Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2016) 25, 951–963

A Journal of Macroecology

Global Ecology and Biogeography

RESEARCH PAPER

Analysing patterns of spatial and niche overlap among species at multiple resolutions

Marcel Cardillo[1]* and Dan L. Warren[1,2]

[1]Macroevolution and Macroecology Group, Research School of Biology, Australian National University, Canberra 0200, Australia, [2]Department of Biology, Macquarie University, Sydney 2109, Australia

ABSTRACT

Aim Analyses of spatial overlap in species distributions are frequently used to test a range of ecological and evolutionary hypotheses, from the role of competition in community assembly to the geography of speciation. Most studies quantify overlap at one spatial resolution. Here we explore the effects of measuring spatial and niche overlap patterns for the same clade (*Banksia*) at multiple resolutions.

Location Australia.

Methods We quantify overlap among species using broad overlap of species range polygons, proximity of occurrence points and co-occurrence within small survey plots. We compare overlap patterns with null models using age–range correlations and the frequency of sympatric sister species. We then use similar methods to examine patterns of overlap in environmental niche dimensions.

Results *Banksia* species show a wide range of overlap values based on range polygons and point proximities, but very low levels of co-occurrence at the local scale. Intercepts of age–range correlations point to higher levels of overlap among recently diverged species than expected. However, comparing the frequency of sympatric sister species with an evolutionary null model supports a prevailing allopatric mode of speciation. In many cases, niche overlap between species exceeds that expected from phylogenetic relatedness or spatial overlap alone.

Main conclusions Patterns of broad geographical overlap among *Banksia* species support a predominantly allopatric mode of speciation, combined with post-speciation range drift. There is more evidence for niche conservatism than for rapid niche divergence among closely related species. This pattern is consistent with broad-scale geomorphic and landscape complexity as a driver of plant speciation in south-west Australia. It is less consistent with finer-scale mechanisms of species divergence such as fire mosaics, or with ecological divergence in sympatry. Analysis of species overlap patterns at different resolutions is a useful approach for revealing the multiple ecological and historical factors that influence species distributions.

Keywords

Age–range correlation, allopatric speciation, co-occurrence, environmental niche, geographical range, sympatric speciation.

*Correspondence: Marcel Cardillo, Macroevolution and Macroecology Group, Research School of Biology, Australian National University, Canberra 0200, Australia. E-mail: marcel.cardillo@anu.edu.au

DOI: 10.1111/geb.12455
http://wileyonlinelibrary.com/journal/geb

INTRODUCTION

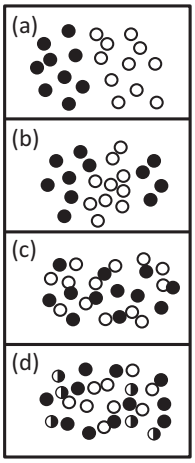
Many comparative studies have sought to test hypotheses about ecological or evolutionary processes by analysing patterns in the spatial overlap of species distributions. In most studies, the spatial resolution at which overlap is measured is predefined by the aims of the study and the processes expected to be at play in determining species distributions. Thus, biogeographical studies that seek to reconstruct the geographical mode of speciation typically measure the degree of overlap in the polygons that define the broad geographical limits of species ranges (e.g. Barraclough & Vogler, 2000; Phillimore *et al.*, 2008), while ecological studies that focus on species interactions define overlap as the degree of co-occurrence of species within small survey plots (e.g. Webb, 2000; Cavender-Bares *et al.*, 2004).

While it is probably broadly true that biogeographical processes are more likely to be reflected in species distributions at broad geographical scales, and ecological interactions at fine scales, Warren *et al.* (2014) argue that there is no real basis for assuming a strict dichotomy of scales at which species overlaps reflect one kind of process or the other. Speciation may occur allopatrically by geographical isolation of populations across a large barrier, but it could also occur if gene flow is disrupted by isolating mechanisms that operate at relatively fine spatial scales, such as fire mosaics for heathland plants (Hopkins & Griffin, 1984; Cowling, 1987). Conversely, ecological interactions can limit co-occurrence of two species within small survey plots, but it can also restrict overlap in species distributions across broad geographical areas (Sexton *et al.*, 2009; Cardillo, 2011; Pigot & Tobias, 2013; Gutierrez *et al.*, 2014). Hence, a sensible approach to analysing species spatial overlap patterns would be to compare overlap at a range of spatial resolutions then ask what processes best explain the overlap patterns, rather than to predefine the scale of overlap together with the expected driving processes.

In this study we present a comparison of the patterns of overlap in species distributions measured at three different spatial resolutions. First, we measure broad overlap in the geographical range polygons that represent the outer limits of species distributions, as is typically done in phylogenetic comparative studies of the geographical mode of speciation. Range polygons are a step removed from the raw data on species distributions, which are typically in the form of occurrence records from a set of sampling points. Our second method is a novel metric that measures overlap between two species using the point occurrence records directly. This method is more sensitive to distributions that might overlap broadly but show allopatry at finer resolutions. For our third method, we follow the approach used in community ecology and measure overlap as the degree of co-occurrence of two species across a set of small ecological survey plots.

Measures of overlap across a set of species are only meaningful when they are compared with appropriate null models. In studies of the geography of speciation, early studies that simply interpreted the relative frequency of allopatric and sympatric sister species were replaced by 'age–range correlations', in which overlaps among species are plotted against their divergence times on a phylogeny (Lynch, 1989; Chesser & Zink, 1994), under the assumption that the degree of overlap immediately following speciation becomes obscured over time by the gradual drift of species range boundaries. Barraclough *et al.* (1998) and Barraclough & Vogler (2000) introduced explicit null models for the expected slope and intercept of age–range correlations under different modes of speciation, while Fitzpatrick & Turelli (2006) presented an age–range correlation method that corrects for the inflated degrees of freedom produced by phylogenetic pseudoreplication across multiple nodes in a phylogeny. Subsequently, Phillimore *et al.* (2008) used simulations of range drift to show that age–range correlations lack power to recover the mode of speciation, and advocated a return to approaches based on the frequency of sympatric sister species. In community ecology, the advent of phylogenetic approaches saw the independent development of a method very similar to age–range correlations (plotting species divergence times against their degree of co-occurrence), even though the aim of this approach is to detect the signature of ecological processes such as competition rather than the mode of speciation (e.g. Slingsby & Verboom, 2006; Cardillo, 2012).

As a case study for our analyses of species overlap patterns, we use the Australian plant genus *Banksia* (Proteaceae). *Banksia* consists of 170 species, about 90% of which are endemic to Australia's Southwest Botanical Province (SWBP), one of the world's mediterranean-climate hotspots of angiosperm diversity. Analyses of spatial overlap among species are potentially important for understanding plant diversity in the SWBP, since this region has seen massive diversification of many plant taxa in a landscape noted for its topographical homogeneity and lack of prominent geographical barriers (Hopper, 1979; Cowling *et al.*, 1996; Hopper & Gioia, 2004). In the SWBP, plant speciation has been attributed to a number of possible mechanisms, including allopatric subdivision by the formation of edaphic barriers (Hopper, 1979; Hopper & Gioia, 2004; Schnitzler *et al.*, 2011), fine-scale geographical subdivision by frequent fires (Hopkins & Griffin, 1984; Cowling, 1987), adaptation and specialization to different soil types (Cowling *et al.*, 1994), shifts in pollination syndromes (Armbruster *et al.*, 1994) and shifts in fire response strategies (Litsios *et al.*, 2013). These different mechanisms of speciation occur at different spatial resolutions, so they may be distinguishable by comparing patterns of species overlap at different resolutions. For example, if broad-scale landscape features serve as the isolating mechanisms that disrupt gene flow and lead to speciation, then the degree of overlap should be low under all three methods (polygons, points and co-occurrences). If the isolating mechanisms operate at finer scales (fire or edaphic mosaics), then overlap of broad-range polygons could be high, but overlap measured using points or co-occurrences could be low. Finally, if the mode of speciation is truly sympatric (pollinator, phenological or fire-response shifts), there may be high levels of overlap under all



Spatial overlap measure			Example of possible interpretations of overlap pattern
Polygon overlap	Point proximity	Local co-occurrence	
low	low	low	Allopatric speciation with broad geographic barrier as isolating mechanism
high	low	low	Allopatric speciation with finer-scale landscape features as isolating mechanism
high	high	low	Sympatric speciation with habitat filtering
high	high	high	Sympatric speciation with ecological trait divergence

Figure 1 A simplified graphical representation and summary of the way spatial overlap between two species can vary with the spatial resolution at which overlap is measured. Occurrences of the two species are represented by black and white circles, with co-occurrences shown as half-filled circles.

three methods. Figure 1 presents a graphical representation and summary of how overlap patterns could vary at different spatial resolutions, and examples of how different patterns might be interpreted.

