LETTER

Cichlid species-area relationships are shaped by adaptive radiations that scale with area

Catherine E. Wagner,[1,2,3]* Luke J. Harmon[4,5] and Ole Seehausen[1,2]

Abstract

A positive relationship between species richness and island size is thought to emerge from an equilibrium between immigration and extinction rates, but the influence of species diversification on the form of this relationship is poorly understood. Here, we show that within-lake adaptive radiation strongly modifies the species-area relationship for African cichlid fishes. The total number of species derived from *in situ* speciation increases with lake size, resulting in faunas orders of magnitude higher in species richness than faunas assembled by immigration alone. Multivariate models provide evidence for added influence of lake depth on the species-area relationship. Diversity of clades representing within-lake radiations show responses to lake area, depth and energy consistent with limitation by these factors, suggesting that ecological factors influence the species richness of radiating clades within these ecosystems. Together, these processes produce lake fish faunas with highly variable composition, but with diversities that are well predicted by environmental variables.

Keywords

Adaptive radiation, cichlid, diversity-dependent diversification, ecological limits, island biogeography, species-area relationship.

Ecology Letters (2014) **17**: 583–592

INTRODUCTION

Exceptional species richness within geographically restricted clades is a hallmark of adaptive radiation. Although understanding the factors influencing geographical patterns of species richness is a fundamental goal of ecological research, adaptive radiations have rarely been studied from macroecological perspectives. Decades of ecological research have uncovered strong and predictable relationships between ecological variables (e.g. area, productivity) and species richness (MacArthur & Wilson 1963; Gaston 2000). These same variables may influence processes of species origination and the occurrence and richness of adaptive radiations. For example, habitat size may influence *in situ* speciation and adaptive radiation (e.g. Losos & Schluter 2000; Kisel & Barraclough 2010; Rabosky & Glor 2010). In addition, adaptive radiations are often characterised by slow-downs in diversification rate through time (e.g. Schluter 2000), assumed to derive from niche-filling processes (e.g. Phillimore & Price 2008; Rabosky & Lovette 2008). If adaptive radiations have reached this slowdown phase, there should be clear relationships between species richness and environmental variables related to resource availability, if these variables limit species richness. However, rarely have such environmental variables been identified for adaptive radiations.

In classic island biogeography theory, MacArthur & Wilson (1963) viewed species richness as an equilibrium between immigration and extinction rates, and treated newly evolved species as equivalent to new immigrants to the system. The implicit assumption is that species derived from these different sources would not influence equilibrium diversity in distinct ways. However, empirical evidence and theoretical modelling suggest that speciation within island systems results in a steeper rate of increase in species richness with area than community assembly purely by dispersal (Losos & Schluter 2000; Triantis *et al.* 2008; Rabosky & Glor 2010; Rosindell & Phillimore 2011). Furthermore, regions wherein the majority of species have arisen *in situ* ('biological provinces') are well known to have steeper species-area relationships than regions in which immigration and emigration play a proportionally larger role in determining species richness (Rosenzweig 1995).

Although in its simplest formulation MacArthur & Wilson's (1963) equilibrium theory posits that the relationship between extinction rate and island size is the main driver of the species-area relationship, another likely influence on this pattern is area's function as a proxy for resource availability. If so, other variables linked to resource availability, or those that more precisely approximate it, should explain additional variation in species-area relationships (MacArthur & Wilson 1967). Such variables are numerous and include elevation and/or direct metrics of habitat diversity (e.g. Ricklefs & Lovette 1999), and energy availability (reviewed in Evans *et al.* 2005). Island age might also act to limit species richness if islands are young, an effect commonly predicted in isolated

[1]*Department of Fish Ecology & Evolution, EAWAG, Centre for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland*
[2]*Department of Aquatic Ecology, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland*
[3]*Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA*

[4]*Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA*
[5]*Institute for Bioinformatics and Evolutionary Studies (IBEST), University of Idaho, Moscow, ID 83844, USA*
Correspondence: E-mail: catherine.wagner@eawag.ch

systems where diversity arises predominantly via *in situ* speciation (e.g. Parent & Crespi 2006).

Lakes, as ecosystems with well-circumscribed boundaries, are an aquatic equivalent to oceanic islands. There is evidence for positive species-area relationships in freshwater fish faunas in large North American and African lakes (Barbour & Brown 1974), and in smaller temperate lake systems (Magnuson *et al.* 1998). However, species-area patterns in fishes have not been studied with regard to the influence of *in situ* speciation.

Evolutionary diversification of cichlid fishes is widespread across the African continent. The East African Rift Lakes host the most species-rich radiations (Tanganyika: ~ 250 spp.; Malawi: 451–600 spp.; Victoria: 447–535 spp.; Genner *et al.* 2004), but diversification has also occurred in lakes from Cameroon to the Eastern Rift, and from Namibia to the Middle East (Fig. 1). Multiple cichlid lineages often coexist and diversify within the same lake (Wagner *et al.* 2012). The composition of lake cichlid faunas is therefore a variable assemblage of species derived from immigration and intralacustrine radiation (Fig. 1). Previous work suggests that a combination of environmental variables and species traits best predict whether or not cichlid lineages undergo adaptive radiation upon colonising lakes (Wagner *et al.* 2012). Consistently important environmental predictors of radiation were lake

depth, age and energy availability. We here ask a distinct, yet complementary set of questions: what predicts the species richness of lake cichlid faunas, and the species richness of lineages within lakes that undergo adaptive radiation?

We focus on four variables and their relationship with cichlid species richness: area, depth, energy, and time. We examine the relationships between these variables and the richness of lake communities resulting from adaptive radiation, and the total richness of cichlid assemblages. We ask the following questions: (1) What is the shape of the species-area curve, and how does *in situ* speciation influence it? (2) For total species richness, what is the relationship between species richness and depth, energy and lake age? (3) For intralacustrine radiations, what is the relationship between species richness and depth, energy, and clade age? Together, these questions seek to identify important environmental factors influencing species richness in lake species assemblages, and to understand the influence of adaptive radiation on the shape of species richness-environment relationships.

