Coastal pollution limits pelagic larval dispersal

Jonathan B. Puritz[1,2] & Robert J. Toonen[1]

The ecological impact of large coastal human populations on marine ecosystems remains relatively unknown. Here, we examine the population structure of *Patiria miniata*, the bat star, and correlate genetic distances with a model based on flow rates and proximity to *P. miniata* populations for the four major stormwater runoff and wastewater effluent sources of the Southern California Bight. We show that overall genetic connectivity is high ($F_{ST} \sim 0.005$); however, multivariate analyses show that genetic structure is highly correlated with anthropogenic inputs. The best models included both stormwater and wastewater variables and explained between 26.55 and 93.69% of the observed structure. Additionally, regressions between allelic richness and distance to sources show that populations near anthropogenic pollution have reduced genetic diversity. Our results indicate that anthropogenic runoff and effluent are acting as barriers to larval dispersal, effectively isolating a high gene flow species that is virtually free of direct human impact.

[1] Hawai'i Institute of Marine Biology, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, P.O. Box 1346, Kane'ohe, Hawaii 96744, USA. [2] Department of Zoology, University of Hawai'i at Mānoa, 2538 McCarthy Mall, Honolulu, Hawaii 96822, USA. Correspondence and requests for materials should be addressed to J.B.P. (email: jpuritz@gmail.com).

1

Human beings have impacted every marine ecosystem on earth, and almost half of these systems have been strongly affected by multiple anthropogenic factors[1]. The ecological consequences of anthropogenic impacts have become too large to ignore, and recent efforts have been made to incorporate them into management and conservation strategies[2,3]. However, in addition to recognizing humans as the world's most powerful evolutionary force[4], we have little information on anthropogenic impacts to the evolutionary processes of non-harvested species, such as population connectivity, and this specific information remains a critical missing element in almost all marine conservation[5,6]. Over 17 million people, close to 25% of the entire US coastal population[7,8], inhabit the coastal area of the Southern California Bight (SCB), and it represents one of the most affected marine ecosystems in the world[1]. Besides habitat loss and modification, the next largest environmental impact of this urban population is the discharge of contaminants into the ecosystem[7]. The SCB has more than 60 point sources discharging ~4.7 billion liters of effluent per day, with the two largest sources of contaminants being sewage effluent from municipal treatment plants (wastewater) and runoff from highly modified natural river basins (stormwater)[9]. This effluent transports large loads of known and unknown contaminants, including heavy metals, chlorinated hydrocarbons, petroleum hydrocarbons, nutrients and bacteria[7], and has been shown to be toxic to marine fauna[10]. A wide range of anthropogenic chemicals has been shown to disrupt the hormonal systems of wildlife species and have become a major cause for concern in global efforts to conserve biodiversity[11]. Many of these substances are also known to be toxic to larvae[12–14]. We hypothesized that anthropogenic effluent and runoff and the transported contaminants associated with them are toxic to larvae, therefore affecting marine population connectivity in the SCB by inducing broad-scale larval mortality.

To test whether population connectivity was limited by stormwater and wastewater contaminants, we examined the population structure of *Patiria miniata*, the common bat star, across 16 populations in the SCB. *P. miniata* is a ubiquitous omnivore found along the subtidal and intertidal west coast of North America. It has a 6–10-week pelagic larval development period, a potential year-round breeding season that may peak in summer[15,16], and has been shown to have both high genetic diversity and high genetic connectivity between populations over 1,000-km apart[17]. Moreover, *P. miniata* is a non-fished, non-harvested species relatively free from any direct human impact, making it a conservative system to test for human-mediated effects on connectivity. We found stormwater and wastewater to be effective barriers to larval dispersal, reducing gene flow between populations. Furthermore, wastewater (and most likely stormwater) sources effectively lower genetic diversity. Thus, our study suggests a role for human impact in altering larval dispersal.

Results

Population connectivity. We sequenced 439 individuals at two mitochondrial DNA (mtDNA) genes and genotyped them at seven microsatellite loci (nDNA). As expected, genetic diversity is extremely high (Table 1, Supplementary Fig. S1) and the overall genetic structure of *P. miniata* is low, with global F_{ST} values ranging between 0.001 and 0.007 and only one of those values being statistically significant. However, examining pairwise differentiation values among sites showed that there is a distinct geographic pattern to the genetic structure of *P. miniata* (Fig. 1), and the structure appeared to correlate visually with the major sources of anthropogenic contaminants in the SCB (Fig. 1).

Anthropogenic effects on genetic structure. To specifically test the effects of stormwater and wastewater on genetic structure, we created a simple model based on the average flow rate of the contaminant source and the distance of that source from our sampled populations. That model was then used to create pairwise 'resistance' matrices for all of our populations. We examined how well anthropogenic resistance matrices predicted the pairwise values of genetic structure, connectivity (F_{ST}) and differentiation (D_{EST}), using single and multiple matrix regression (MMR) frameworks[18] and a distance-based canonical redundancy analysis.

Single-variable regression. Examining pollution sources in a single-variable framework revealed consistent and significant limitations to gene flow associated with contaminant sources. Overall, several stormwater and wastewater variables have high and significant correlations to several measures of genetic structure using the low diversity (LD) mtDNA locus (Table 2, Supplementary Table S2). In contrast, no significant predictors were found for the nDNA data sets. Neither geographic distance nor the stormwater input from the Santa Clara River (RSC) was found to be a significant predictor for any data set.