Superimposed onto the signal of speciation mode are the post-speciation processes of range drift, adaptation to local conditions and ecological sorting that may leave an additional signal in species overlap patterns. Local adaptation and ecological sorting are most likely to be detectable by analysing patterns of niche overlap among species, with respect to divergence times and spatial overlap. For example, sister species that have diverged purely as a result of geographical subdivision may show low spatial overlap but high niche overlap, with the degree of spatial overlap increasing and the degree of niche overlap decreasing among increasingly distant relatives, as the result of avoidance of competition or local adaptation. To explore these kinds of processes we analyse patterns of species overlap in a set of environmental variables, using the same methods as for spatial overlap.

METHODS

Phylogenetic, geographical and environmental data

The *Banksia* phylogeny is from Cardillo & Pratt (2013). These trees consist of 158 of the 170 species of *Banksia*, constructed from 3558 bp of chloroplast DNA using a Bayesian relaxed-clock method in BEAST (Drummond & Rambaut, 2007), and calibrated to a real time-scale using three calibrations (two fossil and one secondary). Our analyses were based on a set of 250 trees from the posterior distribution. We obtained records of all herbarium specimens of *Banksia* (including the former genus *Dryandra*, now merged with *Banksia*) from the Australian Virtual Herbarium (http://chah.gov.au/avh/). We then checked the list of records for each species, discarding obvious duplicates and any records that were clearly outside the species' natural distribution, by comparing records with published distribution maps in monographs of *Banksia* (Collins *et al.*, 2008) and *Dryandra* (Cavanagh & Pieroni, 2006). The final list included 10,276 records (mean records/species = 53.8). Spatial data layers of environmental features were from the Atlas of Living Australia (http://www.ala.org.au/data-sets). To handle the hundreds of layers available from this source in a systematic way, we separated layers into eight categories: precipitation, growth index, radiation, temperature, wind, soil chemistry, soil moisture and productivity. Within each category, we extracted the first two axes of a principal components analysis (PCA) across the set of variables, and used these eight pairs of axes in subsequent analyses.

Quantifying spatial overlap

We quantified spatial overlap between clades using: (1) polygon range maps, (2) distances between point occurrences, and (3) a co-occurrence metric. Polygon range maps were constructed from herbarium record locations using fixed-k convex hulls (Getz & Wilmers, 2004), implemented in the function NNCH in the R package adehabitat. This method generalizes point data to a polygonal range map based on a series of localized convex polygons for each point, rather than requiring that the entire inferred range be convex. This procedure results in less overestimation of species ranges (and hence range overlap between species) than with minimum convex polygons, and gives users some control of the overall level of convexity allowed for range estimates. A setting of $k = 0.5N$, where N is the number of occurrence points for each species, was found to perform well for data sets of all sizes. Overlaps between range polygons were calculated as the area of the intersection of the two species ranges divided by the area of the smaller of the two ranges (Fitzpatrick & Turelli, 2006); this metric ranges from zero (no overlap) to one (100% overlap).

To quantify spatial overlap between species using occurrence point data directly, we present a point-proximity metric, O, that compares conspecific with heterospecific nearest-neighbour distances. Although a number of measures of spatial aggregation in species occurrences have been presented in the past (e.g. Condit *et al.*, 2000; Shen *et al.*, 2013), most of these are focused on aggregation in the distribution of single species, and we are unaware of any metric that does what we require here, namely tests for spatial overlap in the distributions of two species using patterns of co-aggregation in their respective point occurrences. To calculate O, we first excluded multiple records of the same species from the same location. For each occurrence point i of a given species x, we calculate:

$$O(x_i) = \frac{w(x_i)}{b(x_i)}$$

where $w(x_i)$ is the Euclidean distance to the nearest conspecific point and $b(x_i)$ is the Euclidean distance to the nearest heterospecific point. For each species we then calculate:

$$p = \frac{n(O > 1)}{n(O)}$$

where $n(O > 1)$ is the number of O values in which the distance to the nearest conspecific point is greater than the distance to the nearest heterospecific, and $n(O)$ is the total number of O values for the species. The overlap between two species x and y is then:

$$O_{xy} = \frac{p(x) + p(y)}{2}.$$

The value of O is bounded between zero and one, with values close to zero indicating little spatial overlap while a value of $O \approx 0.5$ is expected if the occurrence points of the two species are randomly and independently distributed across the same area. Theoretically, values of O between 0.5 and 1 would also be possible if the nearest occurrence point was more frequently a heterospecific rather than a conspecific (which might be the case, for example, under strong intraspecific competition). In the Appendix S1 in the Supporting Information we present results of simulations that demonstrate the behaviour of O under different levels of overlap, different degrees of asymmetry in the number of occurrence points between two species and different levels of sampling completeness. In general, the value of O becomes less powerful at discriminating different degrees of overlap as the asymmetry in the number of points of the two species increases. When sampling is incomplete, values of O seem to be elevated, particularly when numbers of occurrence points are unequal and more aggregated (Appendix S2).

To measure local-scale co-occurrences between species we use data from a large botanical survey of the wheatbelt region of the SWBP (Gibson *et al.*, 2004). This survey recorded the presence of all plant species within 682 10 m × 10 m plots, across an area of >200,000 km^2. The survey dataset includes 58 of the *Banksia* species found in our phylogeny, within 204 plots. Although the plots are small, levels of co-occurrence can be high, with more than six *Banksia* species recorded from many plots. To quantify the degree of co-occurrence among pairs of species across these 204 plots, we use Schoener's index C_{ij} (Schoener, 1970), calculated as $C_{ij} = 1 = 0.5\sum(p_{in} - p_{hn})$, where p_{in} and p_{hn} are proportion of occurrences of species i and h in the nth plot. C was calculated using the species.dist function in the R library picante (Kembel *et al.*, 2010).

There is necessarily some non-independence between the first two overlap measures (polygons and point-proximity) because they are unavoidably based on the same raw data, namely species occurrence records. However, this does not invalidate the logic of the interpretation of overlaps at different resolutions presented in Fig. 1.

Comparing overlap patterns with null models: age–range correlations

To explore the way spatial overlap and environmental similarity vary with divergence time, we used the phylogenetic averaging method of Fitzpatrick & Turelli (2006), implemented in the R package phyloclim. This method calculates an overlap score at each node in the phylogeny by applying nested averages of the overlaps among all pairs of species across the two clades descending from the node. In a fully bifurcating tree each pairwise overlap is inversely weighted by two raised to the power of the number of nodes separating the two species. In this way, the overlap value at the node is a measure of the average overlap between species after a given time since speciation (Fitzpatrick & Turelli, 2006). We then obtained the slope of the regression of node age against overlap, and compared these slopes with null distributions of 500 slopes. In generating null distributions, we wished to: (1) remove the observed association between species distributions and phylogenetic relationships, and (2) account for uncertainty in the phylogenetic topology and branch lengths. We incorporated both of these considerations into a single null model by drawing 50 phylogenies at regular intervals from the Bayesian posterior set and randomizing the data with respect to the phylogeny, generating 10 iterations of the null model for each tree. For the polygon and point-proximity overlap measures, data randomization was done by shuffling species across the tips of the phylogeny. For the co-occurrence metric, data randomization was done by shuffling the species presence–absence matrix 10,000 times, using the independent-swap algorithm from the R library picante. This method maintains the observed row and column totals of the matrix, apparently the least biased way to randomize species occurrence data (Gotelli & Entsminger, 2003). In addition, we were mindful of the fact that *Banksia* species endemic to the SWBP are widely separated from those found elsewhere in Australia, with apparently very few migration events between the SWBP and elsewhere (Cardillo & Pratt, 2013). We therefore constrained tip shuffling to randomize separately across SWBP species and non-SWBP species in order to

generate null models that were more biologically plausible than randomizing across all tips.

