

## LETTERS

# On the importance of habitat continuity for delimiting biogeographic regions and shaping richness gradients

Phillip B. Fenberg<sup>1,2\*</sup> and  
Marcelo M. Rivadeneira<sup>3,4,5</sup>

### Abstract

The formation and maintenance of biogeographic regions and the latitudinal gradient of species richness are thought to be influenced, in part, by the spatial distribution of physical habitat (habitat continuity). But the importance of habitat continuity in relation to other variables for shaping richness gradients and delimiting biogeographic regions has not been well established. Here, we show that habitat continuity is a top predictor of biogeographic structure and the richness gradient of eastern Pacific rocky shore gastropods (spanning *c.* 23 000 km, from 43°S to 48°N). Rocky shore habitat continuity is generally low within tropical/subtropical regions (compared to extratropical regions), but particularly at biogeographic boundaries where steep richness gradients occur. Regions of high rocky shore habitat continuity are located towards the centres of biogeographic regions where species turnover tends to be relatively low. Our study highlights the importance of habitat continuity to help explain patterns and processes shaping the biogeographic organisation of species.

### Keywords

Biogeographic structure, diversity gradient, gastropods, habitat continuity, range limits, rocky intertidal.

*Ecology Letters* (2019) **22**: 664–673

## INTRODUCTION

The availability of physical habitat for a species or a group of species is not evenly distributed across space. Spatial clusters and/or long stretches of unsuitable habitat can limit dispersal (Pearson & Dawson 2005), cause range limits (Sexton *et al.* 2009) and, by extension, delimit biogeographic regions and shape regional species richness gradients (Whitton *et al.* 2012). For example, depending on the habitat specificity of the group of species in question, if habitat is sparse within regions spanning mid to low latitudes where temperate climates coincide with subtropical/tropical climates, then species turnover across these regions may be particularly abrupt due to a combination of thermal stress and low habitat availability, resulting in a biogeographic boundary. Such boundaries should also mark a steep richness gradient given that tropical regions are expected to be more species rich than extratropical regions (Edgar *et al.* 2017; Ibanez-Erquiaga *et al.* 2018; Kinlock *et al.* 2018). Likewise, if biogeographic boundaries and associated species richness gradients are delimited and shaped, in part, by low habitat continuity (sparse habitat), then, relatively speaking, the central portions of biogeographic regions should be characterised by high habitat continuity (plentiful habitat) and comparatively lower species turnover. Thus, the spatial distribution of habitat, and its geographical coincidence with climatic transitions may influence the overall shape

of the richness gradient by affecting the slope and location of gradients.

Despite the importance of the spatial distribution of habitat for affecting ecological and evolutionary processes that likely contribute to the formation and maintenance of the species richness gradient and bioregionalisation (Hanski 1999; Jocque *et al.* 2010; Sanciangco *et al.* 2013; Schluter 2016), there are no studies that have quantitatively examined the importance of habitat continuity in explaining those patterns. This could be partly due to the fact that studies of regional/global datasets often use coarse spatial resolutions (e.g. 1–5° grid cells) to bin biological (species presence/absence) and physical/geological variables, resulting in multiple habitat types and/or depth/elevational gradients within each bin (e.g. Valdovinos *et al.* 2003). In addition, the shapes of richness gradients are usually viewed latitudinally (Hillebrand 2004; Chaudhary *et al.* 2016), but the distribution of available habitat for any group is not likely to be evenly or linearly spread within and across latitudes (Rosenzweig 1995; Fernandez & Marques 2017). This can make attempts to isolate the effect of habitat continuity on biogeographic structure and the shape of richness gradients less clear. One way to potentially assess the role of habitat continuity for predicting these biogeographic patterns is to produce a habitat map that spans tropical and extratropical regions and couple it with the geographical ranges of species restricted to that habitat type.

<sup>1</sup>Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton SO14 3ZH, UK

<sup>2</sup>Department of Life Sciences, The Natural History Museum, London SW7 5BD, UK

<sup>3</sup>Laboratorio de Paleobiología, Centro de Estudios Avanzados en Zonas Áridas (CEAZA), Av. Bernardo Ossaóndón 877, C.P. 1781681, Coquimbo, Chile

<sup>4</sup>Departamento de Biología Marina, Universidad Católica del Norte, Av. Larraondo 1281, Coquimbo, Chile

<sup>5</sup>Departamento de Biología, Universidad de La Serena, Av. Raúl Bitrán 1305, La Serena, Chile

\*Correspondence: E-mail: p.b.fenberg@soton.ac.uk

Rocky shore habitats are spread across the globe and host a diverse set of species, many of which are not found in other habitat types (e.g. soft sediment substrates). Interestingly, marine geomorphologists have noted that the distribution of rocky substrate along the inner continental shelf (< 65 m) is not evenly distributed latitudinally; they tend to prevail around mid to high latitudes but appear sparse in tropical areas, where they are often replaced by unconsolidated sediment habitats, such as coastal plains and sandy beaches (Hayes 1967). This suggests that rocky shore habitat continuity may be particularly low in tropical regions compared to extratropical regions, which in turn, may have implications on the formation and structuring of biogeographic patterns (see above). With the aid of satellite images and new methods, quantitative measurements of the distribution of coastal habitat types at large geographical scales is now underway (see Luijendijk *et al.* (2018) for a global map of sandy beaches).

The eastern Pacific (EP) coast is uniquely suited for exploring the relationship between coastal habitat distribution and biogeographic patterns because it spans uninterrupted across much of the latitudinal extent of the globe. Of the species inhabiting the EP rocky shore, gastropods are particularly suitable for study because they are relatively diverse, extensively collected for natural history museums, and generally well-studied in the ecological/taxonomic literature, making their geographical range extents straightforward to estimate.

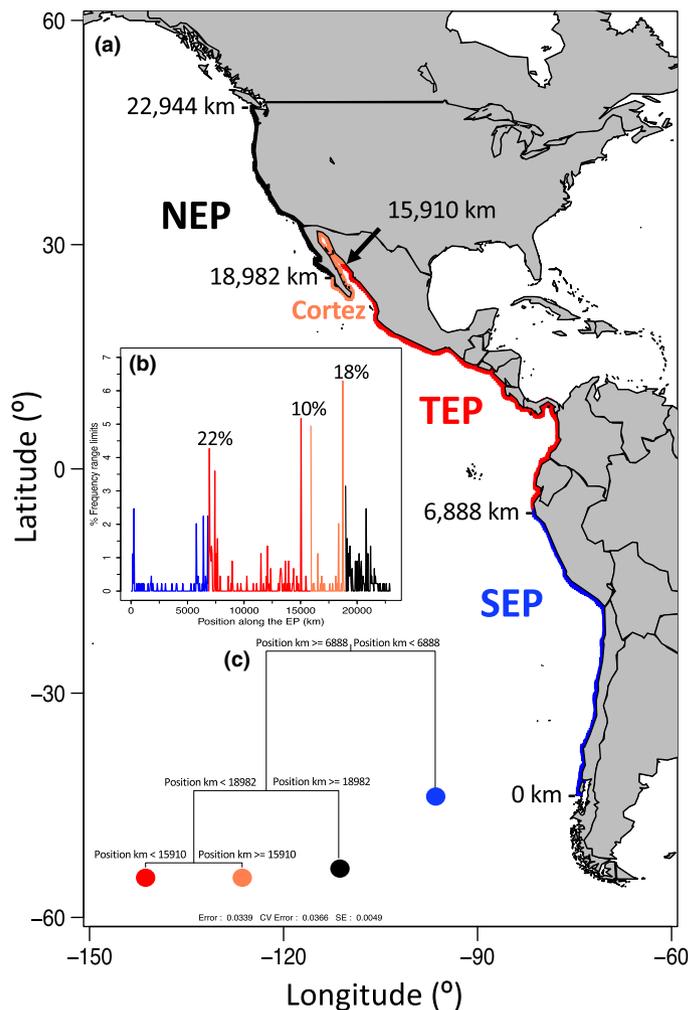
In this study, we hypothesised that rocky shore habitat continuity is a major driver of bioregionalisation and the species richness gradient of rocky shore marine gastropods. We tested this idea by creating a transect that follows the contours of most of the EP coast (> 90° latitude and *c.* 23 000 km) and mapping the distribution of rocky shore habitat along the transect using Google Earth. We show that regional variation in habitat continuity, in combination with sea-surface temperature and other oceanographic variables, is ranked as a top predictor of both biogeographic structure and the richness gradient across the EP transect; a result not previously recognised across such a large geographical scale. Moreover, we find that rocky shore habitat is relatively sparse in the tropical/subtropical regions compared to the extratropical regions. But within and across all biogeographic regions, there is significantly more rocky shore habitat towards the centres of biogeographic regions compared to their edges.