Multivariate regression. Expanding models to include more than one predictive variable greatly improved the model fit for all data sets, highlighting the synergistic effect of multiple anthropogenic sources on the genetic structure of *P. miniata*. Every top model included both stormwater and wastewater variables (Table 3, Supplementary Tables S3–S8), indicating that both are important in driving the observed patterns of genetic structure. Although there were significant models predicting nDNA genetic structure, they often had relatively low R^2 values (0.11–0.14) compared with that using mtDNA loci (0.22–0.57). Models predicting genetic structure values using the LD mtDNA marker showed the highest R^2 values (0.54–0.58) and the best fit (Fig. 2).

Among stormwater sources, three were present in every top mtDNA model: the Los Angeles River, the San Gabriel River and the Santa Ana River. RSC was included in several models, but was usually nonsignificant. Among all loci, the Hyperion and Palos Verdes sewage outfalls were included in all top models, whereas the Point Loma outfall appears in two of the six models. The Huntington Beach outfall is not included in many models and is often insignificant when it does. This is most likely because of its proximity and spatial autocorrelation with Santa Ana River.

Canonical redundancy analysis. MMR is one of the better performing methods of evaluating the influence of landscape (seascape) variables

---

Table 1 | Genetic diversity of *Patiria miniata*.

	LD mtDNA	Full mtDNA		nDNA
Number of haplotypes	33	294	Average number of alleles per locus	13
Number of singletons	24	238	Number of unique genotypes	437
Mean gene diversity	0.4169	0.9914	Mean gene diversity	0.6013
Mean allelic richness	2.47	13.17	Mean allelic richness	5.9

LD, low diversity; mtDNA, mitochondrial DNA.
Diversity statistics for all three loci. Gene diversity and allelic richness are averaged across populations, and across loci for the nDNA markers.
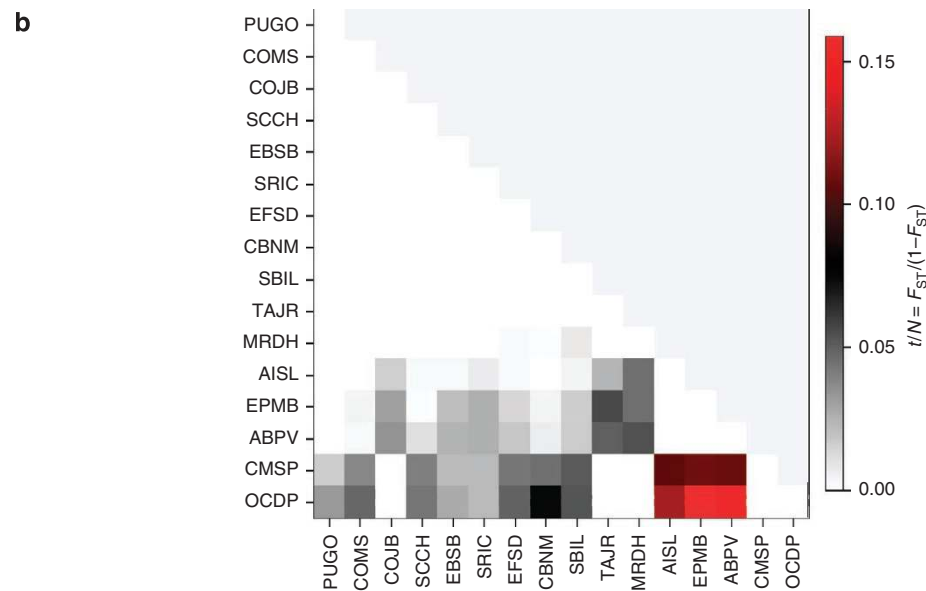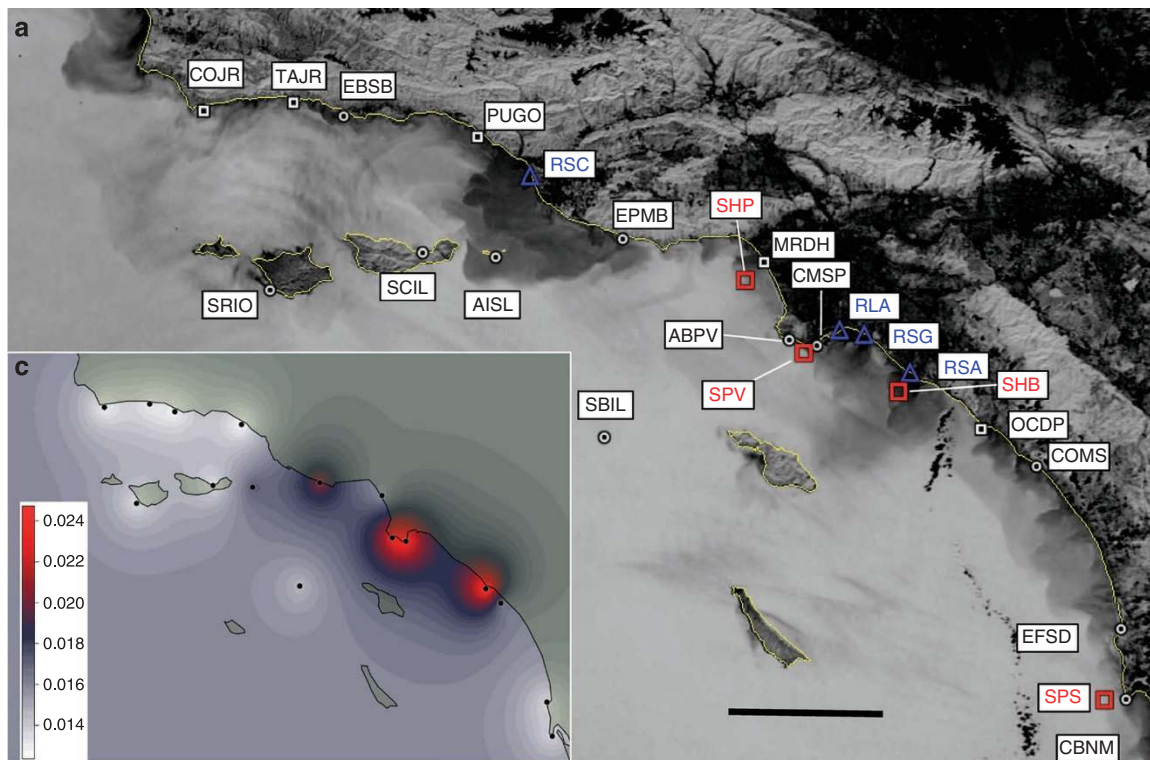
Figure 1 | Visualization of anthropogenic correlations of *P. miniata* population structure. (**a**) Map and satellite image of *P. miniata* population sampling sites highlighting the four wastewater sources (red squares) and the four stormwater sources (blue triangles). The image is a reverse black and white rendering of an image from the NASA MODIS satellite on 1 December 2005 (http://rapidfire.sci.gsfc.nasa.gov/subsets/?subset=AERONET_La_Jolla.2005012.aqua.250m) showing visible sediment from stormwater. White squares represent populations with $n < 20$, whereas white circles are populations with $n > 20$. Each population is also labelled with a four-letter code for reference. Red squares are the locations of wastewater outfalls and blue triangles are the locations of stormwater sources (river mouths). Pollution source abbreviations are RLA, The Los Angeles River; RSA, The Santa Ana River; RSG, The San Gabriel River; SHB, outfall for the Huntington Beach wastewater treatment facility; SHP, Hyperion sewage treatment plant; SPS, outfall for the Point Loma wastewater treatment plant and SPV, outfall for the Rancho Palos Verdes wastewater treatment plant. (**b**) A heat map colour-coded by pairwise linearized F_{ST} values from the LD mtDNA marker. (**c**) A Kriging surface interpolation using the average pairwise F_{ST} value for each population averaged across each of the three marker sets. Scale bar, 50 km.