To highlight individual nodes for which the degree of spatial overlap deviates significantly from the null expectation for that node, we calculated standardized effect sizes. These were calculated as (observed overlap – mean null overlap)/standard deviation of null overlap. For each node, calculation of the effect size was based on the distribution of null overlap values for that particular node only.

Comparing overlap patterns with null models: frequency of sympatric sister species

Age–range correlations have the advantage that they make full use of the data across the phylogeny, but they have also been criticized: some authors have argued that species range shifts are too rapid for any but the shallowest divergences to retain the signal of the speciation mode in their patterns of overlap (Berlocher & Feder, 2002; Losos & Glor, 2003; Fitzpatrick & Turelli, 2006; Bolnick & Fitzpatrick, 2007; see Cardillo, 2015, for a counterpoint). Therefore, we also calculate the degree of overlap among pairs of sister species only, and compare these with a process-based null model generated under a modified form of the range evolution simulation presented by Phillimore *et al.* (2008). Phillimore *et al.* (2008) applied a simulation in which the boundaries of two hypothetical ranges were allowed to drift within a bounded domain for a specified time. Our model builds on this approach by evolving ranges along the branches of the *Banksia* phylogeny, with range divisions occurring at branching points. In this way, the sizes, positions and degree of overlap of ranges at the tips of the phylogeny are a function of the size of the ancestral ranges, the time since divergence and the rate of range boundary drift. Full details of the model are presented in Cardillo (2015). Because we do not currently have a method for simulating occurrence points or local-scale occurrences in a similar way, we apply this model only to polygon range overlaps between sister species.

To generate the null distributions, we ran one set of simulations in which the probabilities of any speciation event being allopatric and sympatric were 0.9 and 0.1, respectively, and another set where the probabilities were 0.1 and 0.9, respectively. We considered sister species ranges to be 'sympatric' using two alternative threshold values for the degree of range overlap: 50% and 90%. For each of these four parameter combinations, we ran the simulation 1000 times. The output for each run was the number of sympatric sister species, and this was the test statistic used to compare the observed level of sympatry against each of the four null distributions.

Quantifying niche overlap and comparison with null models

We measured overlap in environmental niche axes between pairs of species using polygon ranges and the point proximity metric *O*. We did not do the niche overlap analyses for the local-scale survey plot data, as we did not have data for the same set of environmental variables. We overlaid polygons onto environmental data layers from the Atlas of Living Australia, and extracted mean values of each PCA axis for each species. To measure environmental overlap using point occurrence data we applied *O* to the set of records of environmental variables most closely associated with the species occurrence records. Instead of calculating *O* from *x*/*y* spatial coordinates, we used Euclidean distances between the two PCA axes representing each of the eight types of variables (temperature, precipitation, etc.). In this way, we obtained a measure of overlap in each niche dimension between two species, based on point occurrence data. For both polygon and point-based niche overlap scores, we tested the slopes and intercepts of niche overlap against divergence times using the Fitzpatrick–Turelli age–range correlation method.

Niche overlap and spatial overlap relative to divergence times

Next, we asked whether sister clades that show high levels of spatial overlap, given their age, tend to be unusually divergent along environmental niche axes. If this is the case, we expect a negative correlation between spatial and environmental overlap scores across nodes in the phylogeny. To test the significance of these correlations, we compared the observed correlation between average overlap values across nodes with a null distribution of correlation coefficients obtained by shuffling the raw data values across the tips of the phylogeny 500 times and recalculating the nodal overlap values each time.

Because environmental variables tend to be spatially autocorrelated, we expect that spatial and environmental nodal overlap scores will be strongly correlated. The randomization test described above shuffles species values with respect to the phylogeny but retains the pairing of spatial and environmental overlap scores, thereby including this correlation in the null distribution. Significance tests using this method therefore address whether or not the observed correlation exceeds that expected due to geography alone, i.e. whether there is phylogenetic signal suggesting more, or less, environmental similarity than expected given the geographical distributions of the constituent species. In this way, a negative correlation would suggest that highly sympatric clades tend to display levels of environmental niche divergence that are higher than expected given the average level of sympatry within the clade. A positive correlation would suggest that highly sympatric clades tend to be unusually similar in environmental preferences. R code to implement the ARC analysis as well as the randomization of C_{ij} values is available from https://github.com/danlwarren/arc-extensions.

RESULTS

Node ages against spatial and environmental niche overlap

The age–range correlations of divergence times against overlap in range polygons among nodes (Fig. 2a) and against

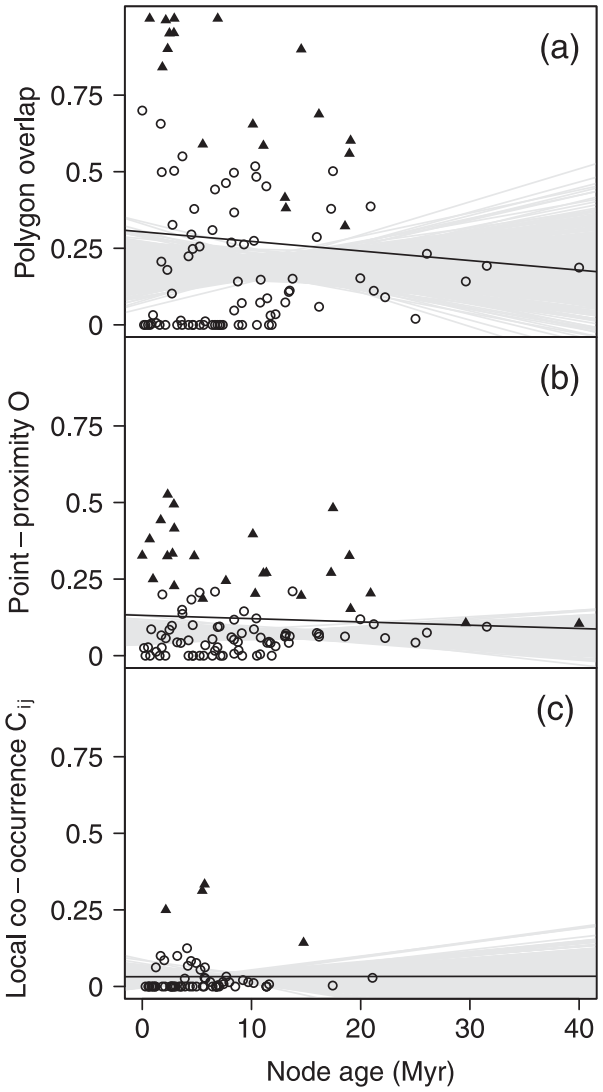


Figure 2 Age–range correlations of node ages and nodal values of spatial overlap, with overlap quantified using (a) range polygon overlaps, (b) the point-proximity metric *O*, and (c) the local co-occurrence metric C_{ij}. The grey lines are the slopes of 500 replicates of the null model (see text for details) and the black lines are the medians of observed slopes from 50 phylogenies from the Bayesian posterior distribution. Points are the nodal values for the maximum clade credibility tree; note that the points are shown simply to illustrate the spread of overlap values and the slopes shown are not from models fitted to these sets of points. Black triangles are the points for which the standardized effect size of the overlap value is ≥ 1.65, indicating a significant departure from the null expectation for that node.