MATERIAL AND METHODS

We compiled information about the cichlid species present in 46 African lakes from published information (Genner *et al.* 2004 and references therein; Lamboj 2004; Seehausen 2006
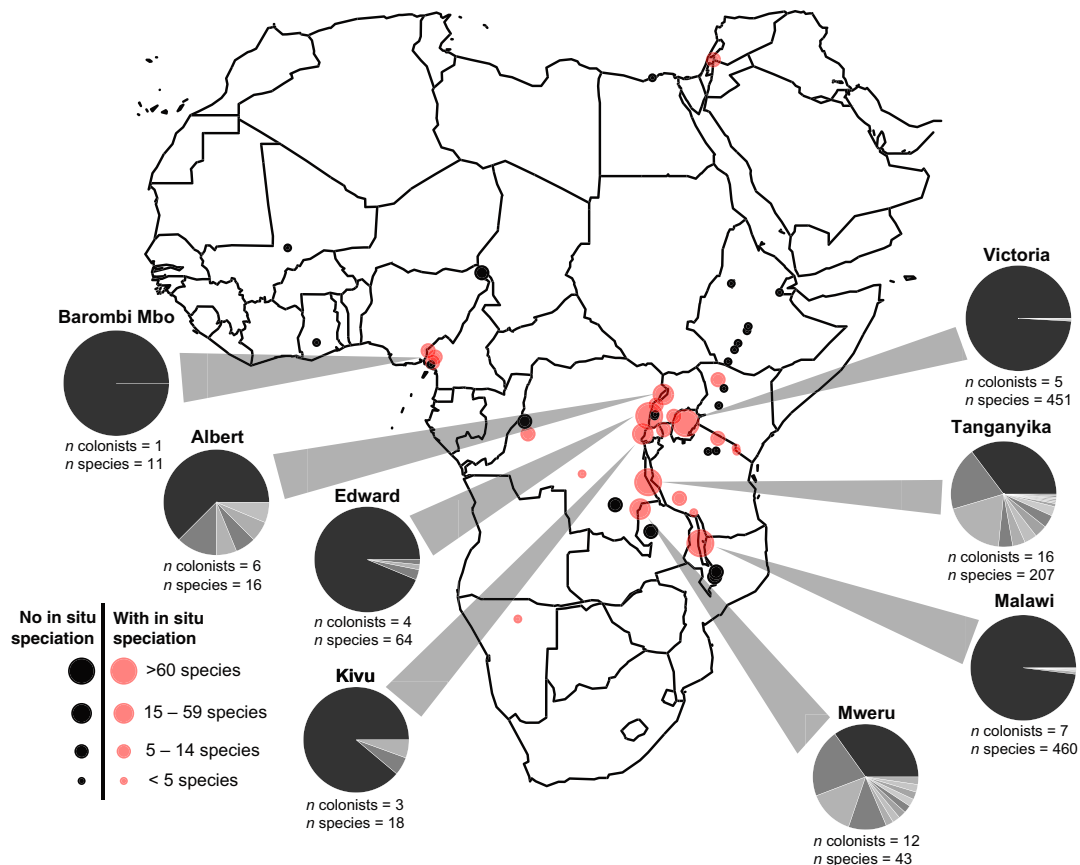

Figure 1 Species richness of the cichlid faunas of 46 African lakes. Dots on the map correspond to lakes, where red indicates lakes with *in situ* speciation and black indicates lakes that have accumulated species richness only via colonisation. Pie charts, shown for selected lakes with radiations, represent how diversity within lakes is distributed among cichlid lineages, where the richness of each lineage present within a lake is shown in a different shade. Lakes accumulate richness both through dramatic diversification of single lineages (e.g. Lakes Malawi, Victoria and Edward) or via accumulation of diversity in many lineages (e.g. Lakes Tanganyika, Mweru).

and references therein), FishBase records (Froese & Pauly 2010), and our own data. We calculated stem and crown clade ages (*sensu* Magallon & Sanderson 2001) for lineages that diversified via within-lake speciation using ultrametric molecular phylogenies of African cichlids (for phylogenetic methods, see Appendix S1 and Wagner *et al.* 2012). Phylogenies included ~ 650 of ~ 2000 species.

We collated information on area, depth, energy and age for lakes. If published records of surface area were unavailable, we measured it from Google Earth images using ImageJ (Rasband 1997–2011). We gathered information about lake ages and/or time since last desiccation from a variety of sources (Table S1). For energy, we used the average of 2010 monthly values of net solar radiation (net flux), from the NASA Langley Research Center ASDC.

Fish diversity in lakes is often higher in the littoral zone than in the pelagic zone (Vadeboncoeur *et al.* 2011), and thus perimeter might better approximate habitat availability than area. However, because of the strong correlation between area and perimeter for these lakes ($r^2 = 0.97$; Fig. S1), these variables perform equivalently. We here use area, consistent with previous studies in island biogeography.

General analytical framework

We examine patterns of species richness at two scales (Fig. 2a). First, as in traditional island biogeography studies, we examine the total per-lake species richness. Total richness is the sum of the diversity of radiations and colonists. Second, we examine patterns of species richness among clades that have diversified by *in situ* speciation within single lakes ('radiations'). We considered as 'radiations' any case where species have arisen via *in situ* cladogenesis within lakes, including any endemic species co-occurring with its sister species, and all fully endemic clades inhabiting a single lake. Single endemic species not co-occurring with a sister taxon were treated as equivalent to non-endemic species. To examine how *in situ* speciation influences the species-area relationship (e.g. Losos & Schluter 2000; Rabosky & Glor 2010; Rosindell & Phillimore 2011), we also analysed the total richness of communities in the subset of lakes with *in situ* speciation, and the difference between the colonising lineages-area relationship (CAR) and the species-area relationship (SAR) (details below).

We used regression models to test for relationships between species richness and environmental variables. Our general approach (Fig. 2b) was to first fit linear regression models and calculate correlation coefficients for richness and environmental variables. We fit quantile regression models for radiations (see details below). In addition, because SARs have been shown to take a two-slope form in systems where *in situ* speciation contributes significantly to richness (e.g. Losos & Schluter 2000), we fit two-slope species-area regression models (details below).