## METHODS

### Transect, habitat map and habitat continuity

Despite the widespread availability of oceanographic and environmental variables for biogeographic research (e.g. Tyberghein *et al.* 2012), there are currently no databases that include the distribution of coastal rocky substrate. To measure this for the EP, we traced around the contours of the outer mainland coastline at a constant elevation of 5 km above sea level (viewed perpendicularly) in cumulative 48 km (*c.* 30 mile) bins using the software program Google Earth (v. 7.1.5; similar to Fenberg & Rivadeneira 2011; Martínez *et al.* 2017; Aguilera *et al.* 2018). The transect does not cover deep coastal inlets such as river estuaries (e.g. San Francisco Bay)

or offshore islands (e.g. the Galapagos, the Channel Islands of California). The starting, ending and midpoint coordinates of each bin were recorded as both latitude/longitude and as position along the transect in kilometres. The transect runs from the southern end of Chiloé island in Chile (43.38° S, 74.12° W; 0 km) to northern Washington State in the USA (48.37° N, 124.72° W; 22 944 km), and consists of 478 bins (Fig. 1a). Some sections of the transect share latitudes with multiple coastlines, most notably along the Gulf of California coast of mainland Mexico and the Baja California peninsula (*c.* 23° N–31° N; 15 000–20 000 km). We did not extend the



**Figure 1** (a) Map of the EP transect with boundaries of the major biogeographic regions identified by position (km) along the transect. The transect extends from Chiloé island, Chile (43.38° S, 74.12° W; 0 km) to northern Washington State, USA (48.37° N, 124.72° W; 22 944 km) and consists of 478 equally spaced bins (each 48 km long) (b) The frequency distribution of range limits of EP rocky shore gastropods. The three largest clusters of range limits mark the spatial transitions between each biogeographic region. The numbers above each range limit cluster are the % of the total number of range limits ( $n = 445$ ) which occur within these biogeographic border regions. Together, these three regions account for 50% of the range limits of EP rocky shore gastropods. (c) The CART analysis (cross-validated error = 0.037; pseudo- $R^2 = 0.963$ ) showing the locations of the major biogeographic breaks based on the species presence-absence matrix.

transect south into Patagonia or north into British Columbia/Alaska because of their convoluted fjord-land coastal geomorphology.

Once the transect was created, we then measured the amount of coastline that consisted of unsuitable habitat within each bin, usually beach or other soft sediment habitat, such as mangrove flats or river deltas. We added the lengths of each stretch of unsuitable habitat, divided the sum by the total length of each bin (48 km) and calculated a percentage of unsuitable habitat per bin, with its reciprocal being the percentage of rocky shore habitat. In addition, we combined the lengths of unsuitable habitat stretches that span adjacent borders of bins to determine the extent of each habitat gap (Fig. S1). Regional variation in the percentage of rocky shore per bin does not significantly differ when compared to a portion of the transect created at an elevation of 500 m (Fenberg & Rivadeneira 2011; Fig. S2).

To estimate regional patterns of habitat continuity, we plotted the percentage of rocky shore per bin and fit a local regression smoothing curve through the data (LOESS; span 0.2). To calculate values useful for prediction purposes (see below), we interpolated data points along this curve for the midpoint of each bin. This method gives us a quantitative estimate of regional patterns of habitat continuity with high predictive value for richness and biogeographic structure analyses while removing some of the bin by bin local scale variability. We refer to regions of high and low habitat continuity based on where the major peaks (high continuity) and dips (low continuity) on the LOESS curve are located.

### Range extents

We used the species list generated by Rivadeneira *et al.* (2015) and a subsequent update to identify a baseline set of shelled rocky shore gastropod species to estimate range extents. The Rivadeneira *et al.* (2015) dataset was compiled from published biodiversity surveys of rocky intertidal field sites spanning the EP. Of the species surveyed ( $n = 368$ ), we only included those if they are known to primarily have their native ranges restricted to the EP mainland, spend their adult lives in rocky shore habitats, and their taxonomy is not currently in question. We used a combination of occurrence databases and visits to natural history museums (see acknowledgements) to identify the northern and southern range limits for each species. If available, we supplemented the museum data with published records of species distributions based on museum collections and field surveys (e.g. Reid 2002; Fenberg *et al.* 2015). If a species was poorly represented in museum collections (e.g. few specimens or poor metadata) or the literature, it was excluded from the dataset. The resulting database includes 246 species with ranges from 55.6° S to 63.5° N. For each species with a range totally within the transect ( $n = 199$ ), we calculated the average % of rocky shore across all bins spanning their range extents. We then plotted these values against the location of their range midpoint along the transect. This allows us to test whether species from extratropical regions have more rocky shore habitat available, as a percentage of their total range extents, compared to species from tropical/subtropical regions. We also tested for differences in

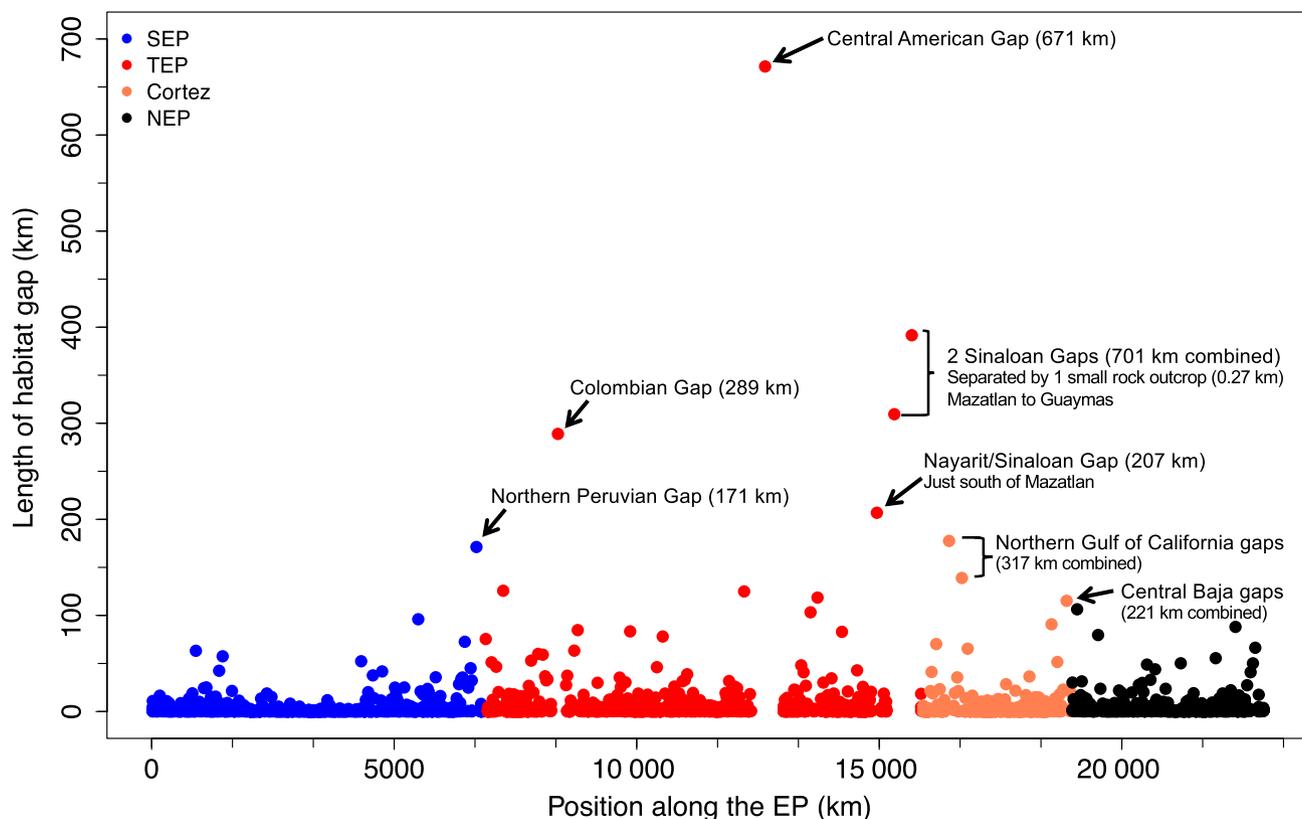
total range extents (total km of coastline within species' range limits) and potentially habitable range extents (total km of coastline – km of unavailable habitat within species' range limits) between species from extratropical vs. tropical regions.