values of *O* (Fig. 2b) both reveal a wide range of overlap values from zero to 100% overlap. Node age has very low explanatory power for overlap (adjusted $R^2 = -0.004$ and -0.006 for the points in Fig. 2a,b, respectively). Note that the overlap score values are not comparable in these two plots: an *O* value around 0.5 is the expectation if both sets of species points are distributed randomly. The slopes of these age–range correlations are no different from those expected under the null model, but the intercepts are in the upper tails of their respective null distributions (Table 1), suggesting that degree of spatial overlap between recently diverged sister species is often higher than would expected by chance. However, when we examine the standardized effect sizes in overlap values it does not appear that overlaps at shallower nodes are any more likely to deviate from the null expectation than those at deeper nodes (the black triangles in Fig. 2). The plot of node age against the level of local-scale co-occurrence (Fig. 2c) reveals very low levels of co-occurrence within ecological survey plots; in fact, across the majority of nodes there is zero co-occurrence. However, the explanatory power

Table 1 Intercepts and slopes of regressions of node ages against spatial and environmental nodal overlap values.

	Range polygons		Point proximity (*O*)	
Variable	Intercept	Slope	Intercept	Slope
Spatial overlap	0.13**	0.03	0.06***	0.01
Temperature	0.18**	0.07	0.08***	0.03
Growth index	0.26	0.17	0.07*	0.05
Soil moisture	0.19***	0.05	0.08**	0.03
Radiation	0.18***	0.06†	0.07***	0.02†
Soil type	0.34	0.29	0.16***	0.10
Wind	0.27***	0.15	0.09***	0.05
Precipitation	0.18***	0.06	0.07***	0.03
Productivity	0.26	0.20	0.13*	0.09

Overlap in spatial and environmental variables is quantified using range polygon overlaps and the point-proximity metric *O*; values shown are the medians of values from 50 phylogenies from the posterior distribution. Significance is indicated with respect to the null model.
†$P \leq 0.1$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.

Table 2 Intercepts and slopes of regressions of spatial overlap against environmental nodal overlap values.

	Range polygons		Point proximity (*O*)	
Variable	Intercept	Slope	Intercept	Slope
Temperature	0.46***	0.22***	0.49***	0.10***
Growth index	0.43***	0.43***	0.39***	0.12***
Soil moisture	0.43†	0.21***	0.50***	0.10***
Radiation	0.40**	0.19***	0.50***	0.08***
Soil type	0.49*	0.44***	0.44**	0.24***
Wind	0.44	0.33***	0.47***	0.13***
Precipitation	0.44*	0.18***	0.48***	0.09***
Productivity	0.46	0.39***	0.37***	0.22***

Overlap in spatial and environmental variables is quantified using range polygon overlaps and the point-proximity metric *O*; values shown are the medians of values from 50 phylogenies from the posterior distribution. Significance is indicated with respect to the null model.
†$P \leq 0.1$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.

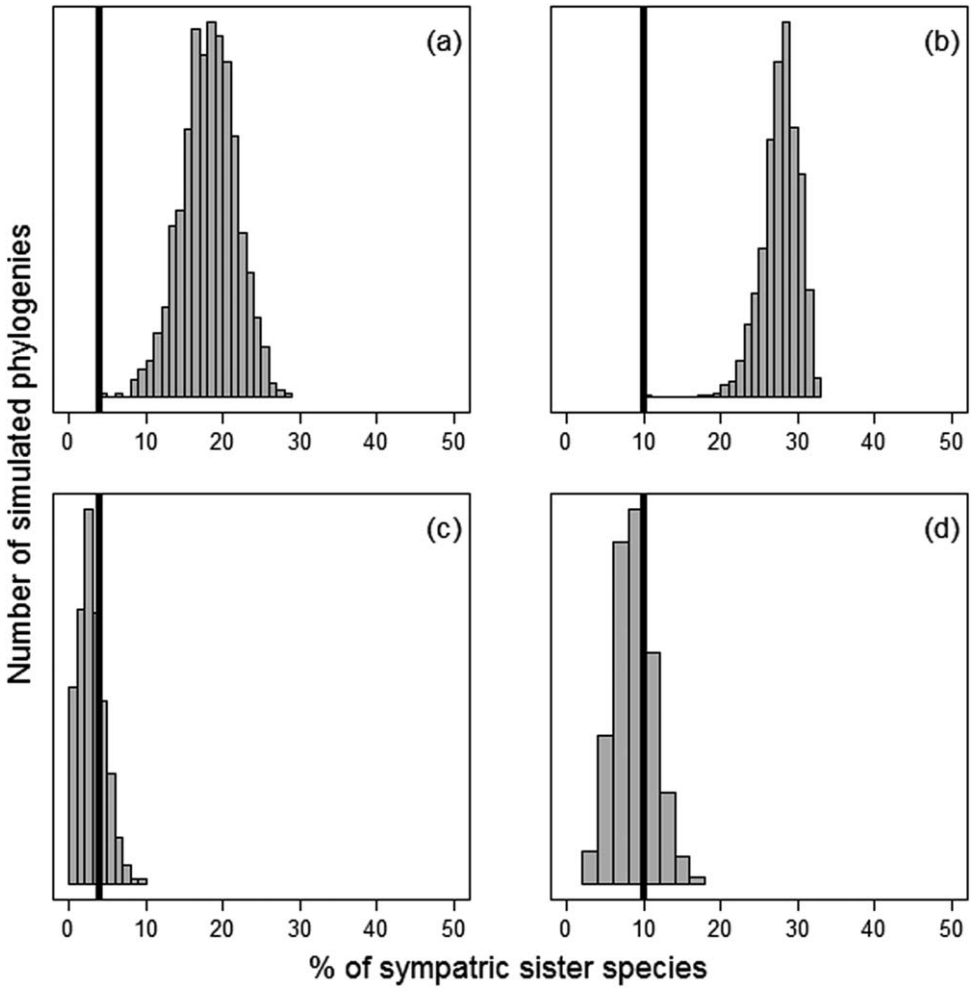


Figure 3 Distributions of 1000 values for the number of sympatric sister species generated under a simulation of geographical range evolution along the *Banksia* phylogeny. Distributions are shown for simulations in which the probability of speciation events being sympatric was 0.9 (a, b) and 0.1 (c, d); and for which sympatry was defined as >90% overlap (a, c) and >50% overlap (b, d). Vertical black lines show the observed numbers of sympatric sister species of *Banksia*.

of this model is also very low (adjusted $R^2 = 0.02$), and neither the slope nor the intercept of this age–range correlation differ from the null expectation.

When the frequency of sympatric sister species is compared with an evolutionary null model, a contrasting picture emerges. Under simulations in which sympatric speciation predominates there are significantly fewer sympatric sister species of *Banksia* than expected ($P < 0.001$; Fig. 3a,b). Under simulations in which allopatric speciation predominates, the number of sympatric sister species is approximately in the middle of the null distribution ($P = 0.66$ and $P = 0.65$; Fig. 3c,d). Hence, the pattern of range overlap among *Banksia* sister species is consistent with predominately allopatric speciation, but not with predominately sympatric speciation.

Overlap in environmental niche axes across nodes (age–niche correlations) also shows a wide range of values, both for polygon-based (Fig. 4) and point-based (Fig. 5) measures of overlap. For all environmental variables, slopes of the age–niche correlations are no different from the null distributions. However, for many variables the intercept is in the upper tail of the null distribution, implying that recently diverged sister species are often more similar in their environmental niche than expected from the null models (Table 1).