Testing for factors potentially limiting species richness

'Limiting factors' in ecology are those where a measured factor imposes a 'ceiling' to a response distribution, but where variance is high under these maximum values due to the influ-
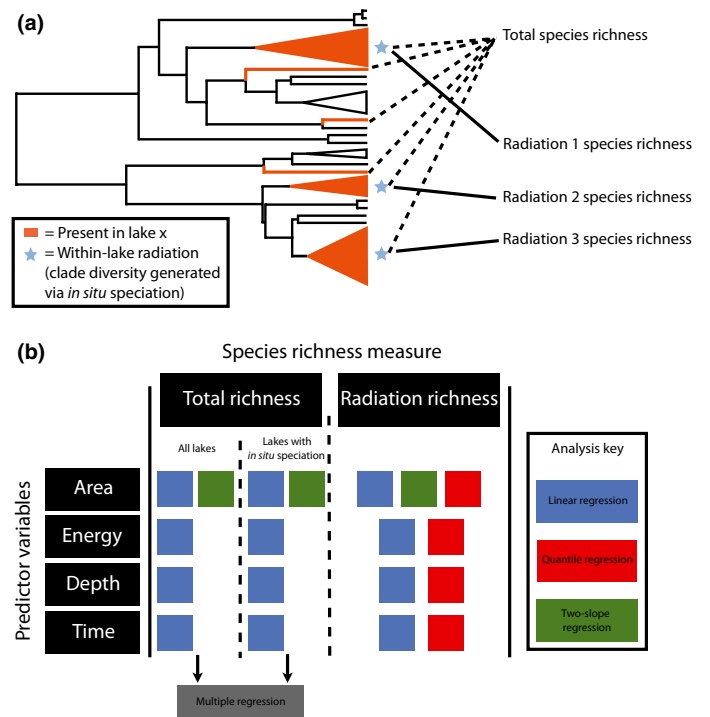

Figure 2 (a) Schematic depicting the difference between total species richness and the species richness of intralacustrine radiations. Total species richness is all cichlid species present within a single lake; some species may have arisen via *in situ* speciation, others are present via immigration into the lake. Radiations are clades that have diversified entirely by *in situ* speciation within single lakes. For African cichlids, it is not uncommon for multiple radiations to have occurred within a single lake. (b) Schematic summarising the analyses performed for assessing relationships between predictor variables and both total species richness and radiation richness.

ence of other unmeasured factors (Thomson *et al.* 1996; Cade *et al.* 1999). We expect environmental variables acting as 'limiting factors' to place an upper bound on species richness, but for no relationship to exist between the predictor variable and lower values of species richness. Quantile regression can test for such relationships between a measured factor and the upper bounds of a response distribution (Chassot *et al.* 2010). We calculated every 5th linear regression quantile from 5 to 95% using the R package quantreg, and estimated standard error with method 'iid' (Koenker 2011).

Species-area relationships

We plotted the SAR for total richness, and compared it to the relationship between area and the number of colonising lineages (the CAR), which we can interpret as the SAR if there were no *in situ* speciation (i.e. immigration-based community assembly). We fitted models with a linear relationship between species richness (S) and surface area (A), the log-transformed equivalent of the well-known power model (Lomolino 2000),

$$\log(S) = c + z_1 \log(A),$$

where c is the intercept of the species-area curve and z_1 is its slope. We compared SARs for immigration-assembled

communities to those for speciation-assembled communities with ANOVA.

Next, to test for an increase in the slope of the SAR for lakes above a threshold size, as in Losos & Schluter (2000), we compared the linear SAR to a model fitting species richness as a function of lake area in a two-slope regression framework. We used the formulation from Losos & Schluter (2000),

$$\log(S) = c + z_1 \log(A) + z_2 [\log(A) - t]\delta,$$

where c is the intercept, z_1 is the slope of the line before the breakpoint, z_2 is the slope of the line after the breakpoint, t is the position of the breakpoint and δ is an indicator variable with a value of 1 if $\log(A) \geq t$ and 0 otherwise. We fit models using maximum likelihood estimation of non-linear least squares parameter estimates, using the function nls in R (R Development Core Team 2011).

To investigate the relationship between *in situ* speciation and lake size, we calculated the proportion of species arising via *in situ* speciation per lake (the 'speciation fraction', Losos & Schluter 2000).

Combined effects of environmental factors on species richness

We examined the combined effects of environmental variables on total species richness by testing among multivariate models including all combinations of predictor variables. First, we assumed linear relationships for all variables and used AICc-based model averaging (Burnham & Anderson 2002) to identify the most important variables from the set of all possible models. We used the Akaike weight of each model to calculate relative importance (RI) as the sum of relative Akaike weights for models in which that given predictor variable appears. RI scales from 0 (not important) to 1 (very important). We included predictor variables with RI above 0.70 in the second step of analysis. We fit models using the function glm with Gaussian error in R (R Development Core Team 2011).

Second, using the reduced variable set, we tested models including both linear and two-slope formulations for species-area effects. To test for effects of ecosystem size on species richness, we included a model for lake volume. Because data on lake volume are sparse, we used surface area multiplied by maximum depth as an approximation. We fit models using the function nls in R (R Development Core Team 2011) and used AICc scores to test among models.

Prior to multivariate analyses we tested for collinearity between predictor variables. We repeated all analyses excluding lake age as a predictor, because data for this variable are sparse.

Phylogenetic signal and diversification rates

In analyses of radiation richness, we tested for phylogenetic signal in species richness and in the residuals of the relationships between richness and area, depth and energy (see Appendix S1).

To test for correlations between diversification rate and environmental variables, we calculated net diversification rate using Magallon & Sanderson's (2001) method implemented in geiger (Harmon *et al.* 2008). We calculated rates based on species richness and median crown clade ages from both relative and time-calibrated trees. We repeated calculations for four values of ε, the extinction rate expressed as a fraction of the speciation rate: 0, 0.1, 0.5 and 0.9. We then correlated diversification rate with area, depth and energy.

RESULTS

Total species richness varies substantially across the 46 lakes, and half of them have intralacustrine speciation (Fig. 1).

Species-area relationships

There is a positive correlation between lake size and the number of colonising lineages (CAR; $r^2 = 0.296$, $P < 0.0005$), and quantile regression suggests that lake size behaves as a 'limiting factor' to colonisation (Fig. S3). The total species richness of lakes with *in situ* speciation is on average an order of magnitude higher than the number of colonising lineages (Fig. 3a). The SAR has a significant positive correlation ($r^2 = 0.195$, $P < 0.005$), which increases when we include only lakes with *in situ* speciation ($r^2 = 0.51$, $P < 0.0005$). The slope of the SAR for lake communities with *in situ* speciation is steeper than that for lakes with immigration-based community assembly (Fig. 3a; $P < 0.05$), and the intercept of the SAR for *in situ* speciation differs from 0, whereas the intercept for the immigration-only SAR does not (*in situ* speciation: $z = 0.250$, $P = 0.0001$, $c = 0.599$, $P = 0.0006$; immigration: $z = 0.132$, $P < 0.0001$, c = 0.113, $P = 0.173$).

