

Environmental heterogeneity explains contrasting plant species richness between the South African Cape and southwestern Australia

Running title: Heterogeneity and species richness

Ruan van Mazijk, Michael D. Cramer and G. Anthony Verboom

Department of Biological Sciences, University of Cape Town, Rondebosch, South Africa

Corresponding author: RvM, ruanvmazijk@gmail.com

Acknowledgements

RvM is grateful to the National Research Foundation and the South African Association of Botanists for bursaries during the course of this work.

Abstract

Aim: Given the importance of environmental heterogeneity as a driver of species richness through its effects on species diversification and coexistence, we aimed to account for the dramatic difference in species richness per unit area between two similar mediterranean-type biodiversity hotspots and whether this difference is explained by differences in environmental heterogeneity.

Location: The Greater Cape Floristic Region, South Africa (GCFR) and Southwest Australian Floristic Region (SWAFR).

Taxon: Vascular plants (tracheophytes).

Methods: Comparable, geospatially explicit environmental and species occurrence data were obtained for both regions and used to generate environmental heterogeneity and species richness raster layers. Heterogeneity in multiple environmental variables and species richness per unit area were compared between the two regions at a range of spatial scales. At each scale, richness was also regressed against these individual axes and against a major axis of heterogeneity, derived by principal component analysis (PCA).

Results: The GCFR is generally more environmentally heterogeneous and species-rich than the SWAFR. Species richness per unit area is significantly related to the major axis of heterogeneity across both regions, the latter describing ca. 38-50% of overall heterogeneity, the slope of this relationship differing between the two regions only at the finest spatial scale. Multivariate regressions, and regressions against the first axes of the PCAs (PC1), revealed variations in the dependence of species richness on environmental heterogeneity between the two regions.

Main conclusions: Notwithstanding some region-specific effects, we present evidence of a common positive relationship between floristic richness and environmental heterogeneity across the GCFR and SWAFR. This is dependent on spatial scale, being strongest at the coarsest level

of sampling. The generally greater richness per unit area of the GCFR compared to the SWAFR is thus explained by the former's generally greater environmental heterogeneity and is concordant with its greater levels of floristic turnover.

Keywords: biodiversity, environmental heterogeneity, fynbos, Greater Cape Floristic Region, kwongan, macroecology, species richness, species turnover, vascular plants, Southwest Australian Floristic Region

1: Introduction

The species richness of a region is a function of its biogeographic context (e.g. proximity to potential immigrant sources), its diversification history and locally-deterministic, environmental features (e.g. productivity, heterogeneity) that influence species persistence and coexistence (Bøhn & Amundsen, 2004; Ricklefs, 1987, 2004). Potentially influencing these effects, environmental heterogeneity may be an important driver of regional species richness variation, with physically heterogeneous regions tending to be more species-rich (Kreft & Jetz, 2007; Laliberte et al., 2014; Stein, Gerstner & Kreft, 2014). As the recruitment success of immigrant lineages into a region is often dictated by their pre-adaptations (Ackerly, 2009; Crisp et al., 2009; Donoghue, 2008), a physically heterogeneous environment may promote diversity by admitting a more functionally diverse array of lineages. Environmental heterogeneity is also a critical requirement for speciation under most models (Sobel, Chen, Watt, & Schemske, 2010; Wiens, 2004a,b), promoting adaptive divergence and/or population isolation. Likewise, in the context of long-term environmental change, physical heterogeneity may offer refugia to a wider array of lineages, buffering against lineage extinction (Byrne, 2008). Finally, environmental heterogeneity has been shown to facilitate species coexistence at a variety of scales, enhancing regional species richness (Hart et al., 2017). Differences in heterogeneity may, therefore, be critically important in accounting for variation in regional richness, particularly when comparing regions with similar areas, physical properties and timeframes within which their biotas have assembled.

The floristically rich South Western Australian Floristic Region (SWAFR; Hopper & Gioia, 2004) and Greater Cape Floristic Region of South Africa (GCFR; Born, Linder, & Desmet, 2007) constitute a case in point. Situated on the southwestern corners of their respective continents, the climates of both these regions have been oceanically-moderated at least since the Cretaceous, and both are dominated by a contemporary mediterranean-type climate whose origin can be traced to the Early-Middle (SWAFR: Rundel et al., 2016; Lamont & He, 2017) or Late Miocene (GCFR: Dupont, Linder, Rommerskirchen, & Schefuss, 2011; Hoffmann, Verboom, & Cotterill, 2015). Additionally, both regions have been unglaciated since the Permian and are dominated by ancient, weathered landscapes whose soil-nutritional status is amongst the lowest of any landscape on Earth (Stock & Verboom, 2012), hence their designation as old, climatically-buffered infertile landscapes (OCBILs; Hopper, 2009). Owing to these environmental similarities, the SWAFR and GCFR floras are very similar with respect to their plant functional trait spectra (Cowling, Witkowski, Milewski, & Newbey, 1994), although the presence of a significant tree component in the SWAFR underpins a striking difference in vegetation physiognomy (Milewski, 1981; Beard, Chapman, & Gioia, 2000). Moreover, the long-term climatic and geological stability of the two regions ensures that the native floras of both reflect long histories of assembly, extending back to the Eocene and Palaeocene (Hopper, 1979; Verboom et al., 2014). There is also evidence of a long history of transoceanic dispersal between them (Bergh & Linder, 2009). In this context, it is unsurprising that the two floras show

strong taxonomic affinities and are both species-rich with high levels of regional endemism (Ackerly, 2009; Beard et al., 2000; Cowling, Rundel, Lamont, Arroyo, & Arianoutsou, 1996; Gioia & Hopper, 2017).

Notwithstanding these similarities, the SWAFR and GCFR differ markedly in terms of their vascular plant species richness, particularly when considered in relation to geographical area. Where the SWAFR accommodates ca. 7,380 species in an area of ca. 302,600 km² (Hopper & Gioia, 2004), the GCFR houses ca. 11,430 species in an area of ca. 189,700 km² (Snijman, 2013). The difference is even more dramatic when the SWAFR is compared to the hyper-diverse core Cape Floristic Region (CFR; Goldblatt, 1978), the latter accommodating ca. 9,400 species in an area of ca. 90,800 km², and so having roughly 25% more species per unit area than the SWAFR (Cowling et al., 2015). One possible explanation for this striking species richness difference relates to differences in the physical heterogeneity of the two regions. Where much of the GCFR, particularly the CFR, is rugged and mountainous, the SWAFR landscape is much more subdued, comprising an ancient, weathered plateau. Indeed, of the world's five mediterranean-climate regions, the CFR has the second highest median topographic heterogeneity, being surpassed only by the Mediterranean (Bradshaw & Cowling, 2014). Crucially, since the strong relief of the GCFR underlies steep climatic and edaphic gradients (Bradshaw & Cowling, 2014; Jiménez & Ricklefs, 2014), it is probable that the GCFR eclipses the SWAFR both in terms of the magnitude of climatic and edaphic heterogeneity and the spatial scale at which this heterogeneity is expressed, the comparative uniformity of the SW Australian landscape probably underpinning a fairly coarse spatial grain of environmental heterogeneity.

The central aim of this paper, then, is to test the hypothesis that the observed difference in species richness (per unit area) between the SWAFR and GCFR is a consequence of differences in the physical heterogeneity of these regions. We investigate this across a range of spatial scales, as we expect this to affect the species richness patterns we observe (Hart et al., 2017). Focusing on the quarter-degree square (QDS), half-degree square (HDS) and degree square (DS) scales (*sensu* Larsen, Holmern, Prager, Maliti, & Røskoft, 2009), we first compare the distribution of species richness between the two regions, and in each region decompose broader-scale richness into average finer-scale richness and between-square turnover. Thereafter, we compare environmental heterogeneity between the two regions at a range of spatial scales. Finally, we use linear models to assess whether differences in environmental heterogeneity are sufficient to explain observed differences in species richness between the two regions. Although the SWAFR has most often been compared to the CFR rather than the GCFR (e.g. Cowling et al., 1996, 2015; Beard et al., 2000; Rundel et al., 2016), we elected to use the latter as a basis for comparison both because it is floristically more coherent, at least in terms of endemism (Born et al., 2017), and because it is more extensive and comparable in size to the SWAFR. The latter is important to ensure that our models relating species richness to environmental heterogeneity are not disproportionately influenced by data from the SWAFR, and to provide enough data points to enable a meaningful assessment of scale effects in the Cape.

2: Materials and methods

2.1: Comparing species richness

To compare vascular plant species richness between the GCFR and SWAFR, geospatially explicit occurrence records of tracheophytes from within the borders of each region were obtained from the Global Biodiversity Information Facility (GBIF; Table S1). Occurrence data

were cleaned using the “taxize” package (Chamberlain et al., 2016) for R (R Core Team, 2019) that was also used for all other analyses (see Supporting Information). Despite spatial variability in collection effort in both regions, we used raw species counts to estimate QDS-scale species richness on the basis that the application of rarefaction techniques severely distorts known richness patterns when applied to the South African flora (Cramer & Verboom, 2016). The final numbers of unique species thus identified as occurring in the GCFR and SWAFR, respectively, were 9,419 and 6,696. To ensure that species richness was compared and decomposed (see below) across equally sized area units, we included only squares comprising four constituent sub-squares (e.g. four QDS in an HDS). While this resulted in the loss of several coastal squares, which is unfortunate because the coastal floras of both the GCFR and SWAFR are rich in endemic taxa, it is not expected to introduce any systematic biases. Overall, we retained 362 of ca. 449 QDS in the GCFR and 624 of ca. 737 in the SWAFR (ca. 81% and 85% sampling, respectively).