Spatial overlap against environmental niche overlap

There is a phylogenetic component to the relationship between geographical and environmental overlap, under both range polygon (Fig. 6) and point-proximity (Fig. 7) methods of measuring overlap (Table 2). It is generally expected that spatial overlap will be positively correlated with niche overlap across phylogenetic nodes. However, the departure of the slope and intercept from the null distribution in this case indicate that the phylogenetic relationship between spatial and niche overlap exceeds that expected based on environmental autocorrelation alone; clades that are sympatric tend to be more similar environmentally than randomly selected non-clade members that are sympatric.

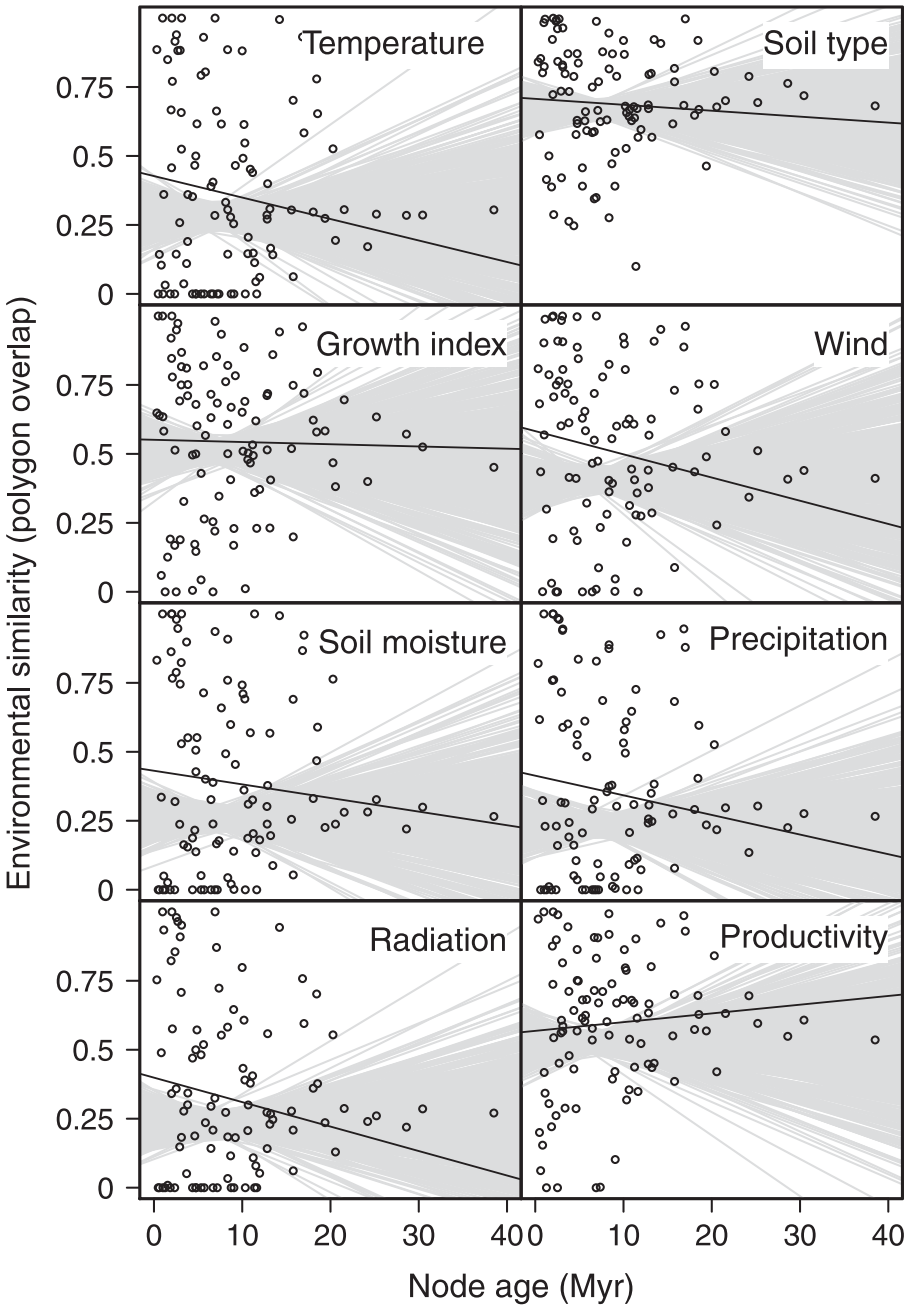


Figure 4 Age–niche correlations of node ages and nodal values of environmental niche similarity, with niche similarity quantified using range polygons. The grey lines are the slopes of 500 replicates of the null model (see text for details) and the black lines are the medians of observed slopes from 50 phylogenies from the Bayesian posterior distribution. Points are the nodal values for the maximum clade credibility tree; note that the points are shown simply to illustrate the spread of overlap values and the slopes shown are not from models fitted to these sets of points.

DISCUSSION

Typically, patterns in the overlap of species distributions are quantified and analysed at either 'regional' or 'local' scales, according to the processes that each study seeks to investigate. Whether such a dichotomy of spatial scales is justified, however, has been questioned (Warren *et al.*, 2014). In principle, patterns of overlap in present-day species distributions should carry the signal of the processes that modify distributions over time (random drift, adaptation or ecological interactions), superimposed onto the historic signal of speciation (Pigot & Tobias, 2013). Hence, analysing patterns of overlap measured at multiple spatial resolutions, together with information on overlap of niche axes and divergence times, should offer a powerful approach to inferring the processes that have shaped species distributions.

We have applied age–range correlations to the analysis of overlap patterns because this approach makes full use of the data on species distributions and divergence times, and because the temporal dimension offers additional ways of comparing the data with the patterns expected under alternative processes (Lynch, 1989; Chesser & Zink, 1994; Barraclough *et al.*, 1998; Barraclough & Vogler, 2000). However, we acknowledge that some authors consider that post-speciation range movement is likely to be too rapid for the signal of speciation to be retained in any but the shallowest divergences, and that approaches based on the frequency of sympatric sister