Fitting two-slope regressions to the SAR provides support for a two-slope model compared to a linear model for the full data set ($\Delta AICc = 28$; Fig. 3b). The pre-breakpoint slope does not differ from zero ($P = 0.411$), the breakpoint occurs at 1030 km^2 and the post-breakpoint slope is positive (1.289, $P < 0.00001$). For the subset of lakes with *in situ* speciation, the two-slope regression model also outperforms the linear model ($\Delta AICc = 9$), and again the pre-breakpoint slope does not differ from zero ($P = 0.444$), the breakpoint falls at a threshold of 1470 km^2 and the post-breakpoint slope is positive (0.985, $P < 0.001$). For the subset of lakes without *in situ* speciation a one-slope regression model significantly outperforms a two-slope regression model ($\Delta AICc = 5.5$).

Lakes in which *in situ* speciation occurred span the full size range of lakes studied (Fig. 3c).

For individual radiations, there is a positive correlation between richness and lake area (RAR: $r^2 = 0.22$, $P = 0.003$). Linear and two-slope regression models perform equivalently ($\Delta AICc$ 0). In the two-slope model, the breakpoint occurs at lake size 2225 km^2, the pre-breakpoint slope does not differ from zero ($P = 0.76$) and the post-breakpoint slope is positive (0.529, $P = 0.04$). As lake area increases, so does the maximum number of species observed within radiations (Fig. 3d). In quantile regression, the estimated slope generally increases with increasing quantiles (Fig. S4). There is a positive correlation between lake size and the number of species that have arisen via *in situ* speciation ($r^2 = 0.36$, $P = 0.002$).
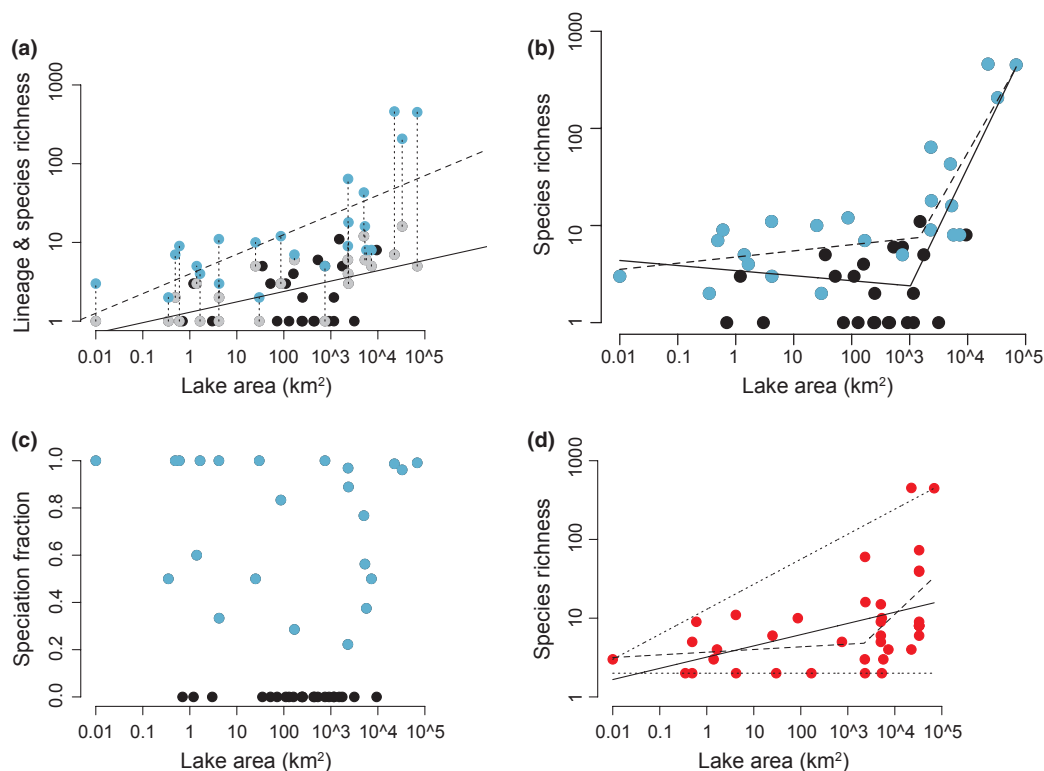
Figure 3 The species-area relationship is strongly modified by adaptive radiation (blue = lakes with speciation; black = lakes without speciation; red = individual radiations). (a) The species-area relationship for lakes with speciation (dashed line) is significantly steeper than that for immigration-based community assembly (solid line). Dotted lines show the change in species richness due to within-lake speciation (grey = number of colonists to lakes with speciation). (b) Two-slope species-area models indicate an increase in the slope of the species-area curve in lakes above 1030 km^2 (solid line = whole data set; dashed line = lakes with speciation only). (c) The proportion of species arising *in situ* ('speciation fraction') is greater than zero in lakes spanning the size range studied. (d) Species richness of radiations exhibits a "limiting" relationship with lake area (solid and dashed lines = linear and two-slope regression models; dotted lines = 5th and 95th quantiles of the linear model).

Species richness and lake depth, energy and time

There are significant correlations between total species richness and depth ($r^2 = 0.33$, $P < 0.0005$), energy ($r^2 = 0.095$, $P < 0.05$), and lake age ($r^2 = 0.21$, $P < 0.05$; Fig. 4a–c). When we exclude lakes without *in situ* speciation, these relationships remain for all three variables (depth: $r^2 = 0.33$, $P < 0.05$; energy: $r^2 = 0.18$, $P < 0.05$; age: $r^2 = 0.23$, $P < 0.05$). There are no significant relationships between any of these variables and the number of colonising lineages (Fig. S5).

For individual radiations, energy and depth produced significant albeit weak positive correlations (depth: $r^2 = 0.18$, $P < 0.05$; energy: $r^2 = 0.25$, $P < 0.05$; Fig. 4d,e). There is no evidence for correlation between relative clade age and richness (stem ages: $r^2 = 0.03$, $P = 0.39$; crown ages: $r^2 = 0.02$, $P = 0.47$; Fig. 4f, Fig. S2). Quantile regressions show a general positive increase in estimated slope as quantiles increase from 5 to 95% for both energy and depth (Fig. S4). There is no consistent change in quantile slope for the clade age-richness relationship.