### Biogeographic structure and richness

We used a classification and regression tree analysis (CART; Fenberg *et al.* 2015) to determine the biogeographic structure of EP rocky shore gastropods. CART is a divisive clustering method for multivariate response data (distance matrix of species' presence-absence) that bases binary divisions on explanatory variables (i.e. position along the transect), which act as constraints on the formation of classification groups. Each species was entered into a presence-absence matrix based on its range extent along the EP transect. A species was recorded as present in a bin if its range extent included that bin and if the bin did not consist of 99–100% unsuitable habitat ( $n = 69$ ). A CART, based on the Bray-Curtis similarity index after performing a Hellinger's transformation (needed in order to deal with the excess of zeroes; Legendre & Gallagher 2001; Legendre & Borcard 2018), was carried out on the species matrix using the midpoint of each spatial bin as a splitting variable, with a cross validation (xv) procedure, selecting the best predictive tree (xv = 'min') using the default complexity parameter (cp) of 0.01 (De'ath & Fabricius 2000; De'ath 2005; Fig. S3). Analyses were carried out using the package mvpart in R (De'ath 2005). Species were subsequently matched to their respective biogeographic regions based on where their range midpoint occurs.

Once biogeographic regions were identified, we tested whether the central portions of each region contain more rocky shore habitat compared to their edges. Within each region, we categorised each 48 km bin into one of four equal sections (using quartiles). Bins within the two outer sections of a biogeographic region are categorised as the edges and bins within the two inner sections are categorised as the centre. We also tested for differences in the percentage of rocky shore between centre and edge sections pooled across the whole transect (regardless of region) and overall between tropical/subtropical regions vs. extratropical regions.

Species richness is defined as the sum of species estimated to be present within in each bin based on their range extents, excluding bins that consist of 99–100% unsuitable habitat. This produces an estimate of regional variation in species richness across the transect (i.e. not based on local species occurrences, which would require standardised field surveys of each bin). The overall shape of the richness gradient would look different if we were to use latitude as a spatial bin. For example when viewed using 1° latitudinal bins, the tropical/subtropical regions appear to cover a much smaller proportion of the EP than they do in reality because the coastline in these regions have some longitudinal trends and the Baja California peninsula shares the same latitudinal range with a large portion of mainland Mexico. In contrast, the Chilean, Peruvian and much of the North American coast is relatively linear, covering more degrees of latitude per bin compared to Central America (Figs S4 and S5).



**Figure 2** Position and length of each habitat gap along the EP transect (coloured by biogeographic region), with the largest gaps identified by name. Each point represents a single gap, ranging in size from 0.03 km to 671.3 km (CAG). The largest habitat gaps occur within the TEP/Cortez and towards the edges of the SEP and NEP.

### Random forest analyses

A random forest analysis (Breiman 2001; Liaw & Wiener 2002) was used to relate the biogeographic structure (i.e. biogeographic regions defined by the methods explained above) and species richness with the predictor variables (see below). Random forest is a powerful machine-learning method of growing use in ecology and biogeography (Cutler *et al.* 2007; Fenberg *et al.* 2015) based on an assemblage of bootstrapped classification or regression trees. Random forest is a nonparametric method that can deal with continuous (richness gradient) and categorical (biogeographic region) responses and makes no assumptions about the residuals of the models, like traditional multivariate regressions. Analyses were carried out using the library `randomForest` in R (Liaw & Wiener 2002). The accuracy of the model (pseudo- $R^2$ ) was assessed as  $100 - [\% \text{ out-of-bag error}]$  for classification, and  $\left(1 - \frac{\text{MSE}_{\text{OOB}}}{\delta^2}\right)$  for regression (where  $\text{MSE}_{\text{OOB}}$  is the mean squared residuals of the out-of-bag error and  $\delta^2$  the variance of the predicted values). We paired each spatial bin with a set of 13 predictor variables commonly thought to be important for predicting biogeographic structure and richness gradients (Belanger *et al.* 2012; Fenberg *et al.* 2015). Oceanographic variables used were: (1) sea-surface temperature (SST), (2) salinity, (3) chlorophyll-a, (4) dissolved oxygen, (5) nitrate, (6) phosphate, (7) silicate, (8) calcite, (9), photosynthetically available radiation (PAR), (10) average tidal amplitude, (11) primary

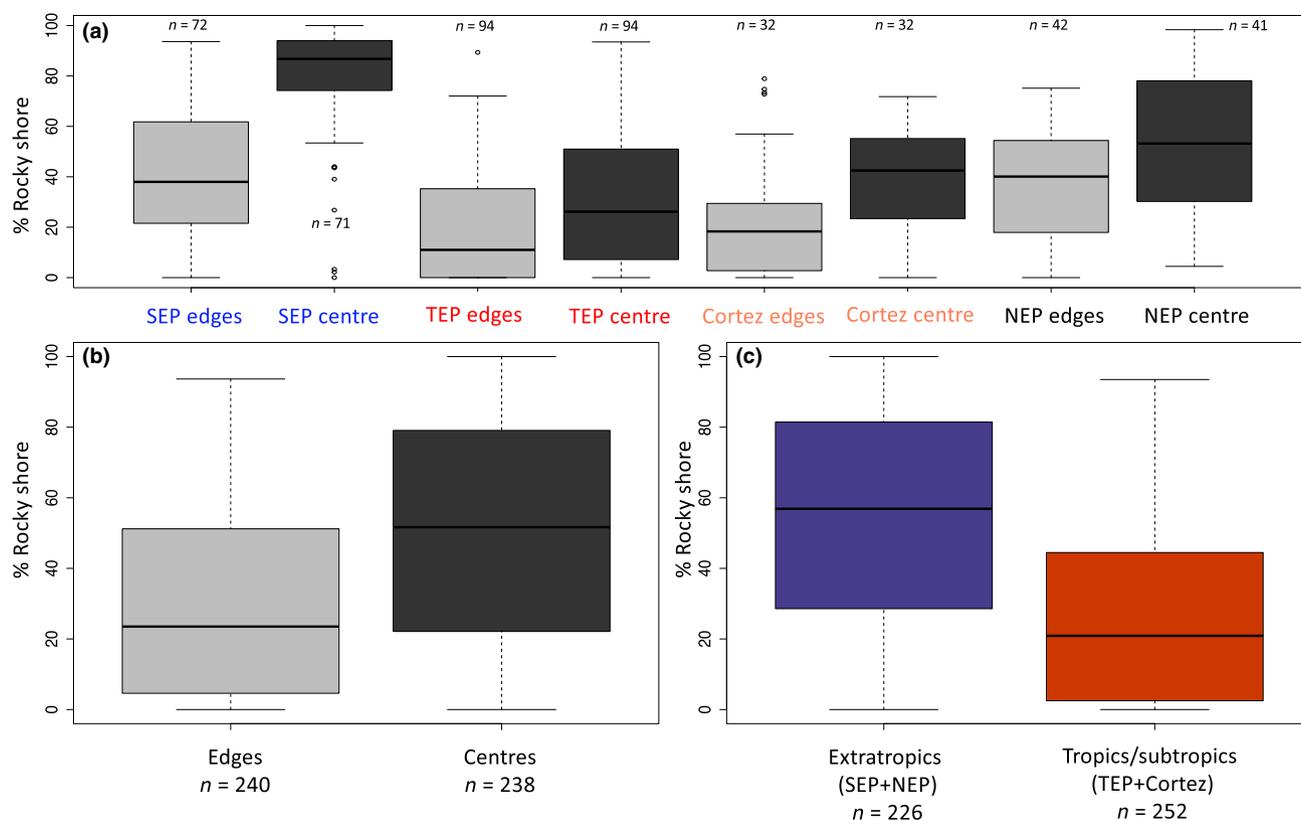
productivity, (12) pH and (13) the habitat continuity variable. The oceanographic variables (1–12) were obtained from the Bio-ORACLE database (Tyberghein *et al.* 2012) and the Hexacoral/LOICZ database ([http://hercules.kgs.ku.edu/hexacoral/envirodata/hex\\_modfilt\\_firststep3dev1.cfm](http://hercules.kgs.ku.edu/hexacoral/envirodata/hex_modfilt_firststep3dev1.cfm)). The spatial resolution of variables was uniformed to 0.5°.