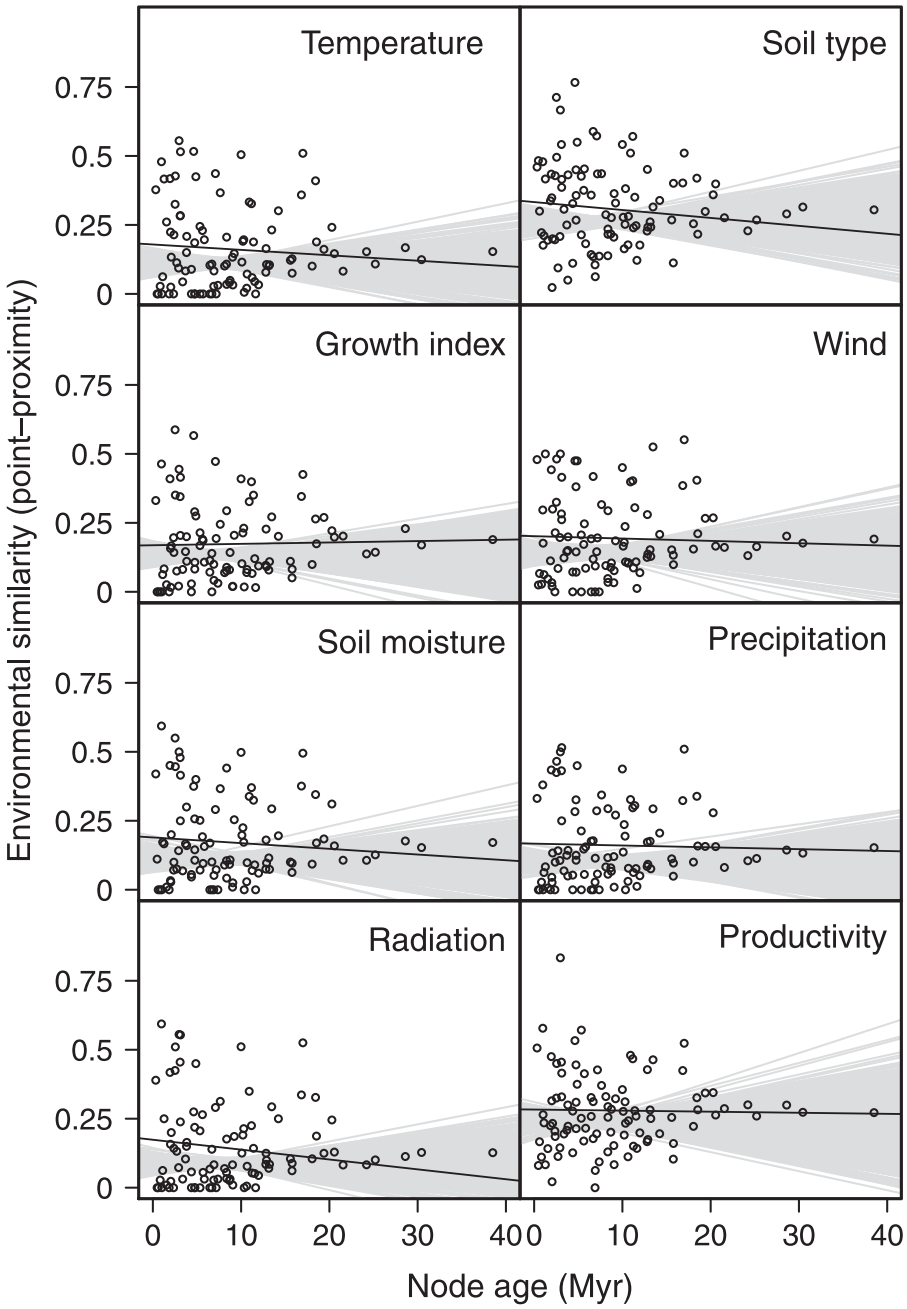


Figure 5 Age–niche correlations of node ages and nodal values of environmental niche similarity, with niche similarity quantified using the point-proximity metric *O*. The grey lines are the slopes of 500 replicates of the null model (see text for details) and the black lines are the medians of observed slopes from 50 phylogenies from the Bayesian posterior distribution. Points are the nodal values for the maximum clade credibility tree; note that the points are shown simply to illustrate the spread of overlap values and the slopes shown are not from models fitted to these sets of points.

species may be more informative (Phillimore *et al.*, 2008). The inference of speciation mode from our data is made difficult by the contrasting results we obtain under these two approaches. A considerable number of shallow nodes in the *Banksia* phylogeny show high levels of spatial overlap, and by comparison with a null model, the significantly high intercept of the age–range correlation implies a greater tendency to a sympatric mode of speciation than expected. A very similar pattern in plants of another mediterranean climate hotspot (California) was recently interpreted in this way (Anacker & Strauss, 2014). Nonetheless, we are cautious about drawing a conclusion of prevalent sympatric speciation from this result because the lack of a clear linear relationship between node ages and degree of spatial overlap (Fig. 2a,b) means that the *y*-intercepts of these plots do not have strong predictive value for the degree of overlap at a node age of zero.

In fact, the opposite conclusion about the prevailing mode of speciation may be drawn from the comparison of the frequency of sympatric sister species with the expectation under our simulation of range evolution along the *Banksia* phylogeny. Under this model, the number of sympatric sister species (whether sympatry is defined conservatively or liberally) is consistent with a process of range evolution in which allopatric speciation predominates, with range boundaries permitted to drift randomly in between speciation events. Two other results also point towards a predominantly allopatric mode of speciation. The first is the large number of shallow nodes with a high degree of overlap in environmental niche

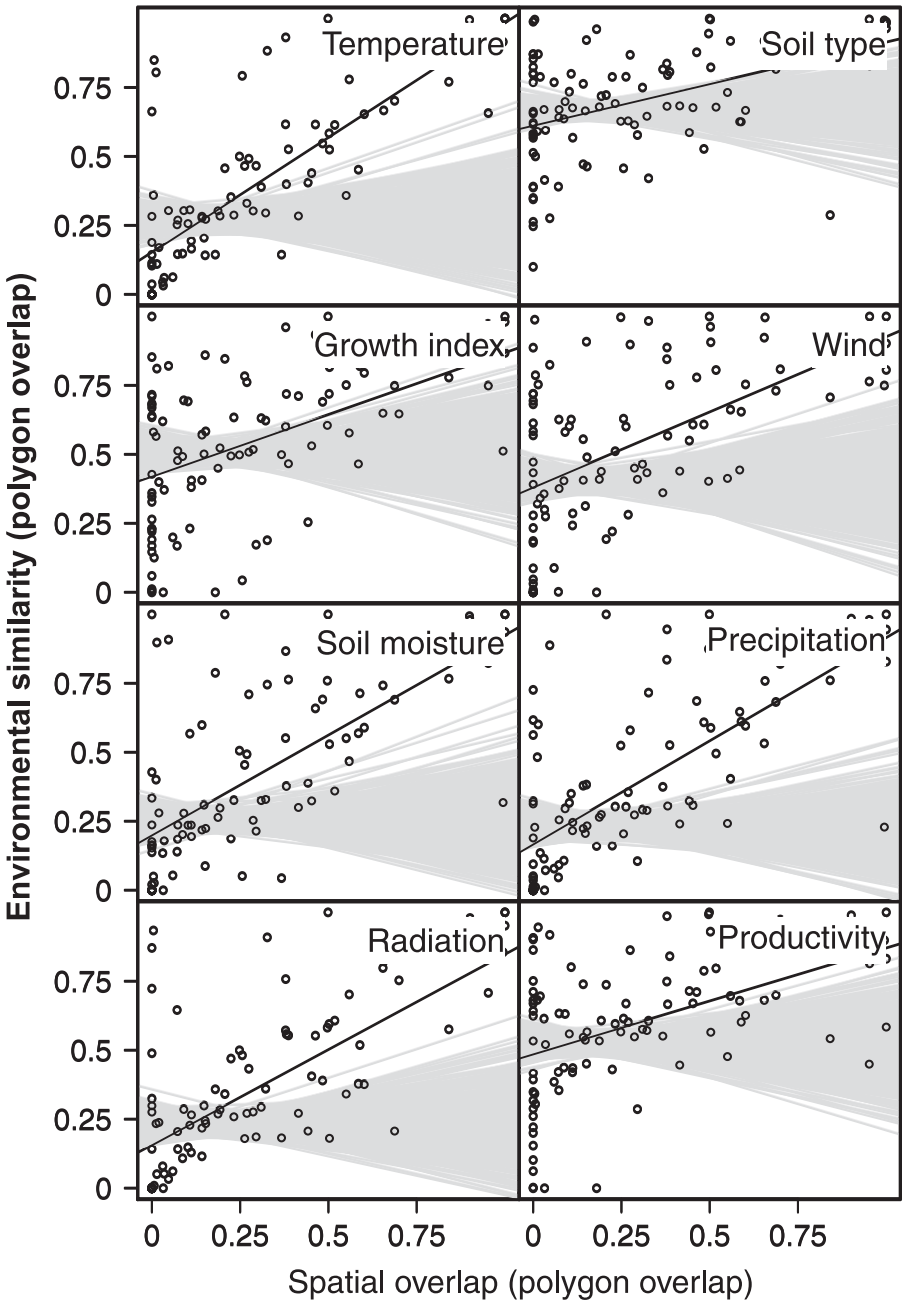


Figure 6 Correlations between spatial overlap and environmental niche similarity across phylogenetic nodes, both quantified using range polygons. The grey lines are the slopes of 500 replicates of the null model (see text for details) and the black lines are the medians of observed slopes from 50 phylogenies from the Bayesian posterior distribution. Points are the nodal values for the maximum clade credibility tree; note that the points are shown simply to illustrate the spread of overlap values and the slopes shown are not from models fitted to these sets of points.

dimensions (Figs 4 & 5), and the significantly high *y*-intercepts of these age–niche correlations (although the above caveat also applies to these intercepts). The second result is the positive relationship between spatial and niche overlap across nodes, in excess of that expected based on spatial autocorrelation in the environment (Fig. 6). Both of these results are consistent with speciation events that result from geographical isolation rather than niche divergence between close relatives or between highly sympatric sister species. For these reasons, we believe that our data more strongly support a mode of speciation in *Banksia* based on geographical subdivision rather than ecological divergence in sympatry. Indeed, these results are suggestive of environmental niche conservatism in *Banksia*. The fact that sister species seem to be more similar in their environmental affinities than their relatedness would suggest, and that sympatric species seem to be more similar than their geographic overlap would suggest, is consistent with a view of niche conservatism as the result of stabilizing selection as opposed to simple phylogenetic inertia (Losos, 2008; Wiens *et al.*, 2010).