Combined effects of environmental factors on species richness

All pairwise correlations between predictor variables had correlation coefficients less than 0.33 except for that

between depth and age ($r = 0.66$, 0.67, respectively, for all lakes and for those with speciation; Fig. S6). Multiple regressions revealed area and depth as the most important predictors of species richness. For the data set including all lakes, area and depth had substantially higher RI scores than energy and lake age (Table 1). When lake age is excluded, depth becomes the most important predictor of richness, followed by area; energy is a poor predictor of richness (RI = 0.07). For lakes with speciation, area alone is a strong predictor of richness when all four variables are considered (RI = 0.91). When lake age is excluded, area and depth are both strong predictors (RI = 0.99 and 0.94 respectively), whereas energy is a poor predictor (RI = 0.02).

We included depth, one-slope and two-slope formulations of the area effect, and approximate lake volume, in additional model testing. Among these, the best model was one that included a two-slope area term plus lake depth (Table 2). For the data set including all lakes, models including two-slope area with and without depth performed best, and were almost equivalent (ΔAICc = 1); all other models performed substantially worse. When including only lakes with speciation, these same two models performed almost equivalently, as did the linear area and depth model and the approximate volume model (ΔAICc = 2–3).
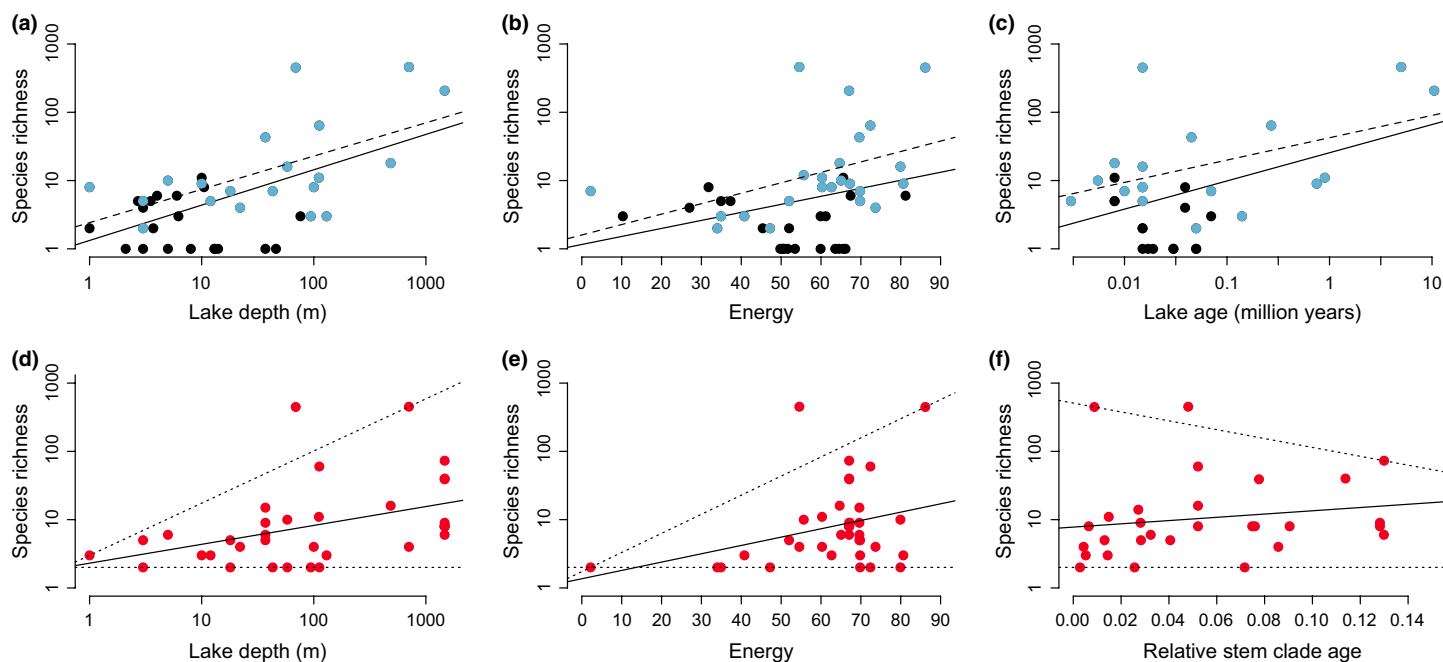
Figure 4 There are significant positive correlations between species richness of cichlids in African lakes and (a) lake depth ($r^2 = 0.330$, $P < 0.0005$), (b) energy ($r^2 = 0.095$, $P < 0.05$), and (c) lake age ($r^2 = 0.210$, $P < 0.05$). For individual cichlid radiations (panels d–f), species richness shows evidence of limitation by both (d) depth and (e) energy. For panels (a–c), solid lines are the result of linear regression for the entire data set; dashed lines are those for the subset of lakes wherein there is speciation (blue dots). For panels (d–f), solid lines are the results of linear regression; dotted are the 5th and 95th quantiles of these data sets. (f) For cichlid radiations, there is no relationship between clade age and species richness ($r^2 = 0.022$, $P = 0.429$), nor is there evidence for clade age-based constraint on lineage species richness.

Table 1 Multiple regression analyses provide evidence for depth in addition to surface area as important predictors of species richness of the cichlid faunas of African lakes. Because estimates of lake age are relatively sparse, we fit models with (left columns in each lake set) and without (right columns in each lake set) this variable. RI = Relative Importance

	All Lakes						Lakes with Speciation					
	($n = 29$)			($n = 40$)			($n = 16$)			($n = 20$)		
Variable	RI	Estimate	SE	RI	Estimate	SE	RI	Estimate	SE	RI	Estimate	SE
SA	**0.82**	0.17	0.07	**0.89**	0.16	0.05	**0.91**	0.21	0.07	**0.99**	0.22	0.05
Depth	**0.70**	0.29	0.10	**1.00**	0.48	0.11	0.32	0.10	0.05	**0.94**	0.36	0.11
Energy	0.49	0.01	0.00	0.07	0.00	0.00	0.54	0.01	0.00	0.02	0.00	0.00
Age	0.45	0.15	0.07				0.61	0.19	0.07			

Phylogenetic signal and diversification rates

No K values differed significantly from zero in tests for phylogenetic signal (Table S2, Fig. S8). This implies that phylogenetic non-independence does not confound estimates of richness-environment relationships, and that these relationships cannot be explained by phylogenetic effects.

There were no strong correlations between diversification rate and area, depth or energy (Fig. S7, Table S3).