The relative importance of each predictor variable for each model was evaluated as the mean decrease in accuracy (structure analyses) and % increase in mean squared error (richness analysis), and its statistical significance was tested using the permutation algorithm implemented in the library `rfrPermute` in R using 10 000 runs (Archer 2016). In order to minimise the effect of multi-collinearity among predictor variables, which could bias assessments of variable importance (Strobl *et al.* 2008), we ran a variance inflation factor (VIF) analysis and selected only variables with a correlation threshold of  $r < 0.80$ , using the package `usdm` in R (Naimi 2015). If any variables have a  $\text{VIF} > 10$ , they were assessed with and without their inclusion in the models.

## RESULTS

### Biogeographic structure

There is clear biogeographic structure of rocky shore gastropods along the transect (Fig. 1b and c). The CART analysis produced a tree with a very low cross-validated error of



**Figure 3** (a) Boxplots of the percentage of rocky shore within bins after equally subdividing each biogeographic region into their edge sections and the centre section (see Methods). Within each biogeographic region, there is significantly more rocky shore in the centre section compared to the edge sections ( $P < 0.005$  for all comparisons) and (b), when centre and edge sections are pooled across all biogeographic regions,  $P < 0.0001$ ). (c) Boxplots showing that the percentage of rocky shore within bins spanning extratropical regions (SEP+NEP) is significantly higher ( $P < 0.0001$ ) compared to bins spanning tropical/subtropical regions (TEP+Cortez).

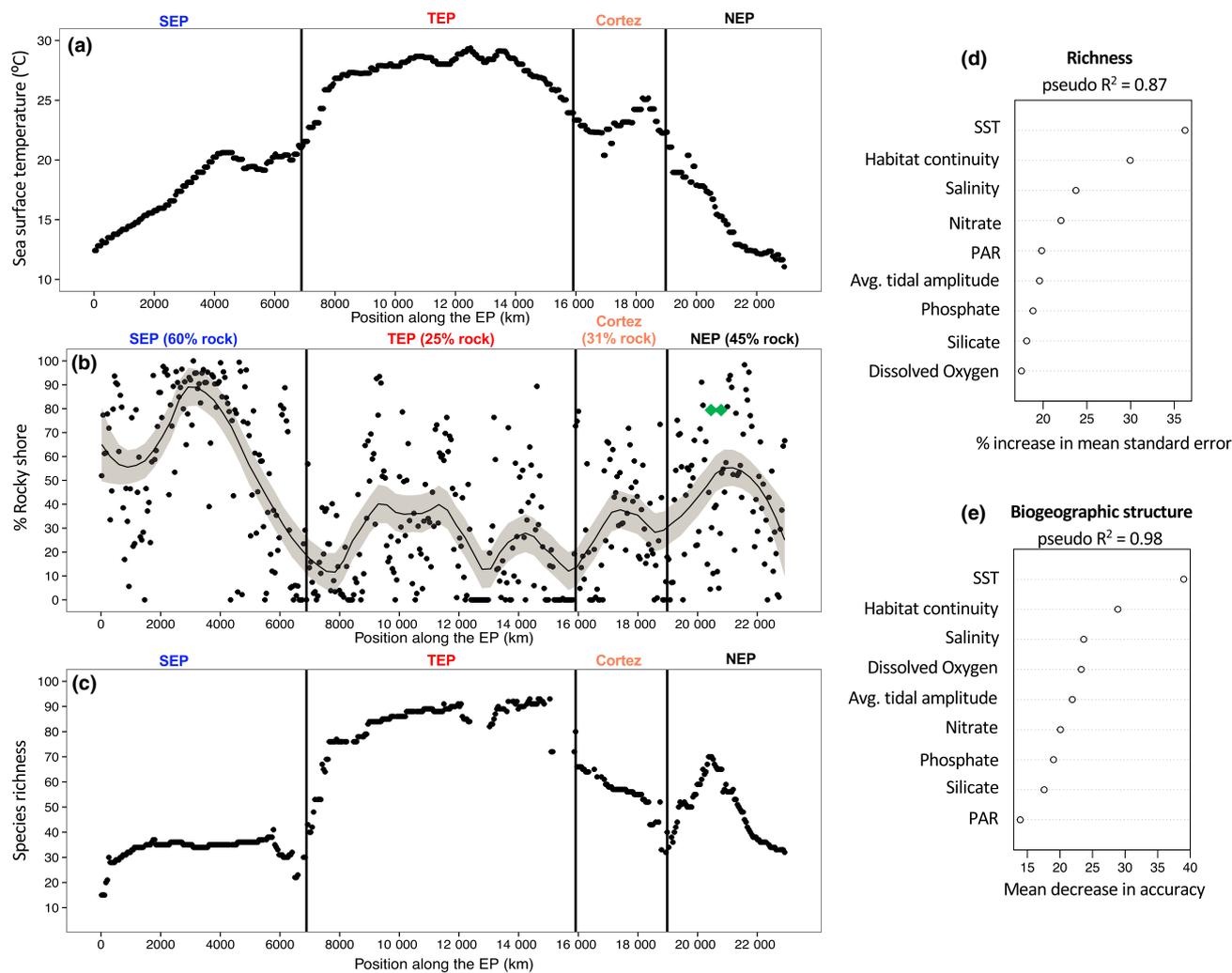
0.037 (i.e. pseudo- $R^2 = 0.963$ ), with three splits (i.e. four biogeographic regions; Fig. 1c). The primary split is at the 6,888 km position ( $5.52^\circ$  S,  $80.95^\circ$  W) in the Piura region of northern Peru. The beginning of the transect to this boundary is categorised as the southeastern Pacific (SEP). Following the coast north from this boundary, the next split is at 15 910 km ( $27.91^\circ$  N,  $110.97^\circ$  W), located in the area near Bahia San Carlos on the Gulf of California coast of mainland Mexico. This region, between 6888 and 15 910 km, is the Tropical Eastern Pacific (TEP). The final split at the 18 982 km position ( $26.17^\circ$  N,  $112.42^\circ$  W) is located north of Magdalena Bay on the Pacific coast of Baja California, near San Juanico. This region (between 15 910 and 18 982 km) constitutes the Cortez province, which has long been known to harbour a subset of TEP fauna with some endemic subtropical marine fauna (Brusca 1980; Reid 2002; Robertson & Cramer 2009). The coast from 18 982 km to the end of the transect constitutes the northeastern Pacific (NEP) region. The region just south of the transect, from ~ Chiloló island and around Cape Horn is often delimited as the Magellanic region (Spalding *et al.* 2007), whereas the region north of the transect is often thought of as a transition zone between the Columbian and Aleutian provinces (Fenberg *et al.* 2015). Thus, we are confident that the EP transect covers four major biogeographic regions/units. The four-region biogeographic structure is also notable when plotting

total range extents or potentially habitable range extents vs. range midpoints (Fig. S6).