The cleaned species occurrence record data were collated into QDS, HDS and DS (sensu Larsen et al., 2009). In addition, following the additive decomposition (Veech et al., 2002) of Whittaker’s (1960) γ -diversity, we decomposed the species richness of each HDS (S_{HDS}) and DS (S_{DS}) into its α (“plot” richness) and β (turnover) components, using the equations

$$S_{\text{HDS}} = \bar{S}_{\text{QDS}} + T_{\text{QDS}}$$

$$S_{\text{DS}} = \bar{S}_{\text{HDS}} + T_{\text{HDS}}$$

where \bar{S}_{QDS} and \bar{S}_{HDS} are the average species richness of the four constituent squares in each HDS and DS, respectively, and T_{QDS} and T_{HDS} represent the residual (i.e. turnover-based) β richness, determined as $\gamma - \alpha$. These data were compared between the GCFR and SWAFR using common language effect sizes (*CLES*) using the R package “canprot” (Dick, 2017). The *CLES* is the proportion of all pairwise comparisons between two categories in which the value in one category exceeds that in the other. Additionally, we tested for differences in regional α -values and β/γ -proportions using two-sided Mann-Whitney *U*-tests in R, as these data were non-normally distributed.

2.2: Comparing environmental heterogeneity

To compare environmental heterogeneity between the GCFR and SWAFR, we acquired a suite of nine geospatially-explicit environmental variables (Table S1) in the form of raster layers to represent topographic (elevation), climatic (surface temperature, *T*; mean annual precipitation, MAP; precipitation in the driest quarter, PDQ), edaphic (clay content; soil carbon, *C*; pH; cation exchange capacity, CEC) and vegetational gradients (normalized difference vegetation index, NDVI). Wherever possible, we made use of remote sensing derived layers that are comparable between the two regions. As far as possible (see Supporting Information), these variables were selected to represent environmental axes which are considered regionally important and independent (Figure S1–3). Soil variables were summarized as depth-interval weighted averages and climatic and spectral variables as annual means using the “raster” package for R (Hijmans, 2016). All layers were then projected to a common coordinate reference system (WGS84) using the “rgdal” package (Bivand et al., 2017) and resampled bilinearly to 0.05° resolution.

In order to quantify heterogeneity in these environmental variables, we developed an index that would account for the spatial configuration of environmental conditions. Making use of raster data, this employs nested squares at various spatial scales (see Supporting Information). We quantified the environmental heterogeneity of a given square (i.e. 0.10°×0.10°, QDS-, HDS- and

DS-scale) as the variance of the environmental conditions of its four sub-squares (i.e. $0.05^\circ \times 0.05^\circ$ -, eighth-degree square-, QDS- and HDS-scale). Since our index measures within-square heterogeneity at each spatial scale, it can be related directly to species richness at the QDS-, HDS- and DS-scales.

We used principal components analysis (PCA), applied to the nine environmental variables across both regions, to extract a major axis of environmental heterogeneity. For this purpose, the layers describing heterogeneity in the nine environmental variables at each spatial scale were first \log_{10} -transformed to ensure normality. A separate PCA was then run at the four spatial scales. The first axis (PC1) from each represents a major axis of heterogeneity across the nine environmental heterogeneity variables considered (see Figure S4).

To compare heterogeneity in the nine environmental variables and in the major axis (PC1) of heterogeneity between the two regions, as with species richness and turnover, we employed common language effect size (*CLES*) and two-sided Mann-Whitney *U*-tests, as some forms of heterogeneity were non-normally distributed. Both analyses were repeated at all four spatial scales. This enabled us to assess scale-dependence in heterogeneity and to ascertain the spatial scale at which regional environmental heterogeneity differences are most pronounced.

2.3: Environmental heterogeneity as an explanation of species richness

We used linear models to assess associations between heterogeneity (in the nine selected environmental variables and the major axis, PC1) with species richness across the two regions. We did not incorporate spatial autocorrelation into our models of species richness because the aim of our analyses was to expose the associations between different forms of spatial environmental heterogeneity and species richness, rather than to predict species richness or to partition sources of its variation. The fact that we focused on heterogeneity between neighboring pixels is at odds with spatial autocorrelation which occurs when pixels are more related to each other than more distant pixels due to their proximity.

We first used analyses of covariance (ANCOVA), at the QDS-, HDS- and DS-scales, to relate species richness (*S*) to each environmental heterogeneity variable (including the major axis) across the two regions. In order to assess whether these relationships are identical in form across the two regions, we fitted three nested models for each heterogeneity predictor *X*, as follows: a “main effect only” model ($S \sim \beta_0 + \beta_1 X$), a “main effect + region” model ($S \sim \beta_0 + \beta_1 X + \beta_2 \text{Region}$) and a “main effect \times region” model, which includes an interaction between heterogeneity and region ($S \sim \beta_0 + \beta_1 X + \beta_2 \text{Region} + \beta_3 (X \times \text{Region})$). For each of the ten predictors, the best fitting model was determined using Akaike’s information criterion (*AIC*), as the simplest model with $\Delta AIC < 2$. Where the “main effect only” model describes heterogeneity as having a uniform effect on richness across the two regions, the “main effect + region” and “main effect \times region” models describe the relationships of richness to heterogeneity as being region dependent. Specifically, where the additive model describes these relationships as being identical in terms of slope but not intercept, the latter model describes them as differing in both.

We then used multiple linear regressions to model species richness as a function of region and heterogeneity in all nine environmental variables simultaneously. As with the simple regressions, this was done at all three spatial scales. Each starting from a 19-predictor model (i.e. region, the heterogeneities of all nine environmental variables and the interactions of those with region), models were simplified using reverse stepwise model selection based on *AIC*-scores. In the final model, the retention of significant heterogeneity-region interaction terms indicates that the dependence of richness on heterogeneity differs between the two regions. Conversely, where

only the main effect is significant, we infer the dependence of richness on heterogeneity to be uniform across the two regions.

2.4: Species richness hotspots

To identify hotspots of exceptional richness, i.e. squares whose species richness exceeded that expected given their environmental heterogeneity, we used the residuals associated with the PC1-based ANCOVA models and the multiple regression models, at all three spatial scales, to identify outlier points (i.e., points with residual S more than two standard deviations from the mean predicted S at a given level of heterogeneity). We also used F -tests to assess whether the variances of these residuals differed between the GCFR and SWAFR. Finally, to assess whether the exceptional richness of hotspots is best explained by factors other than heterogeneity, and the sensitivity of results to these hotspots, we repeated the ANCOVA and multiple regression analyses with hotspots omitted and compared the coefficients of determination from each.

3: Results

3.1: Comparing species richness

Vascular plant species richness varies spatially in both the GCFR and SWAFR, with each region possessing a hotspot of exceptional richness (the Kogelberg Centre in the GCFR; Greater Perth in the SWAFR) and declining richness towards its interior margin (Figure 1a, b). Comparisons of species richness between the regions using Mann-Whitney U -tests reveal that species richness per unit area is similar at the QDS- (Figure 2c; $P = 0.624$, $CLES = 0.490$), somewhat greater in the GCFR at HDS- (Figure 2b; $P = 0.061$, $CLES = 0.579$) and significantly greater in the GCFR at DS-scales (Figure 2a; $P = 0.050$, $CLES = 0.717$).

Partitioning S_{HDS} into its α and β components (\bar{S}_{QDS} and T_{QDS} respectively; Figure 2g), we found that the average proportional contribution of floristic turnover (i.e. T_{QDS}/S_{HDS} ; Figure 2e) is greater in the GCFR (0.64) than in the SWAFR (0.60; Mann-Whitney U -test, $P < 0.001$; $CLES = 0.750$). This is also the case at the DS-scale (Figure 2d; T_{HDS}/S_{DS} , GCFR: 0.57, SWAFR: 0.49; Mann-Whitney U test, $P < 0.001$, $CLES = 0.967$). Thus, at both scales, plant species richness in the GCFR owes more to turnover between sub-squares (Figure 2f, g: points further above $y = x$ line) than is the case in the SWAFR (Figure 2f, g: points closer to $y = x$ line).

3.2: Comparing environmental heterogeneity

With the exception of a few variables at the DS-scale (i.e. MAP, NDVI, CEC, clay, soil C), $CLES$ comparisons revealed the GCFR to be more heterogeneous than the SWAFR in all nine environmental variables, and across the full range of spatial scales studied (Figure 3). The same was true for the major axis of heterogeneity described by PC1 (Figure 3j), which accounted for between 38% ($0.10^\circ \times 0.10^\circ$ -scale; Figure S4a) and 50% (DS-scale; Figure S4d) of the variance across heterogeneity variables. The greater overall heterogeneity of the GCFR is striking when compared visually (e.g. at the HDS-scale: Figure 1c vs d). In general, the disparity in heterogeneity between the two regions seems greater for topographic and climatic ($CLES$ between ca. 0.60 and 1.00, respectively; Figure 3a-d) than edaphic variables ($CLES$ between ca. 0.60 and 0.80; Figure 3f-i).