on genetic structure; however, it has recently been suggested that the best analytical strategies involve using multiple types of analyses[19]. To this end, we used a distance-based canonical redundancy analysis, that has been shown to have not only relatively low power for detect-ing effects but also low type-1 error rate[19], to redo all single variable regressions (Supplementary Table S9) and to test the six top models from the MMR analysis (Supplementary Table S10). The results were nearly identical to the MMR and corroborate our initial findings.

Table 2 | Single anthropogenic source regression results.

LD mtDNA Full data set	R^2	F_{ST} N>20 subset	R^2	nDNA Full data set	R^2	F_{ST} N>20 subset	R^2
RSG	0.4051***	RLA	0.6376***	SPV	0.0307	SPS	0.2915***
SPV	0.3785***	RSG	0.6288***	RLA	0.0235	DIST	0.0739*
RLA	0.3625***	SHB	0.6015***	DIST	0.0207	RSC	0.0459
SHB	0.3344***	RSA	0.5856***	RSG	0.0174	SPV	0.023
RSA	0.3334***	SPV	0.4004***	SHP	0.0157	SHB	0.0158
SHP	0.2629***	SHP	0.1844**	RSC	0.0121	RLA	0.0138
RSC	0.0137	RSC	0.0935*	SPS	0.0032	RSA	0.0134
SPS	0.0116	DIST	0.0759	SHB	0.0027	RSG	0.0112
DIST	0.0031	SPS	0.0504	RSA	0.0019	SHP	0.0019

LD mtDNA Full data set	R^2	D_{EST} N>20 subset	R^2	nDNA Full data set	R^2	D_{EST} N>20 subset	R^2
RSG	0.3653***	RSG	0.8710***	RSC	0.0273	RSC	0.1579**
RLA	0.3545***	RLA	0.8166***	DIST	0.0197	SPS	0.1199**
SPV	0.2922***	RSA	0.8138***	SPV	0.0121	DIST	0.0156
SHB	0.2523***	SHB	0.7661***	SHP	0.0078	SPV	0.0092
RSA	0.2450***	SPV	0.2689***	SPS	0.0071	SHP	0.0032
SHP	0.1802***	DIST	0.0337	RLA	0.0058	RLA	0.0027
SPS	0.0850**	RSC	0.0319	SHB	0.0040	SHB	0.0021
RSC	0.0306	SHP	0.0177	RSG	0.0030	RSG	0.0017
DIST	0.0000	SPS	0.0039	RSA	0.0008	RSA	0.0011

DIST, geographic distance; LD, low diversity; mtDNA, mitochondrial DNA; RLA, The Los Angeles River; RSA, The Santa Ana River; RSC, The Santa Clara River; RSG, The San Gabriel River; SHB, outfall for the Huntington Beach wastewater treatment facility; SHP, Hyperion sewage treatment plant; SPS, outfall for the Point Loma wastewater treatment plant; SPV, outfall for the Rancho Palos Verdes wastewater treatment plant.

R^2 values for single-variable matrix regression tests using the LD mtDNA and nDNA F_{ST} and D_{EST} values. Each measure of genetic structure was tested with the full data set and the $n>20$ data subset. *$P<0.05$, **$P<0.01$, and ***$P<0.001$.