#### Habitat map and continuity

The largest habitat gaps are within tropical/subtropical regions (TEP and Cortez) and towards the edges of the extratropical regions (SEP and NEP; Fig. 2). 60% (13 784 km) of the transect consists of unsuitable habitat for rocky shore species, meaning that at least 40% of the transect is composed of rocky shore (9160 km). The majority of this 40%, however, is within the extratropical regions of the transect. The SEP consists of 60% rocky shore (4138/6864 km) and the NEP consists of 45% rocky shore (1801/3984 km). This is in stark contrast to the TEP and Cortez province, which, respectively, contain 25% (2263/90242 km) and 31% (958/3072 km) rocky shore. As a whole, the percentage (and total amount) of rocky shore within extratropical bins (SEP and NEP) is significantly greater than in the tropical/subtropical bins (TEP and Cortez; Kruskal–Wallis test,  $P < 0.0001$ ; Fig. 3c).

The centre of each biogeographic region contains a significantly higher percentage and total amount of rocky shore than the edges (Kruskal–Wallis tests,  $P < 0.005$ ; Fig. 3a). This pattern holds when the centre and edge bins are pooled across regions (Kruskal–Wallis test,  $P < 0.0001$ ; Fig. 3b). This is notable because each biogeographic boundary coincides with



**Figure 4** (a) Sea-surface temperature (SST) along the EP transect. Note the relatively stable SST within much of the TEP and large gradients at biogeographic boundaries (marked by vertical black lines). (b) Percentage of rocky shore habitat within each bin along the transect. The smoothing curve through the data (LOESS, span 0.2, with 95% confidence intervals) reveals the regional patterns of habitat continuity. The regions of low habitat continuity (dips in the curve) coincide with biogeographic boundaries (with the exception of the CAG; 12 360–13 030 km; see text) and steep richness gradients (see 4c below), whereas regions of high habitat continuity (peaks in the curve) are located towards the centres of biogeographic regions. The two green diamonds within the NEP represent the location and approximate % of rocky shore for the southern and northern Channel Islands of California. (c) The gradient in species richness of EP rocky shore gastropods along the transect. Each point represents the sum of the number of species present within each bin based on their range extents (excluding bins without rocky shore). (d) Results from the random forest analyses for predicting the richness gradient and (e) biogeographic structure. Variables are ordered according to their importance using permutation (measured as % increase in mean standard error for richness and mean decrease in accuracy for the structure analysis). Both SST and habitat continuity are the top two predictors of biogeographic structure and richness, with salinity as the third ranked predictor. Only the significant variables are presented, see Fig. S7 for the same results presented with all variables.

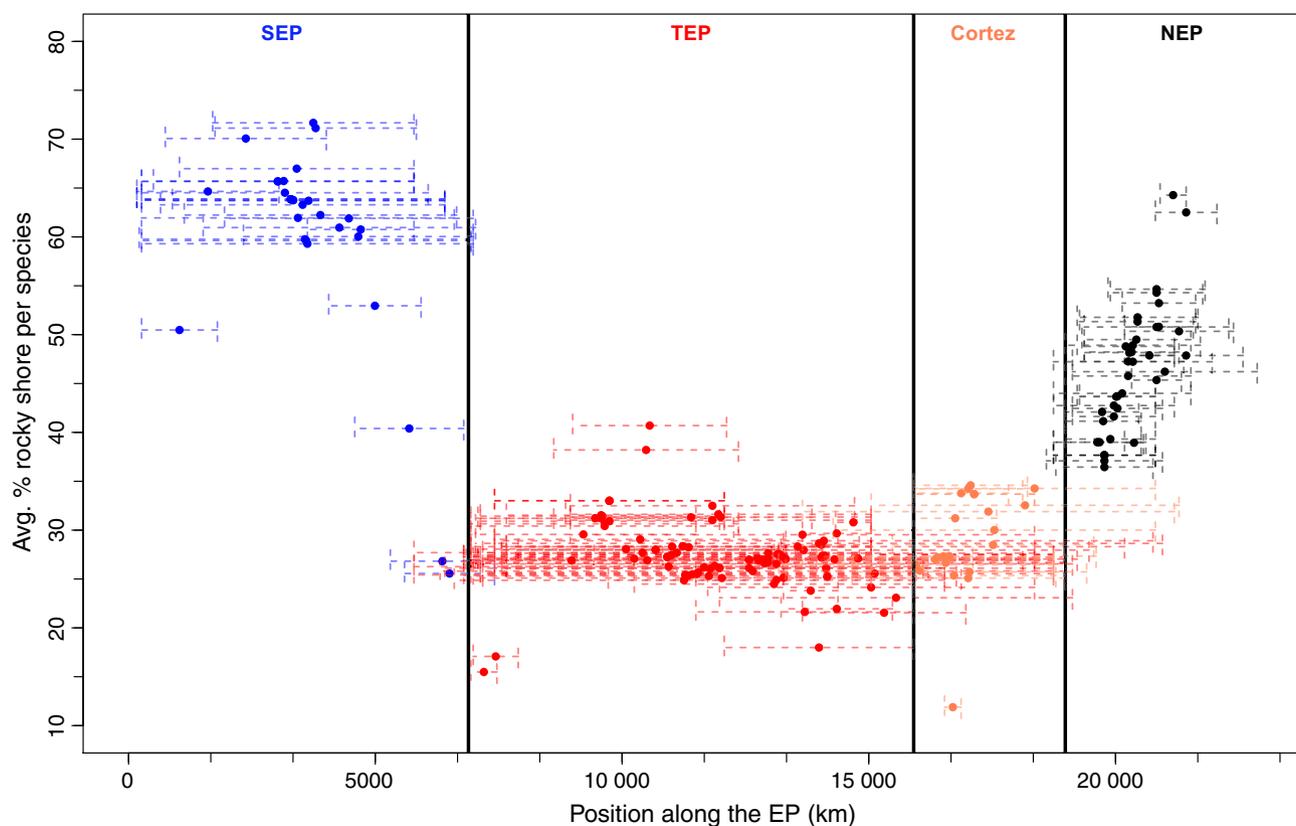
a region of low habitat continuity, whereas the regions of high habitat continuity are located towards the centres of each biogeographic region (Fig. 4b and Table S1). A clear exception to this pattern is the region of low habitat continuity surrounding the Central American Gap (CAG) within the TEP (c. 12 360–13 030 km; Figs 2 and 4b).

On average, the range extents of SEP and NEP species consist of 60 and 46% rocky shore respectively. The range extents of TEP and Cortez species, on the other hand, consist of 27 and 29% rocky shore respectively (Fig. 5). These differences between extratropical (NEP and SEP;  $n = 70$ ) and tropical/subtropical species (TEP and Cortez,  $n = 129$ ) are significant (Kruskal–Wallis test:  $P < 0.0001$ ). The total range extents

(measured as total km of coastline) of tropical/subtropical species are significantly larger than extratropical species (Kruskal–Wallis test:  $P < 0.0001$ ); but, this difference disappears when comparing potentially habitable range extents (total km of coastline–amount of coastline without rocky shore;  $P = 0.169$ ). Note, however, that these measurements must exclude extratropical species with some portion of their range outside of the transect ( $n = 47$ ).