Regressions indicate that the degree to which the GCFR is more environmentally heterogeneous than the SWAFR is scale-dependent (i.e. difference greater at broad scales), with the notable

exceptions of MAP and clay (Figure 3b,h), in which the GCFR is more heterogeneous at fine scales. The major axis of heterogeneity (PC1) displays scale-independence of most forms of heterogeneity, with its *CLES* being similar across spatial scales (Figure 3j).

3.3: Environmental heterogeneity as an explanation of species richness

With just two exceptions (elevation and CEC at the DS-scale), the univariate ANCOVA results (Table 1) show that heterogeneity in each of the nine environmental heterogeneity variables, as well as the main axis of heterogeneity (PC1), influences species richness in a consistently positive manner across the two study regions. At all three scales, the “main effect only” model was favoured for more variables, including the major axis of heterogeneity (PC1), than was any other model (Table 1; Figure 4). This indicates a broadly uniform effect of heterogeneity on species richness across the two regions. Moreover, although the “main effect + region” model was favoured for the major axis of heterogeneity at the QDS-scale (Table 1a; Figure 4a), the difference in intercepts between the GCFR and SWAFR (Table 1a: 92.5 species) is small relative to the spread of residual species richness at the QDS-scale within each region ($SD_{GCFR} = 343.5$; $SD_{SWAFR} = 245.8$; Table S3a). This indicates that regional differences in the form of the QDS-scale richness-heterogeneity relationship, while statistically significant, are subtle. Application of a quadratic “main effect only” model (i.e., $S \sim \beta_0 + \beta_1 X + \beta_2 X^2$) to the richness-major axis of heterogeneity (PC1) relationship provides evidence of non-linearity at the QDS-scale, with a quadratic model being favoured over a linear model ($\Delta AIC_{linear} = 23.78$), and the quadratic and linear coefficients both being positive ($\beta_1 = 69.02$, $\beta_2 = 9.99$; both $P < 0.001$) and describing an upward-opening parabola. At both the HDS- and DS-scales, however, the linear model is favoured. Finally, although the “main effect \times region” model was optimal for some variables, the interaction terms associated with these models either describe positive effects on richness in both regions or a positive effect in one region and a negligible effect in the other (e.g. HDS-scale heterogeneity in clay; Table 1b).

The use of multiple regression models, to assess the effects of heterogeneity in individual variables on species richness when applied jointly, largely corroborates our inference of a generally positive effect of heterogeneity on species richness. The partial effects of heterogeneity predictors retained in the optimal multiple regression models (Figure 5) are mostly positive or neutral in both regions, the only exceptions being heterogeneity in elevation at the DS-scale, PDQ at the QDS- and DS-scales, CEC at the QDS- and DS-scales and pH at all spatial scales. These results add nuance to the generally positive influence of heterogeneity on species richness displayed by the PC1-based ANCOVAs (Figure 4), and in most forms of heterogeneity when compared in isolation (Table 1). Interestingly, DS-scale heterogeneity in elevation and CEC have both negative univariate (Table 1c) and partial (Figure 5c) effects, while heterogeneity in MAP has the greatest explanatory power in the QDS- and HDS-scale models (Table S2a, b). Like the optimal ANCOVA models, multiple regression models suggest region-dependence in the relationships of species richness to environmental heterogeneity at all three spatial scales (Figure 5). Particularly at the DS-scale (Figure 5c), the coefficients associated with the partial effects of heterogeneity in SWAFR relative to the GCFR effectively cancel out the effects of these variables estimated in the GCFR. This indicates weaker or nonexistent partial effects of the different forms of heterogeneity on richness in the SWAFR.

The coefficients of determination associated with the optimal multiple regression models (Figure 5: $R^2 = 0.27, 0.43$ and 0.85) are consistently greater than those associated with the optimal ANCOVA models based on PC1 (Figure 4: $R^2 = 0.13, 0.22$ and 0.49). This indicates that PC1

does not capture all of the variation in environmental heterogeneity relevant to the prediction of species richness.

3.4: Species richness hotspots

Residual-based outlier identification yielded very similar results for the PC1-based ANCOVA and multiple regression models at all spatial scales (Table 2; Figure 1e-h, S9e-h, S10e-h). For both the GCFR and SWAFR, outliers are geographically clustered, in areas corresponding to recognized diversity centres (Figure S5, S6). At the QDS-scale, for example, GCFR outliers are concentrated in the Kogelberg-Hottentots-Holland and Cederberg areas, while SWAFR outliers are concentrated in the Mt Lesueur, Perth, Stirling-Albany and Fitzgerald River areas (Figure S5a, b; S6a, b; S9e-h). Fewer hotspots are resolved at the HDS-scale (Figure S5c, d; S6c, d; 1e-h) and only one (the Hottentots Holland area in the GCFR) at the DS-scale (Figure S5e; S10e-h). Omission of outliers from the PC1-based ANCOVA (Figure S7) and multiple regressions (Figure S8) yielded qualitatively similar models as before, but with improved coefficients of determination. At the DS-scale, the GCFR and SWAFR do not differ significantly in their residual standard deviations (under both the PC1-based ANCOVA and multiple regressions) (Table S3c). This highlights that the commonality of the richness-heterogeneity relationship is recovered more accurately at broader spatial scales following the omission of hotspots.

4: Discussion

Consistent with a recent meta-analysis identifying environmental heterogeneity as a universal driver of species richness (Stein, Gerstner, & Kreft, 2014), we found heterogeneity to have a consistently positive influence on richness in the GCFR and SWAFR. Most significant partial effects associated with heterogeneity terms in our multiple regression models and the overall effect of heterogeneity (PC1) in our ANCOVA results are positive. Thus, we find no evidence for the hump-backed response of richness to heterogeneity anticipated by some authors (Allouche, Kalyuzhny, Moreno-Rueda, Pizarro, & Kadmon, 2012; Carnicer, Brotons, Herrando, & Sol, 2013), at least at the scales considered here. Although a quadratic relationship between richness and heterogeneity is favoured at the QDS-scale, this describes an upward-opening parabola rather than a humped, unimodal curve. Additionally, and also consistent with Stein et al. (2014), we find the strength of the heterogeneity-richness relationship to associate positively with spatial scale, as evidenced by that fact that the coefficients of determination associated with our models are greatest at the DS- and smallest at the QDS-scale. One possible explanation of this effect is the fact that larger areas accommodate more environmental variability (Wüest, Boucher, Bouchenak-Khelladi, Karger, & Linder, 2019) and so facilitate stronger heterogeneity-richness relationships (Van Rensburg, Chown, & Gaston, 2002). In speciation hotspots (see below) like the GCFR or the SWAFR, however, an important additional consideration is the spatial scale of speciation (Stein et al., 2014). Stronger coupling of species richness to heterogeneity at the DS-scale may arise because DS, unlike QDS and HDS, are sufficiently large to capture speciation process, especially allopatric speciation.

We observe species richness to respond to environmental heterogeneity in a relatively uniform manner across the GCFR and SWAFR. Where this is not the case, regional differences in the form of the richness-heterogeneity relationship are subtle. This suggests that the greater vascular plant species richness of the GCFR is partly attributable to the greater physiographic heterogeneity of this region. As noted by Cowling et al. (2015) the lower heterogeneity of SWAFR, both now and during the Tertiary, would have constrained opportunities for radiation,

thereby producing a flora that is generally less diverse. In addition, the greater heterogeneity of the GCFR might facilitate denser species packing, through the provision of greater niche diversity (Cramer, West, Power, Skelton, & Stock, 2014; Hart, Usinowicz, & Levine, 2017). Our data reveal that, despite its significantly greater DS-scale species richness, the GCFR does not have significantly greater QDS- and HDS-scale species richness than the SWAFR. This pattern, which reflects higher rates of turnover between QDS and HDS in the GCFR, implies a greater role for dispersal limitation and local species differentiation in driving high DS-scale richness in the GCFR. This interpretation is consistent with existing evidence for a higher frequency of single-site endemic taxa in the Cape than in the Australian flora (Linder, 2019; see also Figure S11).

Although our data confirm a broad dependence of GCFR and SWAFR plant species richness on environmental heterogeneity, as defined by the nine environmental variables examined, the generally low coefficients of determination associated with these relationships indicates roles for other factors. Firstly, species richness is almost certainly influenced by heterogeneity in other environmental variables. Cramer, Wootton, van Mazijk, & Verboom (2019), for example, recently highlighted the superiority of locally modelled soil layers, including soil chemistry variables, as predictors of vegetation type in the GCFR, compared with the globally modelled layers used in this study. Unfortunately, we were unable to make use of these layers as comparable data are lacking for the SWAFR. Despite the obvious importance of soil variables as determinants of plant species richness patterns, particularly in the two regions under study, where they influence both the genesis and distribution of diversity, often at local scales (Cowling et al. 1994), their inclusion in broad-scale spatial modelling exercises remains problematic since the spatial scale at which soils vary is typically much finer than the spatial resolution of modelled layers (Figueiredo et al., 2018). Similarly, the association of many point-endemics in the GCFR with small, highly localized wetlands whose distributions are geomorphologically rather than climatically determined, presents challenges for species distribution modelling (Born & Linder, 2018). The same may well be true for the SWAFR flora where phylogenetically-relictual species typically inhabit waterlogged situations (Hopper & Gioia, 2004).