What can we infer from patterns of overlap measured using the point-proximity metric O and the local co-occurrence metric C_{ij}, that is not apparent from patterns of overlap in range polygons alone? To a large extent, the

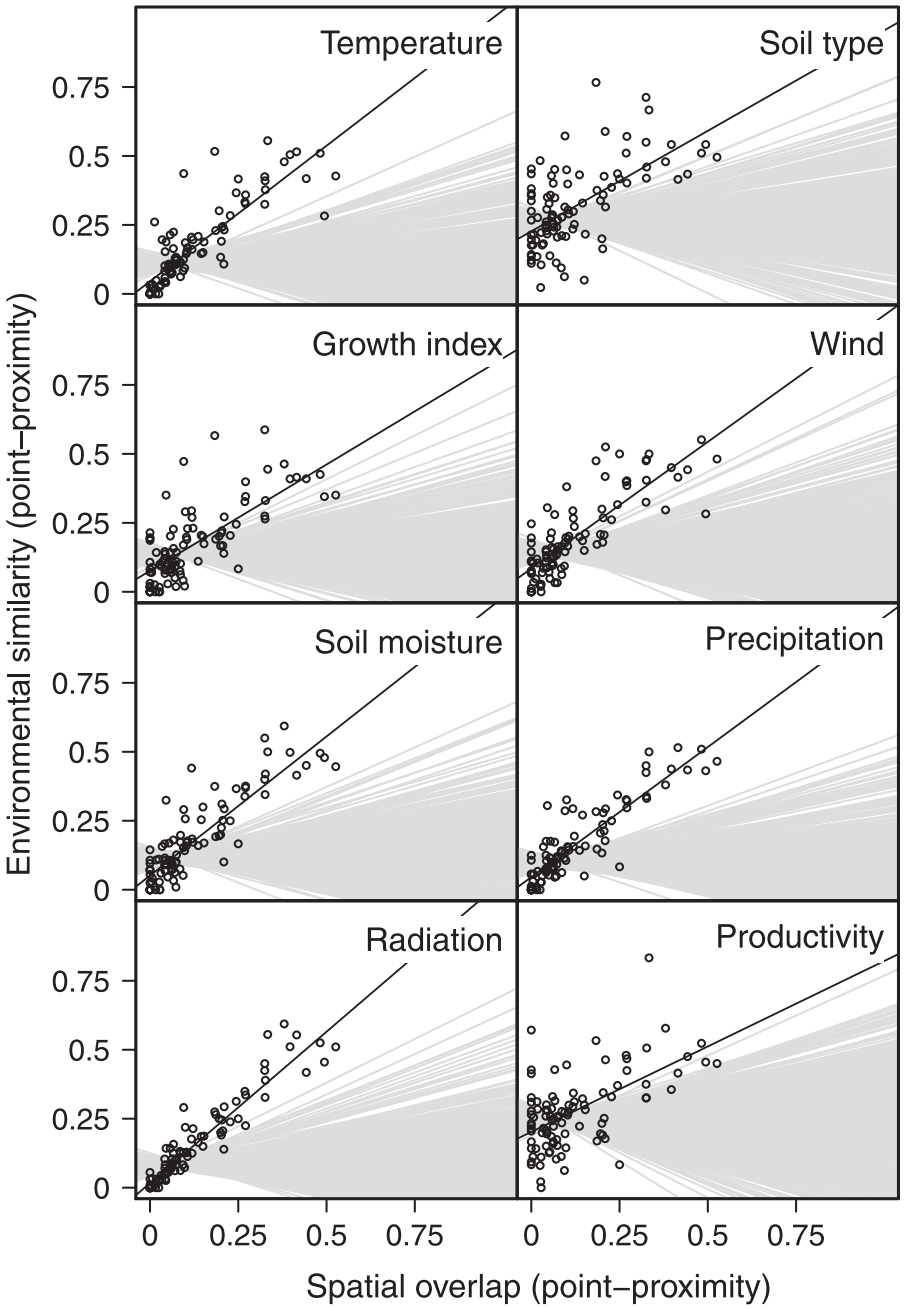


Figure 7 Correlations between spatial overlap and environmental niche similarity across phylogenetic nodes, both quantified using the point-proximity metric *O*. The grey lines are the slopes of 500 replicates of the null model (see text for details) and the black lines are the medians of observed slopes from 50 phylogenies from the Bayesian posterior distribution. Points are the nodal values for the maximum clade credibility tree; note that the points are shown simply to illustrate the spread of overlap values and the slopes shown are not from models fitted to these sets of points.

patterns for point-proximity overlap mirror those for polygons, perhaps because the two overlap measures are based on the same raw data. This is also partly the result of the nestedness of the overlap metrics: two species with low polygon overlap must also have low point-proximity overlap. However, the converse is not true, and the fact that we do not see a large number of nodes with high polygon overlap but low point-proximity overlap might be informative. This could suggest that very fine-scale geographical isolating mechanisms, such as fire mosaics, are not a primary driving force of speciation in *Banksia*. In fact, there is direct evidence that spatial patterns of reproductive isolation in this genus do not always correspond with fine-scale landscape structure or the geographical isolation of populations, probably because occasional long-distance seed dispersal maintains gene flow among populations separated by distances of up to several kilometres (He *et al.*, 2010; Llorens *et al.*, 2012; Merwin *et al.*, 2012).

In a similar way, the low levels of co-occurrence among species at the local scale will be driven, at least partly, by the fact that species that are geographically separated at a broad scale cannot co-occur within small plots. This could, potentially, be overcome by restricting the calculation of C_{ij} to pairs of species that show some range polygon overlap (e.g. Cardillo, 2012), but in the ecological survey dataset we use this would leave very few species pairs for analysis. With

this limitation in mind, there are several ways we can interpret the co-occurrence age–range correlation. A strict interpretation of the slope and intercept in comparison with the independent-swap null model (Fig. 2c) suggests that there is no phylogenetic signal in the factors that determine species co-occurrences within plots (a null slope), and that the average level of co-occurrence at the time of divergence is no different from that expected by chance (a null intercept). On the other hand, we can note that the level of co-occurrence is zero at the shallowest nodes (Fig. 2c), which argues against a purely ecological mode of speciation, such as shifts in flowering time or pollination mode, that would permit sister species to co-occur within small plots.

In conclusion, the results of our analyses of spatial and niche overlap patterns point to a predominantly allopatric mode of speciation in *Banksia*, without any strong evidence for divergence in environmental niche axes among close relatives or highly sympatric species. We also find little evidence that species overlap broadly but are more spatially segregated at finer resolutions. Together, our results are more consistent with a scenario in which species divergences are driven largely by broad-scale geomorphic features or landscape dissection, as suggested by Hopper (1979) and Hopper & Gioia (2004) for south-western Australia. Our results are less consistent with the hypothesis that a finer-scale mechanism such as fire promotes species divergence by disrupting gene flow among populations across relatively small distances.

ACKNOWLEDGEMENTS

This work was funded by an Australian Research Council Discovery grant (DP110103168). The authors would like to acknowledge Michael Turelli, Ben Fitzpatrick and Christoph Heibl for providing valuable feedback on the analyses developed here.