**Table 3 | Multiple anthropogenic source regression results.**

F_{ST} Marker	Full data set Model	R^2	ΔAICc	D_{EST} Marker	Model	R^2	ΔAICc
LD mtDNA	*RSA*, *RLA*, *RSG*, RSC, *SHP*, *SPV*	0.5768	0.00	LD mtDNA	***RSA***, ***RLA***, ***RSG***, ***SHP***, ***SPV***, ***SPS***	**0.5406**	**0.00**
	RSA, ***RLA***, ***RSG***, ***SHP***, ***SPV***	**0.5669**	**0.48**		*RSA*, *RLA*, *RSG*, *SHP*, *SPV*, SHB, *SPS*	0.5445	1.32
	RLA, *RSG*, RSC, *SHP*, *SPV*, *SHB*	0.5743	0.70		*RSA*, *RLA*, *RSG*, *SHP*, *SPV*, *SPS*, DIST	0.5432	1.66
Full mtDNA	***RSA***, ***RLA***, ***RSG***, ***SHP***, ***SPV***, ***SPS***	**0.3259**	**0.00**	Full mtDNA	***RSA***, ***RLA***, ***RSG***, ***SHP***, ***SPV***	**0.2655**	**0.00**
	RSA, *RLA*, *RSG*, RSC, *SHP*, *SPV*, *SPS*	0.3365	0.45		DIST, *RSA*, *RLA*, *RSG*, *SHP*, *SPV*	0.2741	0.88
	RSA, *RLA*, *SPV*, *SHB*, *SPS*	0.3080	0.86		*RSA*, *RLA*, *RSG*, *SPV*, *SHB*	0.2532	1.99
nDNA	***RLA***, ***RSG***, ***RSC***, ***SHP***, ***SPV***	**0.1075**	**0.00**	nDNA	*RSA*, *RSC*, *SHP*, *SPV*	0.1181	0.00
	RSA, *RSC*, *SHP*, *SPV*	0.1233	0.08		***RLA***, ***RSG***, ***RSC***, ***SHP***, ***SPV***	**0.1292**	**0.74**
	RLA, *RSG*, *RSC*, *SHP*, *SPV*	0.1382	0.28		RSC, *SHP*, *SPV*, *SHB*	0.1116	0.88

F_{ST} Marker	N>20 subset Model	R^2	ΔAICc	D_{EST} Marker	Model	R^2	ΔAICc
LD mtDNA	***RLA***, ***RSC***, ***SHP***, ***SPV***	**0.7621**	**0.00**	LD mtDNA	***DIST***, ***RLA***, ***RSG***, ***SHP***, ***SPS***	**0.9369**	**0.00**
	RLA, *RSC*, *SHP*, *SPV*, *SPS*	0.7705	0.11		*DIST*, *RSG*, *SHP*, *SPV*, *SPS*	0.9368	0.12
	RSG, *RSC*, *SHP*, *SPV*	0.7616	0.13		*DIST*, *RLA*, *SHP*, *SPV*, *SPS*	0.9367	0.20
Full mtDNA	***RSA***, ***RLA***, ***RSG***, ***SHP***, ***SPS***	**0.6802**	**0.00**	Full mtDNA	***RLA***, ***RSC***, ***SHP***, ***SPV***, ***SPS***	**0.4288**	**0.00**
	RLA, *RSG*, *SHP*, *SHB*, *SPS*	0.6802	0.00		*RSG*, *RSC*, *SHP*, *SPV*, *SPS*	0.4287	0.01
	RSA, *RSG*, *SHP*, *SHB*, *SPS*	0.6802	0.00		*RLA*, *RSG*, *RSC*, *SHP*, *SPS*	0.4275	0.14
nDNA	***RLA***, ***RSG***, ***SHP***, ***SPS***	**0.4447**	**0.00**	nDNA	***RSC***, ***SPS***	**0.2365**	**0.00**
	SHP, *SPV*, *SHB*, *SPS*	0.4441	0.08		*RSC*, *SPV*, *SPS*	0.2525	0.95
	RSA, *SHP*, *SPV*, *SPS*	0.4421	0.32		*RLA*, *RSC*, *SPS*	0.2425	1.82

DIST, geographic distance; LD, low diversity; mtDNA, mitochondrial DNA; RLA, The Los Angeles River; RSA, The Santa Ana River; RSC, The Santa Clara River; RSG, The San Gabriel River; SHB, outfall for the Huntington Beach wastewater treatment facility; SHP, Hyperion sewage treatment plant; SPS, outfall for the Point Loma wastewater treatment plant; SPV, outfall for the Rancho Palos Verdes wastewater treatment plant.