A second factor affecting the explanatory power of our models is the inclusion of only the heterogeneities of environmental variables, not their absolute values. Although the absolute values of certain variables, particularly those influencing biological productivity (i.e. energy-water theory), have often been found to correlate positively with species richness at broad scales (Currie, 1991; Hawkins et al., 2003; Kreft & Jetz, 2007), we elected to omit these variables from our analyses because we wished to assess explicitly the relative ability of various forms of heterogeneity to account for richness variation across and between the GCFR and SWAFR. Besides, as noted by Cowling, Bradshaw, Colville, & Forest (2017), energy-water theory does not appear to hold in the GCFR where species richness is maximized in cool, nutrient-impooverished areas.

A third factor underpinning the unexplained variance in our models is spatial variation in collection effort and its consequences for species richness estimates. Although some authors have argued for the application of sample- (Gotelli & Colwell, 2001) or coverage-based rarefaction techniques (Chao & Jost, 2012) to correct for variable collection effort (Engemann et al., 2015), the application of these methods to systematically-biased collection data is problematic. Where these techniques assume that the relationship between true species richness and sampling effort is weak, this is typically not the case for herbarium-based species record data. Herbarium collectors commonly focus their efforts on areas having high concentrations of rare species, and many herbaria do not accept material indiscriminately, being reluctant to accept

multiple accessions of the same species from a single area. Collection effort is thus highly nonrandom in relation to species richness. Moreover, rarefaction-methods, while being able to account for spatially non-uniform sampling effort, do not cater for targeted collection efforts, e.g. collecting a specific taxon and disregarding others in a given locality. Whilst acknowledging arguments for applying rarefaction techniques to herbarium-based species data (e.g. Gioia & Hopper, 2017) and recognizing that the jury is still out on whether they are appropriate in this context, we have desisted from applying them here because, at least for the South African flora, they severely distort known species richness (Cramer & Verboom, 2016). On the one hand, use of rarefaction may artificially inflate richness in where collection effort is low, not on account of accessibility but because they are truly species-poor; on the other, it may underestimate species richness in areas which have enjoyed disproportionately high levels of sampling effort by virtue of their exceptional richness. Vindicating their decision not to apply rarefaction techniques to herbarium-based species data, Cramer & Verboom (2016) found that the cumulative species richness estimates yielded by their analyses corresponded well to known biome-level floristic richness, and that these numbers were consistent with results from an exhaustive atlasing project (i.e. Protea Atlas; Rebelo, 2001) with low collecting bias (Merow, Smith, & Silander, 2013).

A fourth and final factor potentially compromising the strength of the heterogeneity-richness relationship is the existence of diversity hotspots whose high species richness is not directly linked to physiographic heterogeneity. The distribution of plant species richness in both the GCFR and SWAFR is highly uneven (cf. Oliver, Linder, & Rourke, 1983; Gioia & Hopper, 2017), with the exceptional richness of some sites potentially being a consequence of long-term climatic and/or hydrological stability there. In the CFR, for example, the higher plant species richness of the west has been attributed to reduced rates of extinction, associated with greater climatic stability through the Pleistocene (Cowling & Lombard, 2002; Cowling et al., 2017). In this study, we used the residuals derived from PC1-based ANCOVA and multiple regression models to identify hotspots whose richness substantially exceeds that expected based on their underlying heterogeneities. Several of the hotspots thus identified correspond to centres of long-term environmental stability. In the GCFR, the southwestern mountains (Kogelberg-Hottentots Holland) have been identified as a long-term climatic and hydrological refugium, especially for species which inhabit the area's numerous bogs and seeps (Linder, 2019; Wüest et al., 2019). Similarly, pollen- and midden-based isotope data provide evidence of relatively muted climate change in the Cederberg during Pleistocene (Meadows, Chase, & Seliane, 2010; Meadows & Sugden, 1991, 1993). The biota of the SWAFR also shows evidence of climatically-forced range contraction in the Pleistocene (Byrne & Hines, 2004; Byrne, 2008), with putative refugia there including the Stirling and Porungurup ranges, whose physical elevation and complex topography offer cooler conditions under temperature elevation (Keppel et al., 2017). Additionally, maps in Byrne (2008; Figure 2) identify a refugial area in the vicinity of Perth and a second in the vicinity of Fitzgerald River. Although the emergence of the Perth area as a species richness outlier was unexpected (see also Gioia and Hopper 2017), this may be a consequence of several factors. These include our failure to capture adequately the variable hydrology and geology of the Swan River Coastal Plain, elevated collection rates associated with Perth being a metropolitan area (Gioia and Hopper 2017), and the potential introduction of plants into the Greater Perth metropolitan area from elsewhere in the SWAFR. Based on a random sample of 300 of the 2,944 species records from the Perth area, roughly 7% are distributional outliers (Figure S12), identifying them as possible introductions.

Given their importance for long-term species persistence (McLaughlin et al., 2017), the accurate identification and effective protection of climatically- and/or hydrologically stable hotspots must be an important conservation objective, particularly in the face of contemporary climatic

deterioration. In the GCFR, therefore, it is important to establish whether the high species richness in the west is attributable to a broad longitudinal effect (Cowling & Lombard, 2002; Cowling et al., 2017; Verboom et al., 2014) or the presence of hotspots of more locally-determined extreme richness. Our data support the latter interpretation, with the Kogelberg-Hottentots Holland area emerging as a particularly important refugium for Cape plant diversity. Given that much of the locally-endemic diversity of this area is associated with bog and seepage habitats (Linder, 2019), recent proposals to abstract water from the Table Mountain Group aquifer, which feeds these bogs, present a serious threat to the continued persistence of Cape floristic diversity (Slingsby et al., 2018). Groundwater abstraction in the Perth area presents a similar threat in the SWAFR (Barron et al., 2014; Froend & Sommer, 2010).

Although the floristic richness of the GCFR has been identified as being globally anomalous, our data reveal that richness in the GCFR and SWAFR varies broadly as a common function of environmental heterogeneity, promoting diversification and facilitating species coexistence. This corroborates the suggestion (Cowling et al., 2015) that the greater richness of the GCFR, relative to that of the SWAFR, is largely a product of its greater physical heterogeneity. Importantly, since environmental heterogeneity is spatially configured, the greater richness of the GCFR flora is associated with higher rates of species turnover in space. The species richness-environmental heterogeneity relationship is, however, punctuated by the existence of local hotspots, whose exceptional richness may be a consequence of historical factors. Though these hotspots may be important for species persistence, e.g. in the face of climate change, we find evidence for a unified conceptual framework for thinking about the species richness of these two mediterranean floras.

Tables

Table 1: Values and significances¹ of coefficients from univariate ANCOVAs of vascular plant species richness against separate axes and the main axis (PC1) of environmental heterogeneity² (log₁₀-transformed) across the GCFR and SWAFR at the (a) QDS-, (b) HDS- and (c) DS-scales. The SWAFR interaction terms describe differences in slope between the GCFR and SWAFR.

Response	Model type	Predictor	Main effect	SWAFR effect	SWAFR interaction
(a) S_{QDS}	Main effect \times region	Elevation	140.2 ***	124.4 ***	-56.1 *
		MAP	172.0 ***	54.5 **	-54.1 **
		PDQ	73.4 ***	55.6 *	61.9 **
		NDVI	154.9 ***	-7.8	-102.2 ***
	Main effect + region	PC1	67.0 ***	92.5 ***	
	Main effect only	Surface T	62.1 ***		
		CEC	14.7		
		Clay	42.1 ***		
		Soil C	62.9 ***		
		pH	21.9 *		
(b) S_{HDS}	Main effect \times region	MAP	399.0 ***	-41.5	-185.0 **
		Clay	-12.8	-216.1 **	173.6 *
	Main effect only	Elevation	163.7 ***		
		PDQ	226.3 ***		
		Surface T	135.9 ***		
		NDVI	246.6 ***		
		Soil C	159.4 ***		
		PC1	123.1 ***		
	Region only	CEC	-26.3	-251.9 **	
		pH	53.8	-193.0 *	
(c) S_{DS}	Main effect \times region	Elevation	-1455.9 *	-2278.4 **	1668.5 *
		MAP	683.3 ***	-519.1 **	-382.1 *
		CEC	-933.3 **	-1043.4 ***	837.1 *
	Main effect + region	Clay	273.0 *	-542.8 *	
		Soil C	246.5 *	-615.4 *	
	Main effect only	PDQ	363.1 **		
		Surface T	336.7 **		
		NDVI	475.3 ***		
		pH	448.4 ***		
		PC1	231.1 ***		

¹ Represented as follows: ***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$; blank, NS.

² Abbreviations are as follows: NDVI, normalized difference vegetation index; T, temperature; MAP, mean annual precipitation; PDQ, precipitation in the driest quarter; CEC, cation exchange capacity; C, carbon.

Table 2: Pearson’s r correlation coefficients comparing the predicted and residual species richness (S) between the PC1-based ANCOVAs (Table 1, Figure 4) and the multivariate regressions (Figure 5) of vascular plant species richness across the GCFR and SWAFR. All correlation coefficients were significant ($P < 0.001$; two-sided t -test).

