y, Zhang HC. 2005. Parallel computation of the regional ocean modeling system. Int J High Performance Comput Appl. 19:375–385. http://dx.doi.org/10.1177/1094342005059115

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 7:256–276. http://dx.doi.org/10.1016/0040-5809(75)90020-9

Wheeler B. 2010. *lmPerm:* permutation tests for linear models.

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ. 2010. Ocean currents help explain population genetic structure. Proc R Soc Lond Ser B-Biol Sci. 277(1688):1685–1694. PMid:20133354. PMCid:PMC2871860. http://dx.doi.org/10.1098/rspb.2009.2214

Whitlock M, McCauley D. 1999. Indirect measures of gene flow and migration: F_{ST} does not equal 1/(4Nm+1). Heredity. 82:117–125. http://dx.doi.org/10.1038/sj.hdy.6884960

Whitlock MC. 2011. G'_{ST} and D do not replace F_{ST}. Mol Ecol. 20(6):1083–1091. PMid:21375616. http://dx.doi.org/10.1111/j.1365-294X.2010.04996.x

Williams ST. 2000. Species boundaries in the starfish genus *Linckia*. Mar Biol. 136:137–148. http://dx.doi.org/10.1007/s002270050016

Williams ST, Benzie JAH. 1993. Genetic consequences of long larval life in the starfish, *Linckia laevigata* (Echinodermata; Asteroidea), on the Great Barrier Reef, Australia. Mar Biol. 117:71–77. http://dx.doi.org/10.1007/BF00346427

Williams ST, Benzie JAH. 1996. Genetic uniformity of widely separated populations of the coral reef starfish *Linckia laevigata* from the east Indian and west Pacific oceans, revealed by allozyme electrophoresis. Mar Biol. 126:99–107. http://dx.doi.org/10.1007/BF00571381

Williams ST, Benzie JAH. 1997. Indo-West Pacific patterns of genetic differentiation in the high-dispersal starfish *Linckia laevigata*. Mol Ecol. 6:559–573. http://dx.doi.org/10.1046/j.1365-294X.1997.00221.x

Williams ST, Benzie JAH. 1998. Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA, and allozyme data. Evolution. 52:87–99. http://dx.doi.org/10.2307/2410923

Williams ST, Duda TF. 2008. Did tectonic activity stimulate oligo-miocene speciation in the Indo-West Pacific? Evolution. 62(7):1618–1634. PMid:18410535. http://dx.doi.org/10.1111/j.1558-5646.2008.00399.x

Williams ST, Jara J, Gomez E, Knowlton N. 2002. The marine Indo-West Pacific break: contrasting the resolving power of mitochondrial and nuclear genes. Integrative Comp Biol. 42:941–952. PMid:21680374. http://dx.doi.org/10.1093/icb/42.5.941

Wright S. 1950. The genetical structure of populations. Ann Eugenics. 15(1):323–354. http://dx.doi.org/10.1111/j.1469-1809.1949.tb02451.x

Yamaguchi M. 1973. Early life history of coral reef asteroids, with special reference to *Acanthaster planci* (L.). *In:* Jones OA, Endean R, editors. Biol Geol Coral Reefs. New York and London: Academic Press, p. 369–389. PMid:4584395. http://dx.doi.org/10.1016/B978-0-12-395526-5.50019-5