The R^2 and AICc values for the top three multivariate models for each marker and for each measure of genetic structure. Full mtDNA refers to using both loci in a concatenated analysis (see Methods). Models were ranked by AICc score. Each variable was italicized if its regression coefficient was significantly different than zero with $P<0.05$. The best model, in bold, was chosen as the model with the lowest AICc score that had all significant variable coefficients. All of the models were significant at $P<0.001$, with the exception of the models using nDNA, which were all significant at $P<0.05$.
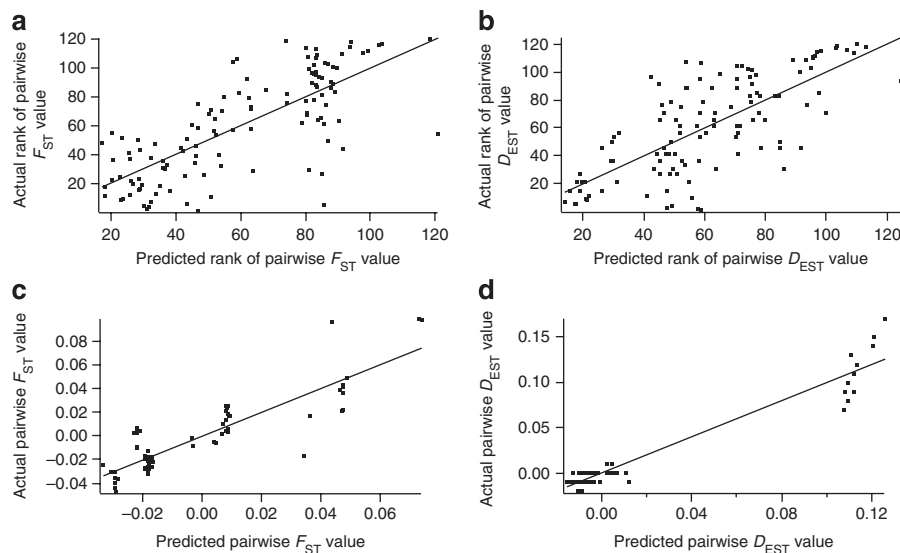
Figure 2 | Graphs of the best multivariate models using the LD mtDNA loci. (**a**) Predicted rank F_{ST} versus actual F_{ST} rank using the full population data set. $R^2 = 0.5768$, P-value < 0.0001. (**b**) Predicted rank D_{EST} versus actual D_{EST} rank, using the full population data set. $R^2 = 0.5406$, P-value < 0.0001. (**c**) Predicted F_{ST} versus actual F_{ST}, using the $n > 20$ data subset. $R^2 = 0.7621$, P-value < 0.0001. (**d**) Predicted D_{EST} versus actual D_{EST} using the $n > 20$ data subset. $R^2 = 0.9369$, P-value < 0.0001. DIST, geographic distance.
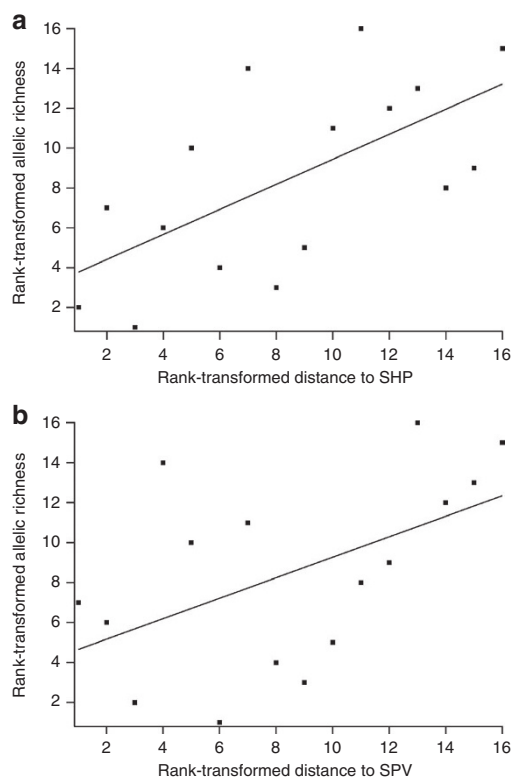
Figure 3 | Graphs of the two significant predictors of nDNA AR. (**a**) $R^2 = 0.3962$, P-value = 0.0090 and (**b**) $R^2 = 0.2619$, P-value = 0.0427.

Population sampling size and high-diversity genetic markers. Despite its ubiquitous nature, *P. miniata* was rare at some sites, and despite extensive survey, some populations only had 15 individuals sampled. For genetic markers with high mutation rates and especially for species with low levels of genetic structure, under-sampling of populations may lead to inaccurate measures of genetic structure[20,21]. To control for this possibility, a subset of the data was

reanalyzed using only populations with more than 20 individuals sampled. This data subset represents a balance between increasing the accuracy of our genetic measures while still retaining 11 of the 16 original populations and maintaining sufficient geographic sampling (Fig. 1, Supplementary Table S1).

For all stormwater and wastewater models, R^2 values increased using the subset of well-sampled populations (Tables 2 and 3, and Fig. 2). In single variable tests, almost all significant variables from the full data set remained significant (Table 2, Supplementary Table S2). For multivariate models, the top models had increased R^2 values and included fewer variables than top models using the full data set; however, they still included both wastewater and stormwater variables (Table 3). For multivariate models, removing populations, especially those bordering or between anthropogenic inputs, lessened geographic sampling, leading to higher correlations between predictive matrices and therefore less variables in the top models. Using the data subset also decreased the large discrepancy between R^2 values of multivariate models for the two mtDNA loci, demonstrating that under-sampling had a large effect on the genetic structure values from the higher polymorphism marker. In all cases, R^2 values were greater for models using either of the mtDNA markers compared with the nDNA markers.

Loss of genetic diversity due to anthropogenic isolation. To examine the effects of wastewater and stormwater sources on genetic diversity, rarefied nDNA allelic richness (AR) for each population was regressed to the distance of each population from a single effluent source. There were positive correlations between AR and distance away from every wastewater and stormwater source; however, only two sources had significant relationships ($P < 0.05$), wastewater sources Hyperion and Palos Verdes (Fig. 3).

Contemporary versus historical effects. Examining the relative roles of contemporary and historical seascape effects on genetic connectivity is essential to any seascape genetic study[22]. The stormwater sources examined in this study are also the mouths of natural rivers. Without accurate estimates of historical river outputs, there is no way of separating contemporary anthropogenic influences of increased urban runoff, more consistent outflow and extensive

Table 4 | Stormwater residual regression results.