Spatial scale	Correlation	
	Predicted S	Residual S
QDS	0.680	0.908
HDS	0.699	0.834
DS	0.723	0.369

476 **Figures**

477

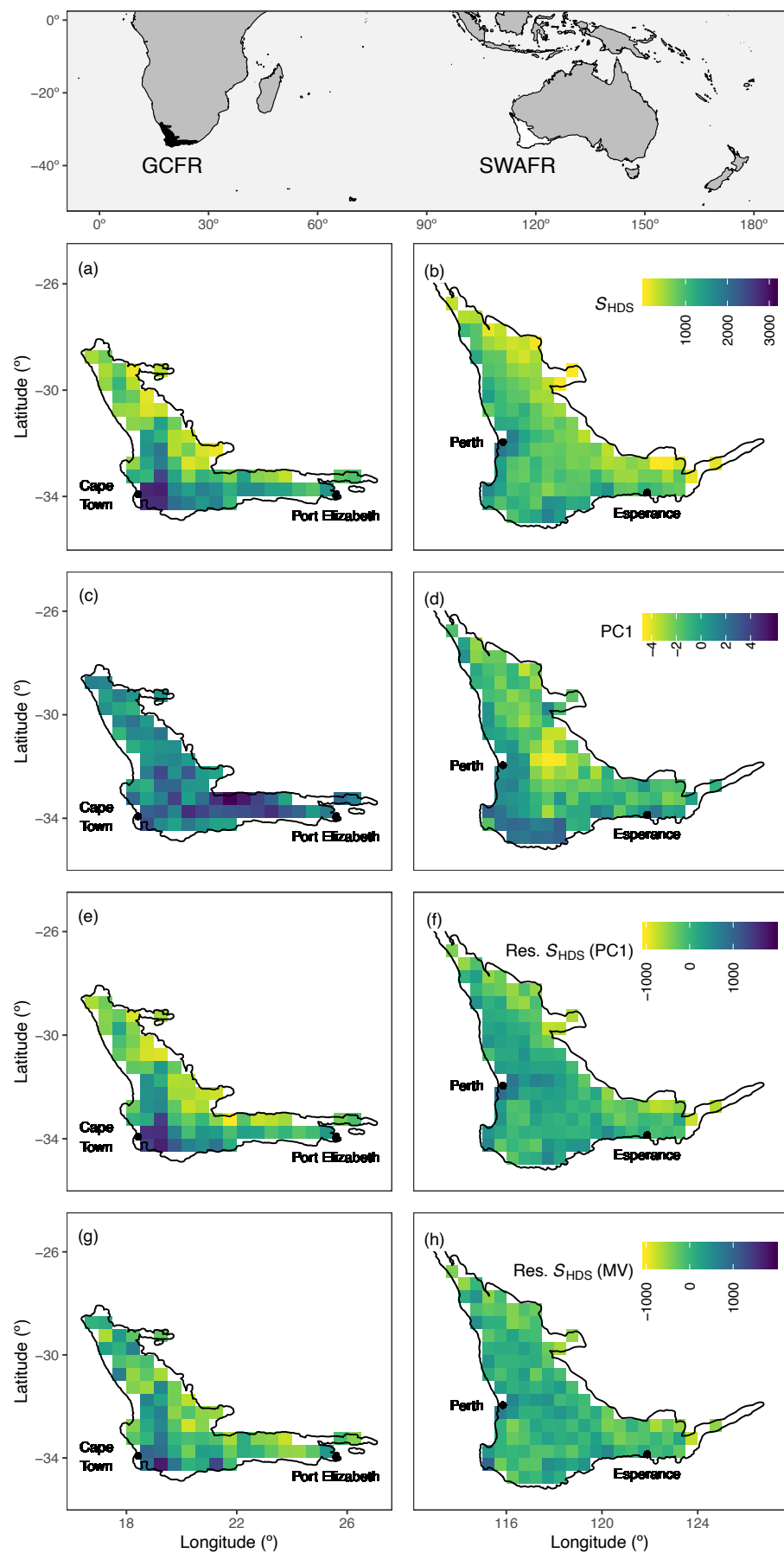


Figure 1 (previous page): HDS-scale maps for the GCFR and SWAFR (illustrated in black and white respectively in the topmost panel) of **(a,b)** vascular plant species richness (S_{HDS}), **(c,d)** values of the major axis of environmental heterogeneity (PC1) from the PCA of nine separate forms (\log_{10} -transformed; Table S1) and residual species richness following regressions against **(e,f)** PC1 (see also Figure 4b) and **(g,h)** the multivariate (MV) regression model (see also Figure 5b). Map projection used: WGS84. Colour versions of these maps, and the QDS- and DS-scale equivalents (Figure S9, S10 respectively) are available in the online version.

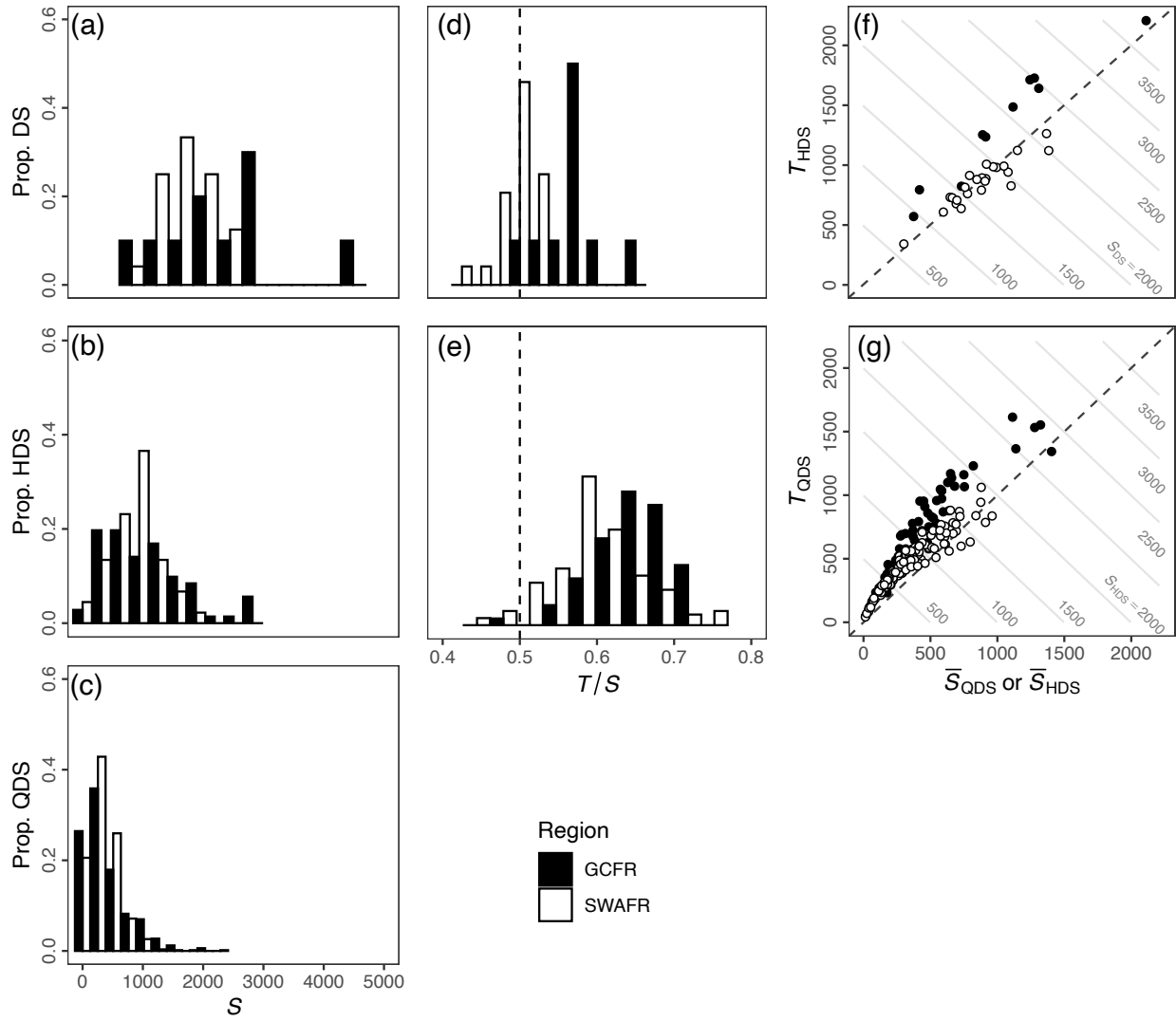


Figure 2: Frequency distributions of vascular plant species richness (S) in the GCFR and SWAFR at the (a) DS-, (b) HDS-, and (c) QDS-scales, and of the proportional contributions of floristic turnover (i.e. T/S) to (d) S_{DS} (i.e. T_{HDS} / S_{DS}) and (e) S_{HDS} . (i.e. T_{QDS} / S_{HDS}). Frequencies are scaled as the proportions of cells within each region. The derivations of T_{DS} and T_{HDS} are demonstrated with scatter plots of mean richness at the (f) HDS- (\bar{S}_{HDS}) and (g) QDS-scales (\bar{S}_{QDS}) against turnover (T_{HDS} and T_{QDS} , respectively), with contour lines denoting the S_{DS} and S_{HDS} that arise as their sums.