#### Richness gradient

EP rocky shore gastropods reach their highest richness within the TEP. Steep gradients in regional richness occur at each



**Figure 5** Average percentage of rocky shore habitat within the geographical range extents of each species with a range totally within the transect ( $n = 199$ ). Solid dots are the range midpoints and dashed lines are the range extents of each species. Species are coloured according to which region the midpoint of its geographical range occurs. Species with midpoints within the TEP and Cortez regions have significantly less rocky shore habitat as a percentage of their total range extents than SEP and NEP species (Kruskal–Wallis test,  $P < 0.00001$ ), despite the fact that they have larger overall range extents (Kruskal–Wallis test,  $P < 0.00001$ ). Note, this figure and analyses must exclude species that also have some portion of their ranges north (NEP) or south (SEP) of the transect ( $n = 47$ ).

biogeographic boundary along the transect (Fig. 4c), which coincide with regions of low habitat continuity (Fig. 4b). On the other hand, regions of relatively high habitat continuity, within the centres of biogeographic regions, are not generally associated with steep richness gradients. The central portion of the SEP contains the highest levels of habitat continuity along the transect, which corresponds to a long plateau in richness and few range limits (Figs 1b and 4b and c). The steep richness gradient spanning the boundary between the SEP and TEP (northern Peru through Ecuador) occurs within a large region of low habitat continuity. The two regions of relatively high habitat continuity within the TEP, Panama through Costa Rica (*c.* 9 300–11 500 km) and southern Mexico (*c.* 13 890–14 800 km) correspond to the two highest sections of richness along the transect, which are bisected by the dip in richness surrounding the CAG. The Cortez is characterised by a richness plateau in its centre where habitat continuity is relatively high, but by steep richness gradients coinciding with regions of low habitat continuity at its southern and northern edges (marked by the Sinaloan and central Baja gaps respectively; Figs 2 and 4b and c). Finally, the central NEP corresponds to a region of high habitat continuity just north of Point Conception in Central California (*c.* 20 800–21 200 km), located near the

peak in NEP richness in southern California (*c.* 20 500 km; Fig. 4c).

#### Predictors of biogeographic structure and richness

The random forest analysis allows us to quantitatively assess and rank the importance of habitat continuity, amongst other environmental variables, in predicting these biogeographic patterns. SST and regional habitat continuity are the top two predictors (out of 12 total variables, nine of which are significant for the richness and structure analyses; Fig. 4d and e). Note, pH was not included in the final models as it exceeded the correlation threshold ( $r > 0.80$ ). The accuracy of both models (pseudo- $R^2$ ) were high: 0.87 for the richness gradient analysis and 0.98 for the biogeographic structure analysis. None of the variables used in the final models had a VIF  $> 10$ , although SST was highest (VIF = 9.9). Further results and figures are in the supplementary information.

#### DISCUSSION

The distribution of habitat across large geographical scales is known to affect dispersal, metapopulation and micro/macroevolutionary dynamics (MacArthur & Wilson 1967;

Hanski 1998; Jocque *et al.* 2010; Norris & Hull 2012). Together, these ecological and evolutionary dynamics interact over space and time to contribute to the origin and maintenance of the latitudinal diversity gradient and to structure biogeographic regions (Mittelbach *et al.* 2007; Roy & Goldberg 2007; Jablonski *et al.* 2017). However, surprisingly little research has been aimed at determining whether habitat continuity can help predict the shape of richness gradients and biogeographic structure. Part of the issue has likely been that (1), estimates of habitat continuity have not been established across geographical scales large enough to be useful for such analyses, (2) richness gradients are classically assessed latitudinally, but habitat continuity cannot be accurately assessed this way and (3), such analyses should focus on species restricted to a common habitat type found throughout the globe, which has not typically been done. By focusing our study on species restricted to rocky coastline along a continuous transect spanning extratropical and tropical/subtropical regions, we are able to identify where regional changes in habitat continuity coincide with regional changes in species turnover/richness and their spatial coincidence with other environmental variables (e.g. SST). Our results provide the first empirical support for the role of habitat continuity as a top predictor of biogeographic structure and the overall shape of the species richness gradient (Fig. 4d and e).

The locations of each biogeographic boundary and associated regional species richness gradients coincide with regions of low habitat continuity – with richness being particularly low, followed by steep gradients, where the extratropics (SEP and NEP) meet the tropics/subtropics (TEP and Cortez; Fig. 4b and c). These areas of low habitat continuity likely act as strong regional filters affecting species occurrence. On the other hand, regions where habitat continuity is relatively high are located within the centres of biogeographic regions and have comparatively lower species turnover. Thus, populations within the centres of biogeographic regions should benefit from higher immigration and larval retention simply due to shorter distances between habitat patches, compared to populations closer to biogeographic edges where habitat availability is comparatively low (Fig. 3a and b). All else being equal, dispersal limitation at the edges of biogeographic boundaries caused by low habitat continuity would contribute to the observed clusters of range limits at these boundaries (Fig. 1b) where steep regional richness gradients occur (Fig. 4b and c). Of course, much of this would be dependent upon the dispersal potential of each species.

While we have not split species into dispersal potential categories given the current lack of knowledge for many species (particularly for tropical species; Marshall *et al.* 2012; Pappalardo & Fernández 2014), we hypothesise that those with low dispersal potential (i.e. short or absent pelagic larval phase) would be more directly affected by regional patterns of habitat continuity (e.g. Fenberg *et al.* 2014) compared to those with a longer pelagic phase. Interestingly, the single largest habitat gap on the transect occurs within the centre of the TEP (CAG; Fig. 2), and does not coincide with many species range limits or a biogeographic boundary (Fig. 1), though a notable dip in richness surrounds the CAG (Fig. 4c). This suggests that many TEP species are capable of dispersing

across the CAG and may have relatively high dispersal potential (Robertson & Cramer 2009). But in order to more fully assess the degree to which the CAG and other tropical habitat gaps may limit dispersal, population genetic studies, coupled with field surveys of abundance, distribution, larval dispersal potential (plus its interaction with temperature; O'Connor *et al.* 2007) and recruitment success for tropical species are needed (Meyers *et al.* 2013; Marchant *et al.* 2015). Nevertheless, the total range extents of tropical/subtropical species tend to be larger than extratropical species (measured as total km of coastline between species' range limits) partly because many TEP species appear capable of dispersing across large habitat gaps such as the CAG. Large geographical range has been linked to lower extinction rates (Jablonski 2005; Gaston & Fuller 2009; Orzechowski *et al.* 2015), which is often one of the strongest components for explaining the latitudinal diversity gradient in diversification dynamics models (Jablonski *et al.* 2013; Rolland *et al.* 2014), including EP rocky shore gastropods (Rivadeneira *et al.* 2015). In addition, range overlap will increase if there are more species with large range extents in a region, which can indirectly influence richness gradients (Colwell & Lees 2000). If large range extent is a partial cause of the tropical richness gradient, then the question arises why low latitude marine species have such broad range extents (Tomašových *et al.* 2015)?

Temperature is generally accepted to be one of the best overall predictors of richness gradients and biogeographic structure (Field *et al.* 2009; Belanger *et al.* 2012; Fenberg *et al.* 2015), which our study corroborates (Fig. 4d and e). Nearshore tropical marine regions often have relatively constant SST (Tomašových *et al.* 2015), also observable across much of the TEP (Fig. 4a). This large region of stable temperatures may help facilitate the comparatively large geographical range extents of tropical/subtropical species (Fig. 5). In addition, the extratropical regions are more linearly oriented in a north–south direction than the tropical/subtropical regions (Fig. S4), and multiple sections of coastline within the tropics/subtropics share the same latitudinal range (e.g. Panama, Gulf of California). Thus, not only do the tropical regions have more stable temperatures, they also contain more total coastline (e.g. 9022 km for the TEP, vs. 6888 for the SEP) compared to extratropical regions. The range extent of a TEP species may therefore span more coastline before it encounters cooler latitudes than it otherwise would if the coastline were oriented more linearly north–south.