F_{ST} Marker	Full data set Model	R^2	P-value	D_{EST} Marker	Model	R^2	P-value
LD mtDNA	SHP, SPV, SHB	0.1217	0.0012	LD mtDNA	SHP, SPV, SPS	0.1290	0.0011
Full mtDNA	SHP, SPV, SPS	0.0771	0.0249	Full mtDNA	SPS	0.0042	0.4815
nDNA	SHP	0.0022	0.6051	nDNA	SHP, SPV, SHB	0.0394	0.1964

F_{ST} Marker	N > 20 subset Model	R^2	P-value	D_{EST} Marker	Model	R^2	P-value
LD mtDNA	SHP, SPV, SHB	0.2104	0.0020	LD mtDNA	SHP, SPV, SHB, SPS	0.1813	0.0146
Full mtDNA	SHP, SPV, SHB, SPS	0.4003	0.0001	Full mtDNA	SPS	0.0849	0.0176
nDNA	SHP, SPV, SHB, SPS	0.4048	0.0001	nDNA	SPS	0.0905	0.0141

LD, low diversity; mtDNA, mitochondrial DNA; SHB, outfall for the Huntington Beach wastewater treatment facility; SHP, Hyperion sewage treatment plant; SPS, outfall for the Point Loma wastewater treatment plant; SPV, outfall for the Rancho Palos Verdes wastewater treatment plant.
The R^2 and P-values values for the top multivariate models for each marker and for each measure of genetic structure correlating wastewater variables to the residuals from top stormwater models. Each variable was italicized if its regression coefficient was significantly different than zero, with $P < 0.05$.

riverbed engineering[7], from the potential historical effects of natural river flow in these variables. Therefore, a conservative approach was taken and all stormwater variables were treated solely as historic variables, and wastewater variables were treated as the only contemporary anthropogenic variables and examined in an MMR framework.

For each measure of population structure, an MMR model using only historic (stormwater) variables was selected by AICc (Akaike information criterion adjusted for small sample sizes) value. The residuals, or unexplained variance, of that model were then used as a response variable in an MMR model using only contemporary (wastewater) variables. This specifically examines whether contemporary variables can explain significant portions of population structure that are unexplained by historic variables. Anthropogenic wastewater variables explained significant amounts of residual variation in three of the six models tested with the full data set and all six models tested with the $n > 20$ subset (Table 4), conclusively showing that contemporary variables have considerable influence on the population connectivity of *P. miniata*.

Discussion

The high and significant correlations of anthropogenic environmental variables with the population structure of *P. miniata* demonstrate the utility of using a spatially explicit model to examine influences on population connectivity. Single variable and multivariate modelling show that human beings have a profound influence on the population connectivity of a ubiquitous, non-harvested species with high levels of genetic diversity and gene flow. This anthropogenic isolation is most likely caused by larval mortality from the substantial contemporary augmentation of freshwater, particulate and/or pollutant load into the SCB.

Most seascape genetic studies have focused on the role of oceanographic factors in determining genetic structure[23–28]. The SCB, however, is an oceanographically unique system with heterogeneous and variable mesoscale eddies influencing short-term circulation and the Southern California Cyclonic Gyre dominating longer term circulation patterns[29]. Ocean-based circulation models for connectivity in the SCB have shown rather stable patterns within single spawning seasons with large amounts of heterogeneity between seasons[30], and most simulations with larvae in the water column for 30 or more days showed high levels of connectivity between most sites, with a general pattern of mainland sites as sources and the Northern and Southern Channel Islands as sinks[31]. Moreover, a comprehensive seascape genetics study of three separate species in the SCB revealed that oceanographic processes were poor predictors of genetic connectivity and diversity in species with

longer pelagic larval durations[32]. Considering the 6–10-week larval development period and year-long breeding season of *P. miniata*[15,16], oceanographic factors are most likely weak influences on the genetic connectivity and diversity of this species and were subsequently not included in any models.

In any study examining anthropogenic effects on evolutionary processes, it is critical to be able to separate historic, natural processes from contemporary anthropogenic ones[22,33]. It could be argued that the genetic structure of *P. miniata* has been shaped historically for thousands of generations by natural rivers. However, several key results from this study point to the anthropogenic nature of the observed effects.

The current generalization in landscape genetics is that microsatellites are the best genetic markers suited for studies of contemporary gene flow[22,34,35]. This is most likely because of the fact that the majority of landscape genetics studies have been on vertebrates and plants in the terrestrial systems[22] in which effective population sizes are small and genetic variation may be difficult to find. However, in an organism with high genetic diversity and a large effective population size, such as *P. miniata*, this is likely not the case and the crucial factor becomes not the mutation rate of the marker but whether or not there is enough variation within the landscape[36]. Recently isolated populations (< 50 generations) that were previously very diverse and well connected will not have time to accumulate much differentiation by mutation, even in fast-evolving markers[37]. Consequently, almost all current spatial genetic information is contained in extant alleles[37]. In fact, most spatial autocorrelation is created over 20–50 generations, and its detectable signal wanes in markers with the highest mutation rates[37]. In these situations, the only evolutionary force acting on allele frequencies in these populations will be genetic drift, which has larger effects on smaller effective population sizes[38], and has been shown to be rampant in artificially fragmented populations[39]. Therefore, considering the high genetic diversity of *P. miniata* and previous connectivity[17], we would expect to see a more pronounced signal of anthropogenic isolation in the mtDNA markers, which have an effective population ~1/4 the size of the nDNA markers[38]. In all population structure models tested, we observed that R^2 values were ~1.5–30 times higher for mtDNA markers relative to nDNA markers (Tables 2–4), although results from nDNA genetic diversity regressions were consistent with anthropogenic isolation (Fig. 3). This indicates that the observed structure is most likely due to the very recent isolation of these previously well-connected populations and not due to a historical pattern of isolation.

Wastewater variables also contain valuable information about the temporal scale of the observed gene flow patterns as they are

completely contemporary and anthropogenic. Here, we observed wastewater variables in all top multivariate models and high and significant correlations to genetic structure in single variable tests. Distance away from two stormwater sources was also significantly correlated to nDNA genetic diversity (Fig. 3). Moreover, when stormwater variables were treated as historical variables, despite the fact that stormwater runoff has been greatly augmented by urbanization and riverbed civil engineering[7], wastewater variables still explained significant amounts of variation in genetic structure not explained by 'historic variables' (Table 4). These results strongly suggest that human-mediated inputs have a profound influence on the genetic structure of *P. miniata*.