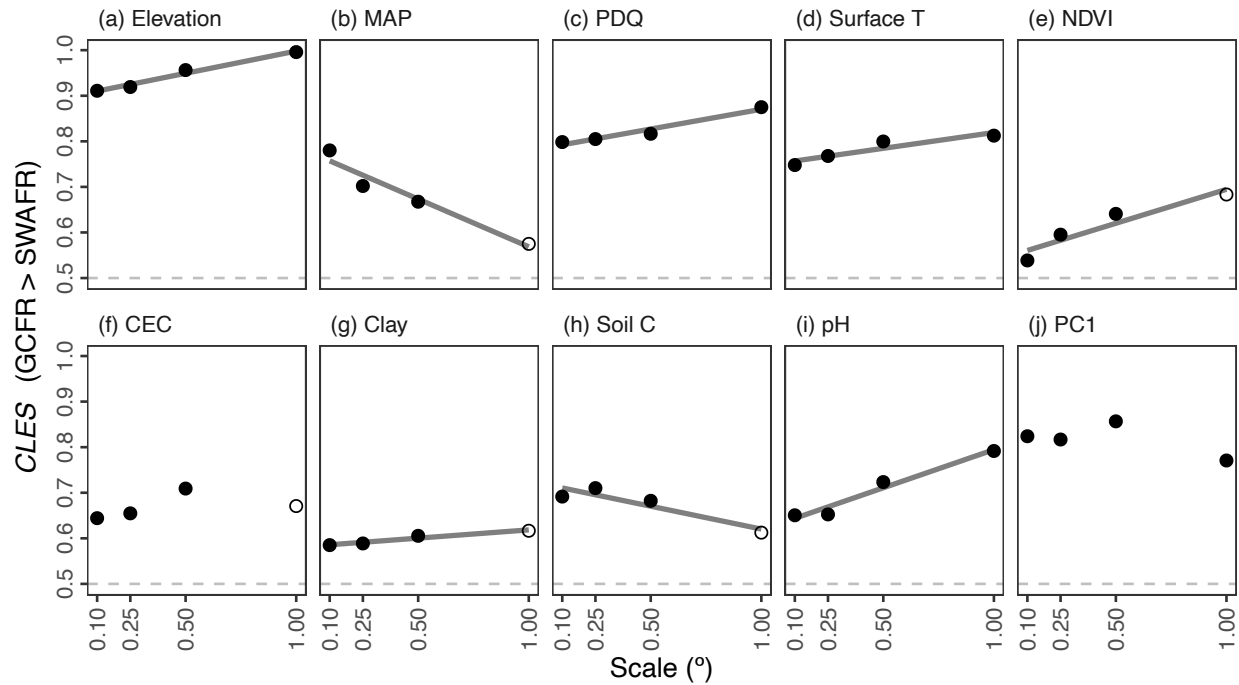


Figure 3: The common language effect size (*CLES*) of (a–i) various forms of environmental heterogeneity (log₁₀-transformed) and (j) the major axis thereof (PC1) in the GCFR and SWAFR. *CLES* here is treated as the effect of GCFR relative to SWAFR values. Filled points represent comparisons where the GCFR and SWAFR significantly differed in heterogeneity ($P < 0.05$; Mann-Whitney *U*-tests), while unfilled points represent those that were not significant. Following simple linear regressions of *CLES* against scale, we found evidence for scale-dependence (depicted by lines) for most variables. Both strongly significant ($P < 0.05$; a–c,g,i) and marginally significant ($P < 0.1$; d,e,h) fits have been plotted, the latter for illustration. Abbreviations are as in Table 1.

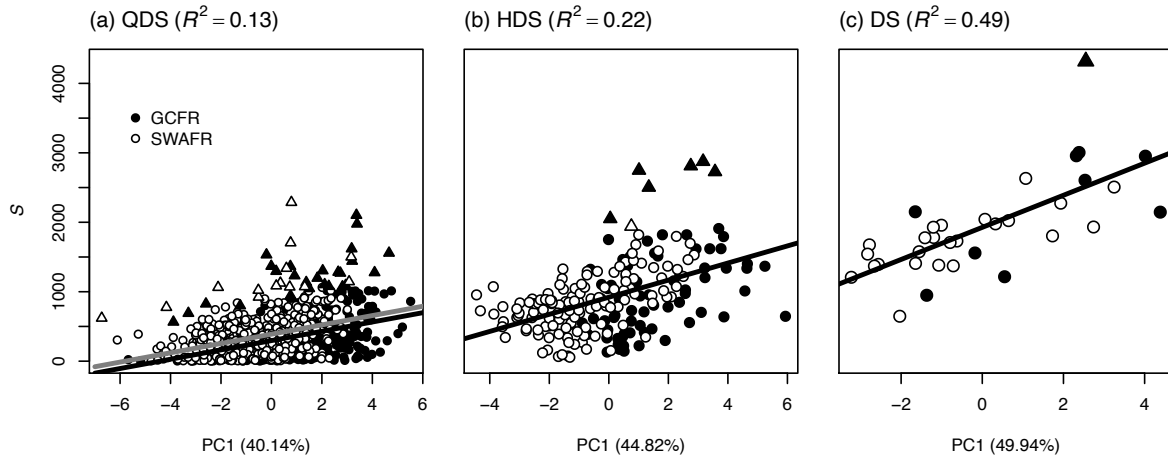


Figure 4: Simple linear regressions of vascular plant species richness ((a) S_{QDS} , (b) S_{HDS} and (c) S_{DS}) against each scale's major axis of environmental heterogeneity (PC1) across the GCFR and SWAFR. Grid-cells representing hotspots of species richness (identified as having residual S greater than two standard deviations from the mean) are identified by triangles (black for the GCFR, white for SWAFR). These three linear models are presented in Table 1, all with highly significant slopes ($P < 0.001$) and were fitted including species richness hotspots. For S_{QDS} , the separate fits for the GCFR (black) and SWAFR (grey) are presented following the best fitting model at that scale: a "main effect + region" model ($S \sim \beta_0 + \beta_1 X + \beta_2 Region$; Table 1a). The R^2 -values of each model and the percentage variation in environmental heterogeneity explained by PC1 (from each of the three PCAs) are noted in parentheses in the panel and horizontal axis headings respectively. The equivalent results when hotspots were excluded is available in the online version (Figure S7).

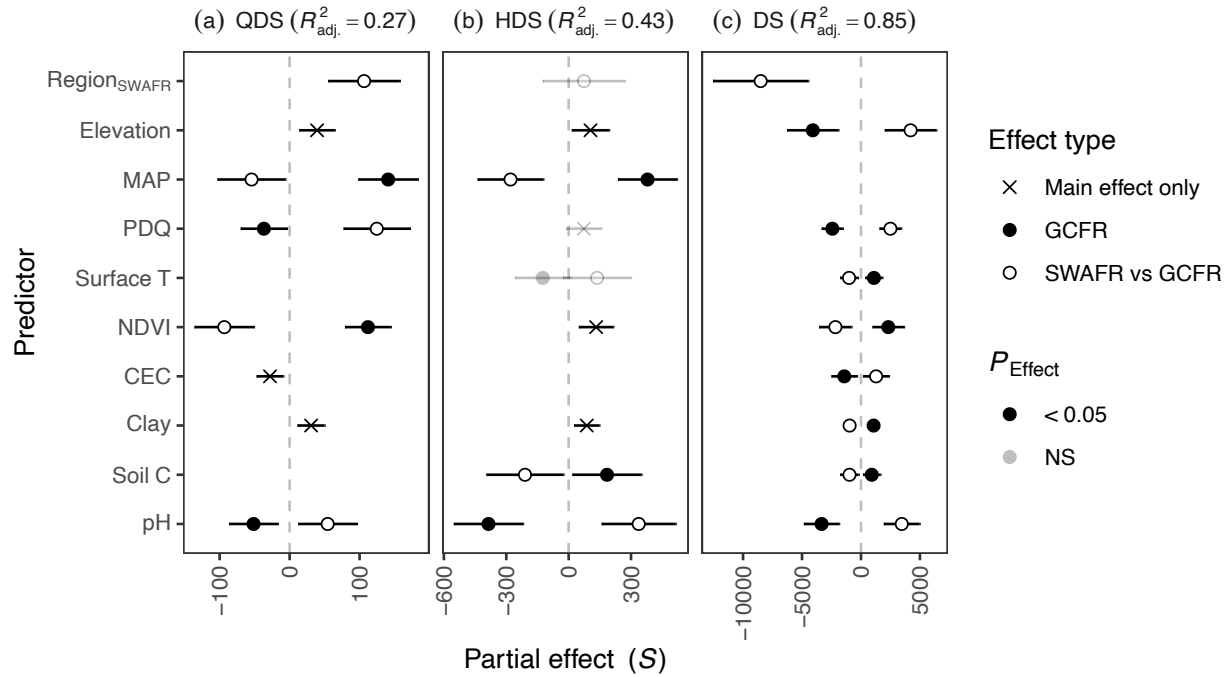


Figure 5: Partial effects from multiple linear regressions of vascular plant species richness ((a) S_{QDS} , (b) S_{HDS} and (c) S_{DS}) against the various forms of environmental heterogeneity (\log_{10} -transformed and re-scaled) across the GCFR and SWAFR. Points with error bars denote slopes and their 95% confidence intervals. Filled points represent effect estimates for the GCFR and empty points represent effect estimates for the SWAFR relative to the GCFR (i.e. where region-interaction terms were retained during stepwise model selection). Crosses represent main effects (i.e. no region-interaction term retained). Estimates illustrated in black were significant ($P < 0.05$), while those in grey were not, but still retained during stepwise model selection. The multiple adjusted R^2 -values of each model are noted in parentheses in the panel headings. Abbreviations of variables are as in Table 1 and Figure 3. The equivalent results when species richness hotspots were excluded is available in the online version (Figure S8).