DISCUSSION

In situ speciation is known to influence the shape of species-area curves both empirically and in theory (Rosenzweig 1995;

Losos & Schluter 2000; Rabosky & Glor 2010; Rosindell & Phillimore 2011). We here show that *in situ* speciation strongly modifies the species-area curve for cichlid fish in African lakes, producing lake faunas that are on average an order of magnitude higher in species richness than faunas assembled by immigration alone. Speciation happens in lakes spanning the size range of lakes we studied (Fig. 3c), but the number of species derived from *in situ* speciation increases with increasing lake size. This produces a significantly steeper species-area curve than that representing immigration-based community assembly alone (Fig. 3a). We also find significant correlations between species richness and lake depth, energy and lake age. Multivariate analyses provide the strongest support for models including both lake area and depth as predictors of species richness.

Table 2 A two-slope area plus depth model outperforms other models which include area, depth, and volume effects in predicting species richness of cichlid lake faunas. The same model performs best both for the data set including all lakes, and for that with only lakes with *in situ* speciation

Model	Formula	All Lakes		Lakes with Speciation	
		AICc	ΔAICc	AICc	ΔAICc
Linear Area	$\log(S) = c + z_1\log(A)$	81.01	29	34.55	10
Two-slope Area	$\log(S) = c + z_1\log(A) + z_2[\log(A) - t]\delta$	52.33	1	26.18	2
Two-slope Area + Depth	$\log(S) = c + z_1\log(A) + z_2[\log(A) - t]\delta + z_3\log(D)$	51.75	0	24.63	0
Depth	$\log(S) = c + z_1\log(D)$	72.92	21	40.11	15
Area + Depth	$\log(S) = c + z_1\log(D) + z_2\log(A)$	66.12	14	27.75	3
Volume	$\log(S) = c + z_1\log(AD)$	69.7	18	26.44	2

Complementary to patterns of species richness at the level of entire lake faunas, area, depth and energy show evidence for being 'limiting factors' (*sensu* Cade *et al.* 1999) to the richness of individual cichlid radiations. We find no relationship between diversity and clade age. These results together suggest that local ecological factors shape the global patterns in species diversity that we observe. We discuss each of these key findings in detail below.

In our data set, the SAR for lakes with *in situ* speciation has a steeper slope (z) and a significantly elevated intercept (c) compared to that for immigration-based species richness alone (CAR; Fig. 3a). This indicates that (1) the richness of faunas assembled through speciation is greater than that for immigration-assembled communities in lakes of equivalent size, and that (2) richness increases with area at a steeper rate in lakes with speciation. These differences likely arise from distinct sources. First, richness arising from dispersal assembly is limited by the diversity of the regional species pool. With the exception of some rivers in southern Africa (Joyce *et al.* 2005), African riverine cichlid diversity is low compared to that in lakes (Seehausen 2006), and thus *in situ* speciation would be expected to increase the intercept of the SAR. Second, evidence from a number of systems suggests that species richness increases with area at a steeper rate in communities assembled by evolution (i.e. within which there is speciation) than those assembled purely by dispersal (Rosenzweig 1995; Triantis *et al.* 2008; Rabosky & Glor 2010), as we observe here. Although this pattern has been noted phenomenologically in other systems and attributed to the effects of *in situ* evolution, its mechanistic basis is poorly understood. Candidate mechanisms include tighter niche packing in communities assembled by evolution (Rabosky & Glor 2010), or simply the effects of increased speciation rates in larger areas (Losos & Schluter 2000; Triantis *et al.* 2008).

We also find support for a two-slope species-area relationship compared to a linear model. However, unlike previous work that has theorised breakpoint models in evolutionary species-area relationships (Losos & Schluter 2000; Losos & Parent 2010), in cichlid fish the threshold does not correspond to the area at which within-lake speciation begins. African cichlids speciate within the smallest (< 1 km^2), and the largest (Lake Victoria, 68 800 km^2) lakes in our data set (Fig. 3c), and the number of species that has arisen by *in situ* speciation increases with lake size (Fig. 3a). Although the onset of *in situ* speciation cannot explain the two-slope species-area relationship for cichlids, the addition of species via different modes of

speciation might contribute to this pattern. Evidence for geographical speciation in cichlids in large lakes comes from evidence of fine-scale genetic differentiation (Allender *et al.* 2003; Wagner & McCune 2009), and the existence of geographically distributed sister species within large African lakes (Seehausen & van Alphen 1999; Genner *et al.* 2004). There is equally strong evidence for speciation without geographical differentiation in cichlids, including some of the best-accepted examples of sympatric speciation (e.g. Schliewen *et al.* 1994; Barluenga *et al.* 2006), and certain mechanisms of cichlid speciation operate at very small spatial scales (e.g. Seehausen *et al.* 2008). If rates of allopatric speciation increase with area, and non-geographical speciation happens consistently at all lake sizes, the interaction between the addition of species to the system via these different modes of speciation, and their cumulative influence on diversification rates, could create a non-linear species-area relationship.

Our data allow us to address questions related to both the factors influencing the diversity of individual radiations, and those influencing the diversity of evolutionarily assembled communities. We find that in communities where *in situ* speciation takes place, total richness is positively correlated with lake area, depth, energy and age. Multivariate models provide the best support for area and depth as richness predictors (Tables 1 and 2). In addition, we find that these same three environmental variables show patterns consistent with being 'limiting factors' (*sensu* Cade *et al.* 1999), of the richness of individual radiations, whereas there is no relationship between clade age and species richness (Fig. 4d–f). Both of these findings are consistent with the idea that ecological limits to diversification exist that render community richness predicted by, and radiation richness limited by, the same environmental factors.

Positive relationships between species richness and environmental variables in communities with *in situ* speciation could result from correlations between diversification rate and environmental variables (e.g. Losos & Schluter 2000). Two pieces of evidence in our analyses conflict with this perspective. First, we find no relationship between clade age and species richness (Fig. 4f; Fig. S2), suggesting limits to clade diversity (Rabosky 2009), and/or complex diversification histories, such as episodes of increased extinction rate and/or wide variation in rate among clades. Because cichlid diversification can be rapid (e.g. 450+ species in 15 000 years for Lake Victoria, Seehausen 2006), if diversification is diversity-dependent, it is plausible that signatures of the effect of age on species richness

could decay rapidly. Second, diversification rates calculated from clade age and richness data, which assume unconstrained diversification, are not correlated with environmental variables (Fig. S7, Table S3). This suggests that the relationships between individual radiation richness and environmental variables are not driven by simple systematic differences in current diversification rate.

In Appendix S2, we show that simple models of diversity-dependent diversification make interesting predictions for the shape of species-area curves when area influences initial speciation rate. In particular, the combination of both area-dependent and area-independent components of speciation rate can interact to create non-linear species-area relationships similar to the one that we observe here, even with linear scaling of maximum carrying capacities (see Appendix S2 for details).