Finally, RSC is the second largest source of stormwater by volume[24]; however, it was only significantly correlated to genetic structure in one single-variable test, and that correlation was negative. RSC was also not included in many of the top multivariate models and when it was included, typically it was nonsignificant (Table 3). RSC is the least populated, least urbanized and least modified watershed of the four stormwater sources[40], and these three factors correlate directly to pollutant and contaminant loading[41]. Thus, we argue that it is contaminant loading and not freshwater river input that drives the significant deviation from panmixia detected in this study.

In summary, our results clearly indicate that coastal pollution is limiting the larval dispersal and population connectivity of *P. miniata*. However, even if the observed effects were purely from historical river input, these results would still have far-reaching implications. This is because both stormwater volume and wastewater volume have increased with the human population of the SCB, augmenting and dwarfing historical natural runoff, especially during dry weather periods[7]. Consequently, the broad, immediate conservation implications of these results would not change; the current paradigm for the incorporation of wastewater and stormwater in ecosystem-based management needs to be rethought.

Methods

Population genetic methods. Microsatellite loci were genotyped according to ref. 17 with the exception of using dye-labelled primers and an ABI 3730XL for genotyping. Two mtDNA loci were sequenced, a tRNA region and part of the cytochrome oxidase subunit I gene (COI)[17] and control region (CR)[42]. The two mtDNA loci were analysed as a LD marker (COI) and as a full mtDNA marker (COI+CR). This was carried out to add more phylogenetic information to the diverse CR loci (268 unique haplotypes). F_{ST} values for mtDNA loci were calculated using the Phi$_{ST}$ estimator of ARLEQUIN[43]. F_{ST} values for microsatellites were calculated using the Phi$_{ST}$ estimator in GenoDive[44]. D_{EST} values were calculated using SMOGD[45] and SPADE[46]. AR was calculated using rarefaction to 15 individuals per population using FSTAT[47].

Seascape genetic methods. Shortest sea distance was calculated between every population and each stormwater and wastewater source using Google Earth (http://earth.google.com). Distances were then entered into the model: PW = flow×exp(−0.1× distance), where flow is the average daily flow (in millions of gallons per day) of the contaminant source. Effluent plumes are subject to oceanographic conditions. Modelling of plumes is well beyond the scope of this project, so PW was based on a simple function of exponential decay. The coefficient of −0.01 was chosen for simplicity and for the property of decreasing ~100-fold by ~50 km and quickly approaching zero thereafter. This was an attempt to match the approximate distance between adjacent populations. Results were also examined using an alternate PW model, PW = flow/(distance)[3], approximating volumetric expansion and were similar (data not shown). PW statistics were summed for each pair of populations to create pairwise matrices of 'ecological distance' or 'resistance' for each anthropogenic source.

Matrices were analysed in an MMR framework[18,33] with F_{ST} and D_{EST} values as dependent values. For the full data set, all matrices, with the exception of full mtDNA values, were rank transformed after examining the residuals of regressions using untransformed data. The n>20 data set was untransformed. Significance was calculated using 10,000 random permutations of matrices. All MMR analyses were performed in the R statistical package[48] using the ecodist package. AICc values were calculated using JMP 8 (SAS).

Canonical redundancy analyses were performed in R using package rdaTest with 10,000 permutations and to assess significance. For single-variable tests, all three F_{ST} or all three D_{EST} were combined into one simultaneous test, and whenever geographic distance was found to be significant, it was included in all other variable tests as a covariable.