534 Data availability statement

535 The raw vascular plant occurrence records used are available from GBIF (GCFR:
536 <https://doi.org/10.15468/dl.n6u6n0>; SWAFR: <https://doi.org/10.15468/dl.46okua>). Analyses in
537 the form of R-scripts, cleaned species lists, raster-layers of the nine forms of environmental
538 heterogeneity, the major axis of heterogeneity (PC1) and species richness (at the three spatial
539 scales) are available in the DRYAD Digital Repository:
540 <https://doi.org/10.5061/dryad.8w9ghx3m8>.

541 References

- 542 Ackerly, D. D. (2009). Evolution, origin and age of lineages in the Californian and
543 Mediterranean floras. *Journal of Biogeography*, 36(7), 1221–1233.
544 <https://doi.org/10.1111/j.1365-2699.2009.02097.x>
- 545 Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M., & Kadmon, R. (2012). Area-
546 heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the*
547 *National Academy of Sciences*, 109(43), 17495–17500.
- 548 Barron, O., Froend, R., Hodgson, G., Ali, R., Dawes, W., Davies, P., & McFarlane, D. (2014).
549 Projected risks to groundwater-dependent terrestrial vegetation caused by changing climate
550 and groundwater abstraction in the Central Perth Basin, Western Australia. *Hydrological*
551 *Processes*, 28(22), 5513–5529.
- 552 Beard, J. S., Chapman, A. R., & Gioia, P. (2000). Species richness and endemism in the Western
553 Australian flora. *Journal of Biogeography*, 27(6), 1257–1268.
554 <https://doi.org/10.1046/j.1365-2699.2000.00509.x>
- 555 Bergh, N. G., & Linder, H. P. (2009). Cape diversification and repeated out-of-southern-Africa
556 dispersal in paper daisies (Asteraceae–Gnaphalieae). *Molecular Phylogenetics and*
557 *Evolution*, 51(1), 5–18.
- 558 Bivand, R., Keitt, T., & Rowlingson, B. (2017). *rgdal: Bindings for the Geospatial Data*
559 *Abstraction Library. R package version 1.2-7*. <https://cran.r-project.org/package=rgdal>
- 560 Bøhn, T., & Amundsen, P.-A. (2004). Ecological interactions and evolution: forgotten parts of
561 biodiversity? *BioScience*, 54(9), 804–805.
- 562 Born, J., Linder, H. P., & Desmet, P. (2007). The Greater Cape Floristic Region. *Journal of*
563 *Biogeography*, 34(1), 147–162. <https://doi.org/10.1111/j.1365-2699.2006.01595.x>
- 564 Born, Julia, & Linder, H. P. (2018). Water availability, fundamental niches and realized niches:
565 A case study from the Cape flora. *Austral Ecology*, 43(6), 696–705.
566 <https://doi.org/10.1111/aec.12616>
- 567 Bradshaw, P. L., & Cowling, R. M. (2014). Landscapes, rock types, and climate of the Greater
568 Cape Floristic Region. In N. Allsopp, J. F. Colville, & G. A. Verboom (Eds.), *Fynbos:*
569 *Ecology, Evolution and Conservation of a Megadiverse Region* (pp. 26–46). Oxford
570 University Press. <https://doi.org/oso/9780199679584.003.0002>
- 571 Byrne, M. (2008). Evidence for multiple refugia at different time scales during Pleistocene
572 climatic oscillations in southern Australia inferred from phylogeography. *Quaternary*

- 573 *Science Reviews*, 27(27–28), 2576–2585. <https://doi.org/10.1016/j.quascirev.2008.08.032>
- 574 Byrne, M., & Hines, B. (2004). Phylogeographical analysis of cpDNA variation in *Eucalyptus*
575 *loxophleba* (Myrtaceae). *Australian Journal of Botany*, 52(4), 459–470.
- 576 Carnicer, J., Brotons, L., Herrando, S., & Sol, D. (2013). Improved empirical tests of area-
577 heterogeneity tradeoffs. *Proceedings of the National Academy of Sciences*, 110(31), E2858-
578 -E2860.
- 579 Chamberlain, S., Szocs, E., Boettiger, C., Ram, K., Bartomeus, I., Baumgartner, J., Foster, Z., &
580 O'Donnell, J. (2016). *taxize: Taxonomic information from around the web. R package*
581 *version 0.7.8*. <https://github.com/ropensci/taxize>
- 582 Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing
583 samples by completeness rather than size. *Ecology*, 93(12), 2533–2547.
- 584 Cowling, R. M., & Lombard, A. T. (2002). Heterogeneity, speciation/extinction history and
585 climate: explaining regional plant diversity patterns in the Cape Floristic Region. *Diversity*
586 *and Distributions*, 8(3), 163–179.
- 587 Cowling, R. M., Witkowski, E. T. F., Milewski, A. V., & Newbey, K. R. (1994). Taxonomic,
588 Edaphic and Biological Aspects of Narrow Plant Endemism on Matched Sites in
589 Mediterranean South Africa and Australia. *Journal of Biogeography*, 21(6), 651.
590 <https://doi.org/10.2307/2846038>
- 591 Cowling, Richard M., Potts, A. J., Bradshaw, P. L., Colville, J., Arianoutsou, M., Ferrier, S.,
592 Forest, F., Fyllas, N. M., Hopper, S. D., Ojeda, F., Procheş, Ş., Smith, R. J., Rundel, P. W.,
593 Vassilakis, E., & Zutta, B. R. (2015). Variation in plant diversity in mediterranean-climate
594 ecosystems: The role of climatic and topographical stability. *Journal of Biogeography*,
595 42(3), 552–564. <https://doi.org/10.1111/jbi.12429>
- 596 Cowling, Richard M., Rundel, P. W., Lamont, B. B., Arroyo, M. K., & Arianoutsou, M. (1996).
597 Plant diversity in mediterranean-climate regions. *Trends in Ecology and Evolution*, 11(9),
598 362–366. [https://doi.org/10.1016/0169-5347\(96\)10044-6](https://doi.org/10.1016/0169-5347(96)10044-6)
- 599 Cowling, Richard M, Bradshaw, P. L., Colville, J. F., & Forest, F. (2017). Levyns' Law:
600 explaining the evolution of a remarkable longitudinal gradient in Cape plant diversity.
601 *Transactions of the Royal Society of South Africa*, 72(2), 184–201.
- 602 Cramer, M. D., & Verboom, G. A. (2016). Measures of biologically relevant environmental
603 heterogeneity improve prediction of regional plant species richness. *Journal of*
604 *Biogeography*, 44(3), 1–13. <https://doi.org/10.1111/jbi.12911>
- 605 Cramer, M. D., West, A. G., Power, S. C., Skelton, R., & Stock, W. D. (2014). Plant
606 ecophysiological diversity. In *Fynbos: Ecology, Evolution and Conservation of a*
607 *Megadiverse Region* (pp. 248–272). Oxford University Press.
608 <https://doi.org/oso/9780199679584.003.0011>
- 609 Cramer, M. D., Wootton, L. M., van Mazijk, R., & Verboom, G. A. (2019). New regionally
610 modelled soil layers improve prediction of vegetation type relative to that based on global
611 soil models. *Diversity and Distributions*, 25(11), 1736–1750.
- 612 Crisp, M. D., Arroyo, M. T. K., Cook, L. G., Gandolfo, M. A., Jordan, G. J., McGlone, M. S.,
613 Weston, P. H., Westoby, M., Wilf, P., & Linder, H. P. (2009). Phylogenetic biome

614 conservatism on a global scale. *Nature*, 458(7239), 754–756.
615 <https://doi.org/10.1038/nature07764>

616 Currie, D. J. (1991). Energy and large-scale patterns of animal-and plant-species richness. *The*
617 *American Naturalist*, 137(1), 27–49.

618 Donoghue, M. J. (2008). A phylogenetic perspective on the distribution of plant diversity.
619 *Proceedings of the National Academy of Sciences*, 105(Supplement 1), 11549–11555.
620 <https://doi.org/10.1073/pnas.0801962105>

621 Engemann, K., Enquist, B. J., Sandel, B., Boyle, B., Jørgensen, P. M., Morueta-Holme, N., Peet,
622 R. K., Violle, C., & Svenning, J.-C. (2015). Limited sampling hampers “big data”
623 estimation of species richness in a tropical biodiversity hotspot. *Ecology and Evolution*,
624 5(3), 807–820.

625 Figueiredo, F. O. G., Zuquim, G., Tuomisto, H., Moulatlet, G. M., Balslev, H., & Costa, F. R. C.
626 (2018). Beyond climate control on species range: The importance of soil data to predict
627 distribution of Amazonian plant species. *Journal of Biogeography*, 45(1), 190–200.
628 <https://doi.org/10.1111/jbi.13104>

629 Froend, R., & Sommer, B. (2010). Phreatophytic vegetation response to climatic and abstraction-
630 induced groundwater drawdown: examples of long-term spatial and temporal variability in
631 community response. *Ecological Engineering*, 36(9), 1191–1200.