REFERENCES

Anacker, B.L. & Strauss, S.Y. (2014) The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proceedings of the Royal Society B: Biological Sciences*, **281, 20132980**.

Armbruster, W.S., Edwards, M.E. & Debevec, E.M. (1994) Floral character displacement generates assemblage structure of Western Australian triggerplants (*Stylidium*). *Ecology*, **75**, 315–329.

Barraclough, T.G. & Vogler, A.P. (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist*, **155**, 419–434.

Barraclough, T.G., Vogler, A.P. & Harvey, P.H. (1998) Revealing the factors that promote speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **353**, 241–249.

Berlocher, S.H. & Feder, J.L. (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Reviews of Entomology*, **47**, 773–815.

Bolnick, D.I. & Fitzpatrick, B.M. (2007) Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 459–487.

Cardillo, M. (2011) Phylogenetic structure of mammal assemblages at large geographical scales: linking phylogenetic community ecology with macroecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2545–2553

Cardillo, M. (2012) The phylogenetic signal of species co-occurrence in high-diversity shrublands: different patterns for fire-killed and fire-resistant species. *BMC Ecology*, **12**, 21.

Cardillo, M. (2015) Geographic range shifts do not erase the historic signal of speciation in mammals. *The American Naturalist*, **185**, 343–353.

Cardillo, M. & Pratt, R. (2013) Evolution of a hotspot genus: geographic variation in speciation and extinction rates in *Banksia* (Proteaceae). *BMC Evolutionary Biology*, **13**, 155.

Cavanagh, T. & Pieroni, M. (2006) *The Dryandras*. Australian Plants Society Inc., Hawthorn, VIC.

Cavender-Bares, J., Ackerly, D.D., Baum, D.A. & Bazzaz, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist*, **163**, 823–843.

Chesser, R.T. & Zink R.M. (1994) Modes of speciation in birds: a test of Lynch's method. *Evolution*, **48**, 490–497.

Collins, K., Collins, K. & George, A. (2008) *Banksias*. Bloomings Books, Melbourne.

Condit, R., Ashton, P.S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S.P., Foster, R.B., Itoh, A., LaFrankie, J.V., Lee, H.S., Losos, E., Manokaran, N., Sukumar, R. & Yamakura, T. (2000) Spatial patterns in the distribution of tropical tree species. *Science*, **288**, 1414–1418.

Cowling, R.M. (1987) Fire and its role in coexistence and speciation in Gondwanan shrublands. *South African Journal of Science*, **83**, 106–112.

Cowling, R.M., Witkowski, E.T.F., Milewski, A.V. & Newbey, K.R. (1994) Taxonomic, edaphic and biological aspects of narrow plant endemism on matched sites in mediterranean South Africa and Australia. *Journal of Biogeography*, **21**, 651–664.

Cowling, R.M., Rundel, P.W., Lamont, B.B., Arroyo, M.K. & Arianoutsou, M. (1996) Plant diversity in mediterranean-climate regions. *Trends in Ecology and Evolution*, **11**, 362–366.

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

Fitzpatrick, B.M. & Turelli, M. (2006) The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution*, **60**, 601–615.

Getz, W.M. & Wilmers, C.C. (2004) A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*, **27**, 489–505.

Gibson, N., Keighery, G.J., Lyons, M.N. & Webb, A. (2004) Terrestrial flora and vegetation of the Western Australian

wheatbelt. *Records of the Western Australian Museum (Supplement)*, **67**, 139–189.

Gotelli, N.J. & Entsminger, G.L. (2003) Swap algorithms in null model analysis. *Ecology*, **84**, 532–535.

Gutierrez, E.E., Boria, R.A. & Anderson, R.P. (2014) Can biotic interactions cause allopatry? Niche models, competition, and distributions of South American mouse opossums. *Ecography*, **37**, 1–13.

He, T., Lamont, B.B., Krauss, S.L. & Enright, N.J. (2010) Genetic connectivity and inter-population seed dispersal of *Banksia hookeriana* at the landscape scale. *Annals of Botany*, **106**, 457–466

Hopkins, A.J.M. & Griffin, E.A. (1984) Floristic patterns. *Kwongan: plant life of the sandplain* (ed. by J.S. Pate and J.S. Beard), pp. 69–83. University of Western Australia Press, Nedlands.

Hopper, S.D. (1979) Biogeographical aspects of speciation in the southwest Australian flora. *Annual Review of Ecology and Systematics*, **10**, 399–422.

Hopper, S.D. & Gioia, P. (2004) The Southwest Australian Floristic Region: evolution and conservation of a global hot spot of biodiversity. *Annual Review of Ecology and Systematics*, **35**, 623–650.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463–1464.

Litsios, G., Wüest, R.O., Kostikova, A., Forest, F., Lexer, C., Linder, H.P., Pearman, P.B., Zimmermann, N.E. & Salamin, N. (2013) Effects of a fire response trait on diversification in replicated radiations. *Evolution*, **68**, 453–465.

Llorens, T.M., Byrne, M., Yates, C.J., Nistelberger, H.M. & Coates, D.J. (2012) Evaluating the influence of different aspects of habitat fragmentation on mating patterns and pollen dispersal in the bird-pollinated *Banksia sphaerocarpa* var. *caesia*. *Molecular Ecology*, **21**, 314–328

Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, **11**, 995–1003

Losos, J.B. & Glor, R.E. (2003) Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, **18**, 220–227.

Lynch, J.D. (1989) The gauge of speciation: on the frequencies of modes of speciation. *Speciation and its consequences* (ed. by D. Otte and J.A. Endler), pp. 527–556. Sinauer Associates, Sunderland, MA.

Merwin, L., He, T., Lamont, B.B., Enright, N.J. & Krauss, S.L. (2012) Low rate of between-population seed dispersal restricts genetic connectivity and metapopulation dynamics in a clonal shrub. *PLoS One*, **7**(11), e50974.

Phillimore, A.B., Orme, C.D.L., Thomas, G.H., Blackburn, T.M., Bennett, P.M., Gaston, K.J. & Owens, I.P.F. (2008) Sympatric speciation in birds is rare: insights from range data and simulations. *The American Naturalist*, **171**, 646–657.

Pigot, A.L. & Tobias, J.A. (2013) Species interactions constrain geographic range expansion over evolutionary time. *Ecology Letters*, **16**, 330–338.

Schnitzler, J., Barraclough, T.G., Boatwright, J.S., Goldblatt, P., Manning, J.C., Powell, M.P., Rebelo, T.G. & Savolainen, V (2011) Causes of plant diversification in the Cape biodiversity hotspot of South Africa. *Systematic Biology*, **60**, 343–357.

Schoener, T.W. (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, **51**, 408

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415–436.

Shen, G., Wiegand, T., Mi, X. & He, F. (2013) Quantifying spatial phylogenetic structures of fully stem-mapped plant communities. *Methods in Ecology and Evolution*, **4**, 1132–1141.

Slingsby, J.A. & Verboom, G.A. (2006) Phylogenetic relatedness limits co-occurrence at fine spatial scales: evidence from the schoenoid sedges (Cyperaceae: Schoeneae) of the Cape Floristic Region, South Africa. *The American Naturalist*, **168**, 14–27.

Warren, D.L., Cardillo, M., Rosauer, D.F. & Bolnick, D.I. (2014) Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology and Evolution*, **29**, 572–580

Webb, C.O. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist*, **156**, 145–155

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site:

Appendix S1 Illustration of different levels of species overlap generated for tests of the point-proximity metric *O*.

Appendix S2 Simulation of values of the point-proximity metric *O*.

BIOSKETCHES

Marcel Cardillo and **Dan Warren** have research interests in spatial and phylogenetic dimensions of biodiversity, spanning a variety of taxa, regions, spatial scales and methodological approaches. Some current topics of interest include plant diversity hotspots, geographic distributions of mammals and fish brains.

Editor: Richard Field