In contrast with the lack of a clade age-richness relationship, the total richness of cichlid communities in African lakes is correlated with lake age (Fig. 4c). This suggests that community richness may be limited by the rate at which diversity arises (despite evidence that this is not the case for individual clades). However, correlation between lake depth and age ($r^2 = 0.44$; Fig. S6) makes interpretation of this pattern ambiguous. In multivariate models including both depth and age, depth is a much stronger predictor of richness than is age (Table 1). This suggests a stronger role for depth's influence than time. For fish and many aquatic organisms, depth likely mediates habitat availability and heterogeneity, in a manner analogous to elevation as a metric of habitat diversity for terrestrial islands (e.g. Ricklefs & Lovette 1999). Nevertheless, our analyses also hint at an effect of age. If the waiting time until radiation varies among colonisation events, but on average, time increases the probability of observing radiation, we would expect a correlation between total faunal richness and lake age, but not between clade age and richness, as we observe. This is consistent with results of previous work on the factors influencing the occurrence of radiation (Wagner *et al.* 2012), although there, too, collinearity between lake age and depth makes disentangling these effects challenging.

Although the richness of individual radiations exhibits a 'limiting' response distribution to energy (Fig. 4e), and there is a weak but significant correlation between energy and species richness for total richness (Fig. 4b; $r^2 = 0.095$, $P < 0.05$), energy is not an important predictor of total richness in multivariate models (Table 1). An explanation for this pattern is the relative variance of energy versus area in our data set. Despite substantial variance in energy among lakes, this is dwarfed by the extent of variance in lake area. Wright (1983) suggested that the product of energy and area ('total energy'), could replace area in MacArthur & Wilson's (1963) island biogeography models, under the logic that it more succinctly and generally approximates resource availability. In our data set, disparity in the total variance in energy versus that of area means that the product of these variables is very strongly correlated with area alone ($r^2 = 0.96$). Therefore, total energy may be an important predictor of species richness in these systems, but its effect cannot be separated from the effect of area.

Lake area and depth are the strongest predictors of species richness (Table 1). A role for both of these variables might reflect an influence of lake volume. However, the best performing model for both the full set of lakes and those with speciation includes a two-slope area-richness relationship, in addition to a linear term for depth (Table 2). That the simpler volume models do not perform better than more complex models suggests that depth and area influence species richness in distinct ways.

In previous work analysing the factors predicting the occurrence of cichlid radiation (Wagner *et al.* 2012), we used Poisson hurdle regression to ask if the environmental variables used here, in addition to several lineage-specific traits, predict the occurrence and the richness of lineages that radiate in these lakes. Although we identified lake depth, energy and the trait of sexual dichromatism as significant predictors of the occurrence of radiation, no variables were significant predictors of the species richness of radiating clades. Examination of the distribution of species richness in response to these environmental variables (e.g. Figs 3d and 4d–f) clarifies this result. Quantile regression analyses indicate that these variables are not good predictors, but may be 'limiting factors' (sensu Cade *et al.* 1999) to radiation richness. Although upper regression quantiles have a strong positive slope, medium and low-richness radiations occur even in large, deep and high-energy lakes. Such a relationship is perhaps not unexpected given the numerous influences on the evolution of clade richness, including both ecological and evolutionary factors such as resource availability, species interactions, diversification rate differences, complex diversification histories and priority effects. In addition, although direct influences of competition among lineages on diversification are difficult to detect in broad-scale comparative analyses, some patterns in lineage diversity suggest that among certain cichlid clades, competition strongly impacts diversification (Seehausen 2007). Hence, whereas our analyses indicate that the maximum observed values of radiation richness are predicted by environmental variables with known macroecological relevance, realised diversity of individual radiations is influenced by many other variables. More generally, these results together with the results of previous work (Wagner *et al.* 2012) suggest that the factors that predict whether or not cichlid radiation happens are distinct from those influencing species richness.

If adaptive radiations experience slowdowns in diversification rate through time due to niche-filling processes (e.g. diversity-dependent diversification), clade richness should ultimately stabilise at a level reflective of the ecological adaptations of that clade and the resources and environment of the region within which the clade has diversified (Ricklefs 2009; Rabosky 2013). Evidence from the fossil record suggests that after an initial period of increase, clades often reach and remain at steady-state diversities, reflecting diversity-dependent logistic growth (e.g. Alroy 2010). Although molecular phylogenies of extant taxa cannot directly record such a process, its signatures can be tested for (e.g. Morlon *et al.* 2010). Recent studies provide evidence for better fits of diversity-dependent models than unlimited diversity models to species richness data (Rabosky 2009; Rabosky *et al.* 2012), and diversity-dependent models produce estimates of diversification parameters which are better aligned with those known from the fossil record (Etienne *et al.* 2012).

This study shows that species-rich cichlid faunas can be assembled either via dramatic diversification of single lineages or via more moderate diversification of multiple lineages within a lake (Fig. 1). Both scenarios produce faunas that follow a well-supported species-area pattern, but *in situ* speciation strongly alters the slope of this relationship compared to that for colonising lineages alone, generating communities that are an order of magnitude more species-rich. More generally, we argue that environmental factors influence patterns of species richness emerging in adaptive radiations by influencing the dynamics of diversification.

AUTHOR CONTRIBUTIONS

C.E.W., L.J.H. and O.S. designed the study. O.S. and C.E.W. collected the data. C.E.W. conducted the analyses. C.E.W., L.J.H. and O.S. contributed to writing.

REFERENCES

Allender, C.J., Seehausen, O., Knight, M.E., Turner, G.F. & Maclean, N. (2003). Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proc. Natl Acad. Sci. USA*, 100, 14074–14079.

Alroy, J. (2010). The shifting balance of diversity among major marine animal groups. *Science*, 329, 1191–1194.

Barbour, C.D. & Brown, J.H. (1974). Fish species diversity in lakes. *Am. Nat.*, 108, 473–489.

Barluenga, M., Stolting, K.N., Salzburger, W., Mushick, M. & Meyer, A. (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, 439, 719–723.

Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodel Inference*. Springer, New York.

Cade, B.S., Terrell, J.W. & Schroeder, R.L. (1999). Estimating effects of limiting factors with regression quantiles. *Ecology*, 80, 311–323.

Chassot, E., Bonhommeau, S., Dulvy, N.K., Melin, F., Watson, R., Gascuel, D. *et al.* (2010). Global marine primary production constrains fisheries catches. *Ecol. Lett.*, 13, 495–505.