At each biogeographic boundary, however, notable gradients in SST coincide with low habitat continuity (Fig. 4). The gradient in SST at the SEP/TEP boundary is a partial result of the convergence of tropical currents with the cold-water Humboldt current, which moves offshore in northern Peru (Montecino & Lange 2009). Similarly, the gradient in SST surrounding the NEP/Cortez boundary occurs at an upwelling transition region and where the southward flowing California current moves offshore (Hewitt 1981; Zaytsev *et al.* 2003; Herrera-Cervantes *et al.* 2013). Thus, as species approach their presumed thermal limits at biogeographic boundaries, they also encounter changes in coastal oceanography and large habitat gaps, which can result in sharp reductions in abundance and range limits (e.g. Fenberg & Rivadeneira

2011; Fenberg *et al.* 2014). The biogeographic extent of the tropics (i.e. the peak(s) of the overall species richness gradient) and adjacent regions may therefore be geographically constrained by the coincidence of low habitat continuity and large gradients in SST. This affects the overall shape of the richness gradient, which we argue is not well assessed or visualised latitudinally (compare Fig. 4c with Fig. S5).

Although rarely discussed outside of classic geomorphological literature (but see Crame 2000; Gray 2002), it is thought that the relative scarcity of nearshore tropical rocky substrate (Fig. 3c) is at least a partial result of spatial and temporal variation in coastal climate, including riverine output, aridity and glacial history (Hayes 1967). Because these climate/geological processes are ongoing at a global scale, the effect of habitat continuity on coastal biogeographic patterns and processes suggested in this study may not be limited to the EP. We therefore suggest that researchers incorporate habitat continuity and geographical variation in dispersal potential into models/assessments of the latitudinal diversity gradient and biogeographic structure. Of particular interest will be to compare and contrast coastal species that primarily inhabit soft sediment vs. rocky substrate given that habitat continuity may have differing effects on their diversification dynamics. In addition, more research should focus on how patterns of habitat continuity within biogeographic regions affects connectivity between populations and whether inter-regional differences in habitat continuity (tropical vs. extratropical) are related to differences in dispersal potential, metapopulation dynamics, and ultimately, macroevolution.

#### ACKNOWLEDGEMENTS

This project would not have been possible without help from the current and former curators and researchers at The Los Angeles County Museum of Natural History, The Santa Barbara Museum of Natural History, The National Museum of Natural History (Smithsonian), The Gulf of California Invertebrate Database (hosted by the Arizona-Sonora Desert Museum), The California Academy of Sciences and The Natural History Museum (London). We are also thankful to the Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO) for providing occurrence data for the northeastern Pacific (especially Pete Raimondi) and for invaluable conversations (and preliminary data collection) about this project with David Reid, David Paz-García, Michael Hellberg, Suzanne Williams, Moira MacLean, Marla Spencer, Stephan Thorson, John Stewart, Jon Todd and Ken Johnson. The research of MMR was funded by FONDECYT projects #1140841, 1150664, 1151094 and 1181300. The research of PBF was funded by a Royal Society Grant RG160700.

#### AUTHORSHIP

PBF and MMR conceived of the idea for the paper, collected data, conducted analyses and wrote the paper.

#### DATA ACCESSIBILITY STATEMENT

Data available from the Dryad Digital Repository <https://doi.org/10.5061/dryad.1m04tv1>.

#### REFERENCES

- Aguilera, M.A., Valdivia, N., Jenkins, S., Navarrete, S.A. & Broitman, B. (2018). Asymmetric competitive effects during species range expansion: an experimental assessment of interaction strength between “equivalent” grazer species in their range overlap. *J. Anim. Ecol.* <https://doi.org/10.1111/1365-2656.12917>
- Archer, E. (2016). rfPermute: estimate permutation p-values for Random Forest importance metrics. R package (Zenodo), Version, 2.
- Belanger, C.L., Jablonski, D., Roy, K., Berke, S.K., Krug, A.Z. & Valentine, J.W. (2012). Global environmental predictors of benthic marine biogeographic structure. *Proc. Natl Acad. Sci. USA*, 109, 14046–14051.
- Breiman, L. (2001). Random forests. *Mach. Learn.*, 45, 5–32.
- Brusca, R.C. (1980). *Common Intertidal Invertebrates of the Gulf of California*. University of Arizona Press, Tucson, AZ.
- Chaudhary, C., Saedi, H. & Costello, M.J. (2016). Bimodality of latitudinal gradients in marine species richness. *Trends Ecol. Evol.*, 31, 670–676.
- Colwell, R.K. & Lees, D.C. (2000). The mid-domain effect: geometric constraints on the geography of species richness. *Trends Ecol. Evol.*, 15, 70–76.
- Crame, J.A. (2000). The nature and origin of taxonomic diversity gradients in marine bivalves. *Geol. Soc. London. Spec. Publ.*, 177, 347–360.
- Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. *et al.* (2007). Random forests for classification in ecology. *Ecology*, 88, 2783–2792.
- De'ath, G. (2005). The mvpart Package. R package. version 1.1.
- De'ath, G. & Fabricius, K.E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81, 3178–3192.
- Edgar, G.J., Alexander, T.J., Lefcheck, J.S., Bates, A.E., Kininmonth, S.J., Thomson, R.J. *et al.* (2017). Abundance and local-scale processes contribute to multi-phyta gradients in global marine diversity. *Sci. Adv.*, 3, e1700419.
- Fenberg, P.B. & Rivadeneira, M.M. (2011). Range limits and geographic patterns of abundance of the rocky intertidal owl limpet, *Lottia gigantea*. *J. Biogeogr.*, 38, 2286–2298.
- Fenberg, P.B., Posbic, K. & Hellberg, M.E. (2014). Historical and recent processes shaping the geographic range of a rocky intertidal gastropod: phylogeography, ecology, and habitat availability. *Ecol. Evol.*, 4, 3244–3255.
- Fenberg, P.B., Menge, B.A., Raimondi, P.T. & Rivadeneira, M.M. (2015). Biogeographic structure of the northeastern Pacific rocky intertidal: the role of upwelling and dispersal to drive patterns. *Ecography*, 38, 83–95.
- Fernandez, M.O. & Marques, A.C. (2017). Diversity of diversities: a response to Chaudhary, Saedi, and Costello. *Trends Ecol. Evol.*, 32, 232–234.
- Field, R., Hawkins, B.A., Cornell, H.V., Currie, D.J., Diniz-Filho, J.A.F., Guégan, J.F. *et al.* (2009). Spatial species-richness gradients across scales: a meta-analysis. *J. Biogeogr.*, 36, 132–147.
- Gaston, K.J. & Fuller, R.A. (2009). The sizes of species' geographic ranges. *J. Appl. Ecol.*, 46, 1–9.
- Gray, J.S. (2002). Species richness of marine soft sediments. *Mar. Ecol. Prog. Ser.*, 244, 285–297.
- Hanski, I. (1998). Metapopulation dynamics. *Nature*, 396, 41–49.
- Hanski, I. (1999). Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. *Oikos*, 87, 209–219.
- Hayes, M.O. (1967). Relationship between coastal climate and bottom sediment type on the inner continental shelf. *Mar. Geol.*, 5, 111–132.
- Herrera-Cervantes, H., Lluch-Cota, S.E., Lluch-Cota, D.B. & Gutiérrez-de-Velasco, G. (2013). Interannual correlations between sea surface temperature and concentration of chlorophyll pigment off Punta Eugenia, Baja California during different remote forcing conditions. *Ocean Sci. Discuss.*, 10, 853–882.
- Hewitt, R. (1981). Eddies and speciation in the California current. *Calif. Coop. Ocean. Fish. Invest. Rep.*, 22, 96–98.

- Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *Am. Nat.*, 163, 192–211.
- Ibanez-Erquiaga, B., Pacheco, A.S., Rivadeneira, M.M. & Tejada, C.L. (2018). Biogeographical zonation of rocky intertidal communities along the coast of Peru (3.5–13.5° S Southeast Pacific). *PLoS ONE*, 13, e0208244.
- Jablonski, D. (2005). Mass extinctions and macroevolution. *Paleobiology*, 31, 192–210.
- Jablonski, D., Belanger, C.L., Berke, S.K., Huang, S., Krug, A.Z., Roy, K. *et al.* (2013). Out of the tropics, but how? Fossils, bridge species, and thermal ranges in the dynamics of the marine latitudinal diversity gradient. *Proc. Natl Acad. Sci. USA*, 110, 10487–10494.
- Jablonski, D., Huang, S., Roy, K., Valentine, J.W. & Bronstein, J.L. (2017). Shaping the latitudinal diversity gradient: new perspectives from a synthesis of paleobiology and biogeography. *Am. Nat.*, 189, 1–12.
- Jocque, M., Field, R., Brendonck, L. & De Meester, L. (2010). Climatic control of dispersal-ecological specialization trade-offs: a metacommunity process at the heart of the latitudinal diversity gradient? *Global Ecol. Biogeogr.*, 19, 244–252.
- Kinlock, N.L., Prowant, L., Herstoff, E.M., Foley, C.M., Akin-Fajiye, M., Bender, N. *et al.* (2018). Explaining global variation in the latitudinal diversity gradient: meta-analysis confirms known patterns and uncovers new ones. *Global Ecol. Biogeogr.*, 27, 125–141.
- Legendre, P. & Borcard, D. (2018). Box-Cox-chord transformations for community composition data prior to beta diversity analysis. *Ecography*, 41, 1–5.
- Legendre, P. & Gallagher, E.D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129, 271–280.
- Liaw, A. & Wiener, M. (2002). Classification and regression by randomForest. *R. News*, 2, 18–22.
- Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G. & Aarninkhof, S. (2018). The State of the World's Beaches. *Sci. Rep.*, 8, 6641.
- MacArthur, R.H. & Wilson, E.O. (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton.
- Marchant, S., Moran, A.L. & Marko, P.B. (2015). Out-of-the tropics or trans-tropical dispersal? The origins of the disjunct distribution of the gooseneck barnacle *Pollicipes elegans*. *Front. Zool.*, 12, 39.
- Marshall, D.J., Krug, P.J., Kupriyanova, E.K., Byrne, M. & Emlet, R.B. (2012). The biogeography of marine invertebrate life histories. *Annu. Rev. Ecol. Evol. Syst.*, 43, 97–114.
- Martínez, G., Arim, M. & Defeo, O. (2017). Distribution of the isopod *Exciroloana braziliensis* on sandy beaches of the Atlantic and Pacific Oceans. *Mar. Ecol. Prog. Ser.*, 583, 137–148.
- Meyers, M.K., Pankey, M.S. & Wares, J.P. (2013). Genealogical approaches to the temporal origins of the Central American gap: speciation and divergence in Pacific *Chthamalus* (Sessilia: Chthamalidae). *Rev. Biol. Trop.*, 61, 75–88.
- Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B. *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.*, 10, 315–331.
- Montecino, V. & Lange, C.B. (2009). The Humboldt Current System: ecosystem components and processes, fisheries, and sediment studies. *Prog. Oceanogr.*, 83, 65–79.
- Naimi, B. (2015). usdm: Uncertainty analysis for species distribution models. R package version, 1.1-15.
- Norris, R.D. & Hull, P.M. (2012). The temporal dimension of marine speciation. *Evol. Ecol.*, 26, 393–415.
- O'Connor, M.I., Bruno, J.F., Gaines, S.D., Halpern, B.S., Lester, S.E., Kinlan, B.P. *et al.* (2007). Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proc. Natl Acad. Sci. USA*, 104, 1266–1271.
- Orzechowski, E.A., Lockwood, R., Byrnes, J.E., Anderson, S.C., Finnegan, S., Finkel, Z.V. *et al.* (2015). Marine extinction risk shaped by trait-environment interactions over 500 million years. *Global Change Biol.*, 21, 3595–3607.
- Pappalardo, P. & Fernández, M. (2014). Mode of larval development as a key factor to explain contrasting effects of temperature on species richness across oceans. *Global Ecol. Biogeogr.*, 23, 12–23.
- Pearson, R.G. & Dawson, T.P. (2005). Long-distance plant dispersal and habitat fragmentation: identifying conservation targets for spatial landscape planning under climate change. *Biol. Conserv.*, 123, 389–401.
- Reid, D. (2002). The genus *Nodilittorina* von Martens, 1897 (Gastropoda: Littorinidae) in the eastern Pacific Ocean, with a discussion of biogeographic provinces of the rocky-shore fauna. *Veliger*, 45, 85–170.
- Rivadeneira, M.M., Alballay, A.H., Villafaña, J.A., Raimondi, P.T., Blanchette, C.A. & Fenberg, P.B. (2015). Geographic patterns of diversification and the latitudinal gradient of richness of rocky intertidal gastropods: the 'into the tropical museum' hypothesis. *Global Ecol. Biogeogr.*, 24, 1149–1158.
- Robertson, D.R. & Cramer, K.L. (2009). Shore fishes and biogeographic subdivisions of the Tropical Eastern Pacific. *Mar. Ecol. Prog. Ser.*, 380, 1–17.
- Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014). Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.*, 12, e1001775.
- Rosenzweig, M.L. (1995). *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, UK.
- Roy, K. & Goldberg, E.E. (2007). Origination, extinction, and dispersal: integrative models for understanding present-day diversity gradients. *Am. Nat.*, 170, S71–S85.
- Sanciango, J.C., Carpenter, K.E., Etnoyer, P.J. & Moretzsohn, F. (2013). Habitat availability and heterogeneity and the Indo-Pacific warm pool as predictors of marine species richness in the tropical Indo-Pacific. *PLoS ONE*, 8, e56245.
- Schluter, D. (2016). Speciation, ecological opportunity, and latitude. *Am. Nat.*, 187, 1–18.
- Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009). Evolution and ecology of species range limits. *Annu. Rev. Ecol. Evol. Syst.*, 40, 415–436.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdana, Z.A., Finlayson, M. *et al.* (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience*, 57, 573–583.
- Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T. & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, 9, 307.
- Tomašových, A., Jablonski, D., Berke, S.K., Krug, A.Z. & Valentine, J.W. (2015). Nonlinear thermal gradients shape broad-scale patterns in geographic range size and can reverse Rapoport's rule. *Global Ecol. Biogeogr.*, 24, 157–167.
- Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. & De Clerck, O. (2012). Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecol. Biogeogr.*, 21, 272–281.
- Valdovinos, C., Navarrete, S.A. & Marquet, P.A. (2003). Mollusk species diversity in the Southeastern Pacific: why are there more species towards the pole? *Ecography*, 26, 139–144.
- Whitton, F.J., Purvis, A., Orme, C.D.L. & Olalla-Tárraga, M.Á. (2012). Understanding global patterns in amphibian geographic range size: does Rapoport rule? *Global Ecol. Biogeogr.*, 21, 179–190.
- Zaytsev, O., Cervantes-Duarte, R., Montante, O. & Gallegos-García, A. (2003). Coastal upwelling activity on the Pacific shelf of the Baja California Peninsula. *J. Oceanogr.*, 59, 489–502.

## SUPPORTING INFORMATION

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

Editor, Sergio Navarrete

Manuscript received 21 December 2018

Manuscript accepted 29 December 2018