632 Gioia, P., & Hopper, S. D. (2017). A new phytogeographic map for the Southwest Australian
633 Floristic Region after an exceptional decade of collection and discovery. *Botanical Journal*
634 *of the Linnean Society*, 184(1), 1–15. <https://doi.org/10.1093/botlinnean/box010>

635 Goldblatt, P. (1978). An analysis of the flora of Southern Africa: its characteristics, relationships
636 abd origins. *Annals of the Missouri Botanical Garden*, 65, 369–436.

637 Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the
638 measurement and comparison of species richness. *Ecology Letters*, 4(4), 379–391.
639 <https://doi.org/10.1046/j.1461-0248.2001.00230.x>

640 Hart, S. P., Usinowicz, J., & Levine, J. M. (2017). The spatial scales of species coexistence.
641 *Nature Ecology & Evolution*, 1(8), 1066–1073. <https://doi.org/10.1038/s41559-017-0230-7>

642 Hawkins, B. A., Field, R., Cornell, H. V., Currie, D. J., Guégan, J.-F., Kaufman, D. M., Kerr, J.
643 T., Mittelbach, G. G., Oberdorff, T., O’Brien, E. M., & others. (2003). Energy, water, and
644 broad-scale geographic patterns of species richness. *Ecology*, 84(12), 3105–3117.

645 Hijmans, R. J. (2016). *raster: Geographic Data Analysis and Modeling. R package version 2.5-*
646 *8*. <https://cran.r-project.org/package=raster>

647 Hopper, S. D. (1979). Biogeographical Aspects of Speciation in the Southwest Australian Flora.
648 *Annual Review of Ecology and Systematics*, 10, 399–422.
649 <http://www.jstor.org/stable/2096798>

650 Hopper, S. D. (2009). OCBIL theory: Towards an integrated understanding of the evolution,
651 ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes.
652 *Plant and Soil*, 322(1), 49–86. <https://doi.org/10.1007/s11104-009-0068-0>

653 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, Evolution, and Systematics*, 35, 623–650. <https://doi.org/10.1146/annurev.ecolsys.35.112202.130201>
- Jiménez, I., & Ricklefs, R. E. (2014). Diversity anomalies and spatial climate heterogeneity. *Global Ecology and Biogeography*. <https://doi.org/10.1111/geb.12181>
- Keppel, G., Robinson, T. P., Wardell-Johnson, G. W., Yates, C. J., Van Niel, K. P., Byrne, M., & Schut, A. G. T. (2017). A low-altitude mountain range as an important refugium for two narrow endemics in the Southwest Australian Floristic Region biodiversity hotspot. *Annals of Botany*, 119(2), 289–300. <https://doi.org/10.1093/aob/mcw182>
- Kreft, H., & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, 104(14), 5925–5930. <https://doi.org/10.1073/pnas.0608361104>
- Laliberte, E., Zemunik, G., & Turner, B. L. (2014). Environmental filtering explains variation in plant diversity along resource gradients. *Science*, 345(6204), 1602–1605. <https://doi.org/10.1126/Science.1256330>
- Lamont, B. B., & He, T. (2017). When did a Mediterranean-type climate originate in southwestern Australia? *Global and Planetary Change*, 156, 46–58. <https://doi.org/10.1016/j.gloplacha.2017.08.004>
- Larsen, R., Holmern, T., Prager, S. D., Maliti, H., & Røskft, E. (2009). Using the extended quarter degree grid cell system to unify mapping and sharing of biodiversity data. *African Journal of Ecology*, 47(3), 382–392. <https://doi.org/10.1111/j.1365-2028.2008.00997.x>
- Linder, H. P. (2019). Rare species, Restionaceae, and the Cape flora. *Journal of Biogeography*, 1–14. <https://doi.org/10.1111/jbi.13709>
- McLaughlin, B. C., Ackerly, D. D., Klos, P. Z., Natali, J., Dawson, T. E., & Thompson, S. E. (2017). Hydrologic refugia, plants, and climate change. *Global Change Biology*, 23(8), 2941–2961.
- Meadows, M E, Chase, B. M., & Seliane, M. (2010). Holocene palaeoenvironments of the Cederberg and Swartuggens mountains, Western Cape, South Africa: pollen and stable isotope evidence from hyrax dung middens. *Journal of Arid Environments*, 74(7), 786–793.
- Meadows, M E, & Sugden, J. M. (1991). A vegetation history of the last 14,500 years on the Cederberg, SW Cape, South Africa. *South African Journal of Science*, 87(3).
- Meadows, Michael E, & Sugden, J. M. (1993). The late Quaternary palaeoecology of a floristic kingdom: the southwestern Cape South Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 101(3–4), 271–281.
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36(10), 1058–1069.
- Milewski, A. V. (1981). A comparison of vegetation height in relation to the effectiveness of rainfall in the mediterranean and adjacent arid parts of Australia and South Africa. *Journal of Biogeography*, 8(2), 107. <https://doi.org/10.2307/2844553>
- Oliver, E. G. H., Linder, H. P., & Rourke, J. P. (1983). Geographical distribution of present-day

- 694 Cape taxa and their phytogeographical significance. *Bothalia*, 14(3/4), 427–440.
- 695 R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. [https://www.r-](https://www.r-project.org/)
696 [project.org/](https://www.r-project.org/)
- 697 Rebelo, T. G. (2001). *A field guide to the proteas of South Africa*. Fernwood.
- 698 Ricklefs, R. E. (1987). Community diversity: relative roles of local and regional processes.
699 *Science, New Series*, 235(4785), 167–171.
- 700 Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology*
701 *Letters*, 7(1), 1–15.
- 702 Slingsby, J. A., February, E. C., & Rebelo, T. G. (2018). Water: at what cost to our unique flora?
703 *Veld & Flora, June*, 72–75.
- 704 Snijman, D. A. (2013). *Plants of the Greater Cape Floristic Region. 2: The Extra Cape flora*.
705 South African National Biodiversity Institute.
- 706 Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation.
707 *Evolution*, 64(2), 295–315. <https://doi.org/10.1111/j.1558-5646.2009.00877.x>
- 708 Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of
709 species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17(7), 866–880.
710 <https://doi.org/10.1111/ele.12277>
- 711 Stock, W. D., & Verboom, G. A. (2012). Phylogenetic ecology of foliar N and P concentrations
712 and N: P ratios across mediterranean-type ecosystems. *Global Ecology and Biogeography*,
713 21(12), 1147–1156.
- 714 Van Rensburg, B. J., Chown, S. L., & Gaston, K. J. (2002). Species richness, environmental
715 correlates, and spatial scale: a test using South African birds. *The American Naturalist*,
716 159(5), 566–577. <https://doi.org/10.1086/339464>
- 717 Veech, J. A., Summerville, K. S., Crist, T. O., & Gering, J. C. (2002). The additive partitioning
718 of species diversity: recent revival of an old idea. *Oikos*, 99(1), 3–9.
- 719 Verboom, G. A., Linder, H. P., Forest, F., Hoffmann, V., Bergh, N. G., & Cowling, R. M.
720 (2014). Cenozoic assembly of the Greater Cape flora. In N. Allsopp, J. F. Colville, & G. A.
721 Verboom (Eds.), *Fynbos: Ecology, Evolution and Conservation of a Megadiverse Region*.
722 Oxford University Press.
723 [http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199679584.001.0001/acprof-](http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199679584.001.0001/acprof-9780199679584)
724 [9780199679584](http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199679584.001.0001/acprof-9780199679584)
- 725 Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California.
726 *Ecological Monographs*, 30(3), 279–338.
- 727 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Golemund,
728 G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M.,
729 Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019).
730 Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.
731 <https://doi.org/10.21105/joss.01686>
- 732 Wiens, J. J. (2004a). Speciation and ecology revisited: phylogenetic niche conservatism and the

733 origin of species. *Evolution*, 58(1), 193–197.

734 Wiens, J. J. (2004b). What is speciation and how should we study it? *The American Naturalist*,
735 163(6), 914–923.

736 Wüest, R. O., Boucher, F. C., Bouchenak-Khelladi, Y., Karger, D. N., & Linder, H. P. (2019).
737 Dissecting biodiversity in a global hotspot: Uneven dynamics of immigration and
738 diversification within the Cape Floristic Region of South Africa. *Journal of Biogeography*,
739 46(9), 1936–1947. <https://doi.org/10.1111/jbi.13625>

740

741 **Biosketches**

742 **Ruan van Mazijk** is a graduate student interested in macroevolution, macroecology and
743 phylogenetic comparative biology, primarily focusing on plants.

744 **Michael D. Cramer** is an ecophysiologist interested in physiological specialization in the hyper-
745 diverse Cape flora and the link between nutrient-impooverished soils and plant species richness.

746 **G. Anthony Verboom** works on the assembly of the hyper-diverse Cape flora, its vegetation
747 organization and the role of speciation and extinction in its radiation.

748 **Author contributions**

749 MDC and GAV conceived the study question, which RvM investigated and developed under
750 their supervision for his BSc Hons project. RvM collated the data and carried out the GIS work.
751 All authors contributed to the analyses, which were then carried out by RvM, who then wrote the
752 first draft of the manuscript. All authors contributed equally to the writing thereafter.

753 **ORCID**

754 Ruan van Mazijk: <https://orcid.org/0000-0003-2659-6909>

755 Michael D. Cramer: <https://orcid.org/0000-0003-0989-3266>

756 G. Anthony Verboom: <https://orcid.org/0000-0002-1363-9781>