Etienne, R.S., Haegeman, B., Stadler, T., Aze, T., Pearson, P.N., Purvis, A. *et al.* (2012). Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. Biol. Sci.*, 279, 1300–1309.

Evans, K.L., Warren, P.H. & Gaston, K.J. (2005). Species-energy relationships at the macroecological scale: a review of the mechanisms. *Biol. Rev.*, 80, 1–25.

Froese, R. & Pauly, D. (2010). FishBase. Online publication. Available at: http://www.fishbase.org. Last accessed 1 May 2010.

Gaston, K.J. (2000). Global patterns in biodiversity. *Nature*, 405, 220–227.

Genner, M.J., Seehausen, O., Cleary, D.F.R., Knight, M.E., Michel, E. & Turner, G.F. (2004). How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? *J. Biogeogr.*, 31, 93–102.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. (2008). GEIGER: investigating evolutionary radiations. *Bioinformatics*, 24, 129–131.

Joyce, D.A., Lunt, D.H., Bills, R., Turner, G.F., Katongo, C., Duftner, N. *et al.* (2005). An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature*, 435, 90–95.

Kisel, Y. & Barraclough, T.G. (2010). Speciation has a spatial scale that depends on levels of gene flow. *Am. Nat.*, 175, 316–334.

Koenker, R. (2011). Quantreg: Quantile Regression. R package version 4.71. Available at:http://CRAN.R-project.org/package=quantreg. Last accessed 1 June 2013.

Lamboj, A. (2004). The cichlid fishes of western Africa. Birgit Schmettkamp Verlag, Bornheim.

Lomolino, M.V. (2000). Ecology's most general, yet protean pattern: the species-area relationship. *J. Biogeogr.*, 27, 17–26.

Losos, J.B. & Parent, C.E. (2010). The speciation-area relationship. In: *The theory of island biogeography revisited* (Losos, J.B. & Ricklefs, R.E. eds). Princeton University Press, Princeton, pp. 415–438.

Losos, J.B. & Schluter, D. (2000). Analysis of an evolutionary species-area relationship. *Nature*, 408, 847–850.

MacArthur, R.H. & Wilson, E.O. (1963). Equilibrium-theory of insular zoogeography. *Evolution*, 17, 373–387.

MacArthur, R.H. & Wilson, E.O. (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton, NJ.

Magallon, S. & Sanderson, M.J. (2001). Absolute diversification rates in angiosperm clades. *Evolution*, 55, 1762–1780.

Magnuson, J.J., Tonn, W.M., Banerjee, A., Toivonen, J., Sanchez, O. & Rask, M. (1998). Isolation vs. extinction in the assembly of fishes in small northern lakes. *Ecology*, 79, 2941–2956.

Morlon, H., Potts, M.D. & Plotkin, J.B. (2010). Inferring the dynamics of diversification: a coalescent approach. *PLoS Biol.*, 8, e10000493.

Parent, C.E. & Crespi, B.J. (2006). Sequential colonization and diversification of Galapagos endemic land snail genus Bulimulus (Gastropoda, Stylommatophora). *Evolution*, 60, 2311–2328.

Phillimore, A.B. & Price, T.D. (2008). Density-dependent cladogenesis in birds. *PLoS Biol.*, 6, 483–489.

R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.

Rabosky, D.L. (2009). Ecological limits on clade diversification in higher taxa. *Am. Nat.*, 173, 662–674.

Rabosky, D.L. (2013). Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annu. Rev. Ecol. Evol. Syst.*, 44, 481–502.

Rabosky, D.L. & Glor, R.E. (2010). Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc. Natl Acad. Sci. USA*, 107, 22178–22183.

Rabosky, D.L. & Lovette, I.J. (2008). Density-dependent diversification in North American wood warblers. *Proc. Biol. Sci.*, 275, 2363–2371.

Rabosky, D.L., Slater, G.J. & Alfaro, M.E. (2012). Clade age and species richness are decoupled across the Eukaryotic tree of life. *PLoS Biol.*, 10, e1001381.

Rasband, W.S. (1997–2011). *ImageJ*. U.S. National Institutes of Health Bethesda, MD, USA.

Ricklefs, R.E. (2009). Speciation, extinction and diversity. In *Speciation and Patterns of Diversity*. (eds Butlin, R.K., Bridle, J.R., Schluter, D.). Cambridge University Press, Cambridge, pp. 257–277.

Ricklefs, R.E. & Lovette, I.J. (1999). The roles of island area per se and habitat diversity in the species-area relationships of four Lesser Antillean faunal groups. *J. Anim. Ecol.*, 68, 1142–1160.

Rosenzweig, M.L. (1995). Species diversity in space and time. Cambridge University Press, Cambridge.

Rosindell, J. & Phillimore, A.B. (2011). A unified model of island biogeography sheds light on the zone of radiation. *Ecol, Lett.*

Schliewen, U.K., Tautz, D. & Paabo, S. (1994). Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, 368, 629–632.

Schluter, D. (2000). *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proc. Biol. Sci.*, 273, 1987–1998.

Seehausen, O. (2007). Evolution and ecological theory - Chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. *Heredity*, 99, 361–363.

Seehausen, O. & van Alphen, J.M. (1999). Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.*, 2, 262–271.

Seehausen, O., Terai, Y., Magalhaes, I.S., Carleton, K.L., Mrosso, H.D.J., Miyagi, R. *et al.* (2008). Speciation through sensory drive in cichlid fish. *Nature*, 455, 620–626.

Thomson, J.D., Weiblen, G., Thomson, B.A., Alfaro, S. & Legendre, P. (1996). Untangling multiple factors in spatial distributions: lilies, gophers, and rocks. *Ecology*, 77, 1698–1715.

Triantis, K.A., Mylonas, M. & Whittaker, R.J. (2008). Evolutionary species-area curves as revealed by single-island endemics: insights for the inter-provincial species-area relationship. *Ecography*, 31, 401–407.

Vadeboncoeur, Y., McIntyre, P.B. & Vander Zanden, M.J. (2011). Borders of biodiversity: life at the edge of the World's large lakes. *Bioscience*, 61, 526–537.

Wagner, C.E. & McCune, A.R. (2009). Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. *Evolution*, 63, 1312–1326.

Wagner, C.E., Harmon, L.J. & Seehausen, O. (2012). Ecological opportunity and sexual selection together predict adaptive radiation. *Nature*, 487, 366–369.

Wright, D.H. (1983). Species-energy theory – an extension of species-area theory. *Oikos*, 41, 496–506.

SUPPORTING INFORMATION