References

1. Halpern, B. S. *et al.* A global map of human impact on marine ecosystems. *Science* **15,** 948–952 (2008).
2. Halpern, B. S., Selkoe, K. A., Micheli, F. & Kappel, C. V. Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats. *Conserv. Biol.* **21,** 1301–1315 (2007).
3. Crain, C. M. *et al.* Understanding and managing human threats to the coastal marine environment. *Ann. N. Y. Acad. Sci.* **1162,** 39–62 (2009).
4. Palumbi, S. R. Humans as the world's greatest evolutionary force. *Science* **293,** 1786–1790 (2001).
5. Ashley, M. V. *et al.* Evolutionary enlightened management. *Biol. Conserv.* **111,** 115–123 (2003).
6. Bowen, B. W. & Roman, J. Gaia's handmaidens: the Orlog model for conservation biology. *Conserv. Biol.* **19,** 1037–1043 (2005).
7. Schiff, K. C., Allen, M. J., Zeng, E. Y. & Bay, S. M. Southern California. *Mar. Pollut. Bull.* **41,** 76–93 (2000).
8. Culliton, T. J. *et al. 50 Years of Population Change Along the Nation's Coast* (National Ocean and Atmospheric Association, Silver Spring, MD, 1990).
9. Lyon, G. S. & Stein, E. D. How effective has the Clean Water Act been at reducing pollutant mass emissions to the Southern California Bight over the past 35 years? *Environ. Monit. Assess.* **154,** 413–426 (2008).
10. Bay, S. M., Jones, B. H., Schiff, K. C. & Washburn, L. Water quality impacts of storm water discharges to Santa Monica Bay. *Mar. Environ. Res.* **56,** 205–223 (2003).
11. Depledge, M. H. & Billinghurst, Z. Ecological significance of endocrine disruption in marine invertebrates. *Mar. Pollut. Bull.* **39,** 32–38 (1999).
12. Bailey, H. C., Miller, J. L., Miller, M. J. & Dhaliwal, B. S. Application of toxicity identification procedures to the echinoderm fertilization assay to identify toxicity of a municipal effluent. *Arch. Environ. Contam. Toxicol.* **18,** 747–755 (1995).
13. McGibbon, S. & Moldan, A. G. S. Routine toxicity testing of toxicants using a sea urchin gamete bioassay. *Mar. Pollut. Bull.* **17,** 68–72 (1986).
14. Walker, A. N. *et al.* Bioaccumulation and metabolic effects of the endocrine disruptor methoprene in the lobster, *Homarus americanus. Int. Comp. Biol.* **45,** 118–126 (2005).
15. Strathmann, M. F. *Reproduction and Development of Marine Invertebrates of the Northern Pacific Coast* (Univ. Washington Press, 1987).
16. Rumrill, S. S. Population size-structure, juvenile growth, and breeding periodicity of the sea star *Asterina miniata* in Barkley Sound, British Columbia. *Mar. Ecol. Prog. Ser.* **56,** 37 (1989).
17. Keever, C. K. *et al.* Discordant distributions of populations and genetic variation in a sea star with high dispersal potential. *Evolution* **63,** 3214–3227 (2009).
18. Lichstein, J. W. Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecol.* **188,** 117–131 (2007).
19. Balkenhol, N., Waits, L. P. & Dezzani, R. J. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography* **32,** 818–830 (2009).
20. Kalinowski, S. T. Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity* **94,** 33–36 (2005).
21. Ruzzante, D. E. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Can. J. Fish. Aquat. Sci.* **55,** 1–14 (1998).
22. Storfer, A. *et al.* Landscape genetics: where are we now? *Mol. Ecol.* **19,** 3496–3514 (2010).
23. Baums, I. B., Paris, C. B. & Cherubin, L. M. A bio-oceanographic filter to larval dispersal in a reef building coral. *Limnol. Oceanogr.* **51,** 1969–1981 (2006).
24. White, C. *et al.* Ocean currents help explain population genetic structure. *Proc. R. Soc. B Biol. Sci.* **277,** 1685–1694 (2010).
25. Galindo, H. M., Olson, D. B. & Palumbi, S. R. Seascape genetics: a coupled oceanographic-genetic model predicts population structure of Caribbean corals. *Curr. Biol.* **16,** 1622–1626 (2006).
26. Galindo, H. M. *et al.* Seascape genetics along a steep cline: using genetic patterns to test predictions of marine larval dispersal. *Mol. Ecol.* **19,** 3692–3707 (2010).
27. Banks, S. C. *et al.* Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. *Ecology* **88,** 3055–3064 (2007).
28. Kenchington, E. L., Patwary, M. U., Zouros, E. & Bird, C. J. Genetic differentiation in relation to marine landscape in a broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). *Mol. Ecol.* **15,** 1781–1796 (2006).
29. Dong, C., Idica, E. Y. & Mcwilliams, J. C. Circulation and multiple-scale variability in the Southern California Bight. *Prog. Oceanogr.* **82,** 168–190 (2009).
30. Watson, J.R. *et al.* Realized and potential larval connectivity in the Southern California Bight. *Mar. Ecol. Prog. Ser.* **401,** 31–48 (2010).
31. Mitarai, S., Siegel, D. A., Watson, J. R., Dong, C. & McWilliams, J. C. Quantifying connectivity in the coastal ocean with application to the Southern California Bight. *J. Geophys. Res. Oceans* **114,** C10026 (2009).
32. Selkoe, K. A. *et al.* Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Mol. Ecol.* **19,** 3708–3726 (2010).

33. Storfer, A. *et al.* Putting the landscape in landscape genetics. *Heredity* **98,** 128–142 (2007).
34. Anderson, C. D. *et al.* Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol. Ecol.* **19,** 3565–3575 (2010).
35. Wang, I. J. Recognizing the temporal distinctions between landscape genetics and phylogeography. *Mol. Ecol.* **19,** 2605–2608 (2010).
36. Holderegger, R. & Wagner, H. H. Landscape genetics. *BioScience* **58,** 199–207 (2008).
37. Epperson, B. K. Mutation at high rates reduces spatial structure within populations. *Mol. Ecol.* **14,** 703–710 (2005).
38. Avise, J. C. *Molecular Markers, Natural History, and Evolution* 2nd edn (Sinaur Associates, 2004).
39. McCraney, W. T. *et al.* Rampant drift in artificially fragmented populations of the endangered tidewater goby (*Ucyclogobius newberryi*). *Mol. Ecol.* **19,** 3315–3327 (2010).
40. Ackerman, D. & Schiff, K. C. Modeling storm water mass emissions to the Southern California Bight. *J. Environ. Eng.* **129,** 308–317 (2003).
41. Stein, E. D., Tiefenthaler, L. L. & Schiff, K. C. *Understanding Sources, Patterns and Mechanisms of Pollutant Loading from Urban, Arid Watersheds and Land Uses of the Greater Los Angeles, California, USA* (Southern California Coastal Water Research Project, Technical Report #510, 2007).
42. Waters, J. M., O'Loughlin, P. M. & Roy, M. S. Cladogenesis in a starfish species complex from Southern Australia: evidence for vicariant speciation. *Mol. Phylogenet. Evol.* **32,** 236–245 (2004).
43. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **3,** 564–567 (2010).
44. Meirmans, P. G. & Van Tienderen, P. H. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **4,** 792–794 (2004).
45. Crawford, N. G. SMOGD: software for the measurement of genetic diversity. *Mol. Ecol. Resour.* **10,** 556–557 (2010).
46. Chao, A. & Shen, T. J. *Program SPADE (Species Prediction And Diversity Estimation)* (2010) http://chao.stat.nthu.edu.tw.
47. Goudet, J. Fstat version 1.2: a computer program to calculate Fstatistics. *J. Hered.* **86,** 485–486 (1995).
48. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2009).

Acknowledgments

Author contributions

J.B.P. designed the study, performed all data collection and analysis, and prepared the manuscript. R.J.T. supervised all aspects of this research along with contributing extensively to design, data analysis and manuscript revision.

Additional information