

Ecological opportunity and sexual selection together predict adaptive radiation

Catherine E. Wagner^{1,2,3,4}, Luke J. Harmon⁵ & Ole Seehausen^{1,2}

A fundamental challenge to our understanding of biodiversity is to explain why some groups of species undergo adaptive radiations, diversifying extensively into many and varied species, whereas others do not^{1,2}. Both extrinsic environmental factors (for example, resource availability, climate) and intrinsic lineage-specific traits (for example, behavioural or morphological traits, genetic architecture) influence diversification, but few studies have addressed how such factors interact. Radiations of cichlid fishes in the African Great Lakes provide some of the most dramatic cases of species diversification. However, most cichlid lineages in African lakes have not undergone adaptive radiations. Here we compile data on cichlid colonization and diversification in 46 African lakes, along with lake environmental features and information about the traits of colonizing cichlid lineages, to investigate why adaptive radiation does and does not occur. We find that extrinsic environmental factors related to ecological opportunity and intrinsic lineage-specific traits related to sexual selection both strongly influence whether cichlids radiate. Cichlids are more likely to radiate in deep lakes, in regions with more incident solar radiation and in lakes where there has been more time for diversification. Weak or negative associations between diversification and lake surface area indicate that cichlid speciation is not constrained by area, in contrast to diversification in many terrestrial taxa³. Among the suite of intrinsic traits that we investigate, sexual dichromatism, a surrogate for the intensity of sexual selection, is consistently positively associated with diversification. Thus, for cichlids, it is the coincidence between ecological opportunity and sexual selection that best predicts whether adaptive radiation will occur. These findings suggest that adaptive radiation is predictable, but only when species traits and environmental factors are jointly considered.

Adaptive radiations are iconic systems for the study of evolutionary processes because they generate a wealth of ecological and species diversity, often on very rapid timescales^{2,4}. Some of the most spectacular examples of young adaptive radiations occur on oceanic islands or in lakes, but such geographically circumscribed habitats are no guarantee for a radiation to evolve. Why is it that some lineages diversify markedly, whereas closely related lineages in the same habitat do not?

One point of view is that adaptive radiation is a consequence of newly arising ecological opportunity^{1,4}. Extrinsic ecological factors that have been linked to adaptive radiation include a paucity of competing lineages^{1,2}, predation regime⁵, biotic insularity⁶, habitat complexity⁷ and habitat area³. In addition, latitude⁸ and energy (measured as solar radiation or primary productivity)⁹ have been classically linked to variation in broad-scale patterns of diversity (for example, the latitudinal diversity gradient), but these factors have not been previously investigated in the context of adaptive radiations.

Another point of view is that differences in diversification result primarily from variation in lineage-specific traits that affect speciation rates, such as prevalence of sexual selection¹⁰, ecological specialization¹¹,

ecological versatility¹² and spatial vagility³. There is mounting evidence for traits underlying variation in diversification rates, but the overall proportion of variation explained is generally low¹³. A main challenge in explaining the causes of diversification lies in identifying the relative roles of intrinsic and extrinsic factors, and how these factors interact to determine the rate and volume of species radiations.

Rarely have the influences of multiple extrinsic and intrinsic factors been considered simultaneously in the study of adaptive radiation. Since the discovery of the species-rich African lake cichlid faunas, hypotheses for the spectacular diversity of these fishes have proliferated, invoking environmental factors^{14,15}, intrinsic traits^{12,16} and their interactions^{14,17} as influences on radiation. However, these hypotheses remain untested at macroevolutionary scales. Most research has focused on the cichlid radiations in Lakes Victoria, Malawi and Tanganyika, but cichlids have independently diversified within African lakes on more than 30 occasions, and have colonized lakes without diversifying on more than 120 occasions. These replicated cases of both occurrence and absence of diversification provide an opportunity to test which factors predict whether a cichlid lineage will diversify.

We built a molecular phylogeny for African cichlids (Supplementary Information 1), and placed all lacustrine African cichlids included in our data set on this tree (Fig. 1). We then collated information on lake characteristics for 46 lakes harbouring cichlids across the African continent, including lake depth, surface area, net solar radiation (hereafter 'energy'), latitude, elevation, the presence of predatory fishes and time for diversification (Fig. 1 and Supplementary Information 2). We collected data on intrinsic traits of cichlid lineages, including the presence of a polygamous mating system, mouthbrooding, generalized egg dummies and/or morphologically derived 'haplochromine' egg dummies¹⁸ (used in courtship and in fertilization of eggs in the mouth of the female) and sexual dichromatism. We then tested for associations between these predictor variables and cichlid 'diversification state': that is, whether a lineage has diversified upon entering a lake or has failed to do so. We conducted analyses using two thresholds for the endemic diversity required to qualify as a radiation: at the lower threshold, we counted any lineage that had undergone at least one intralacustrine speciation event as radiating; at the higher threshold, we counted lineages as radiating only if they produced five or more endemic species within a single lake. Furthermore, we tested jointly for factors predicting radiation and the species richness of radiations using phylogenetic hurdle Poisson regression (see Supplementary Information 5.3).

We examined relationships between cichlid radiation and single predictor variables (Supplementary Tables 3 and 4) and then assessed the combined influence of predictor variables on diversification state in multiple regression models using Akaike information criterion-based model averaging corrected for small sample sizes (AICc)¹⁹ followed by phylogenetic multiple logistic regression and phylogenetic hurdle Poisson regression of a reduced predictor variable set. The best-supported predictor variables in our multiple regression models

¹Department of Fish Ecology & Evolution, EAWAG Centre for Ecology, Evolution and Biogeochemistry, 6047 Kastanienbaum, Switzerland. ²Department of Aquatic Ecology, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland. ³Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853, USA. ⁴Fuller Evolutionary Biology Program, Cornell University Lab of Ornithology, Ithaca, New York 14850, USA. ⁵Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844, USA.

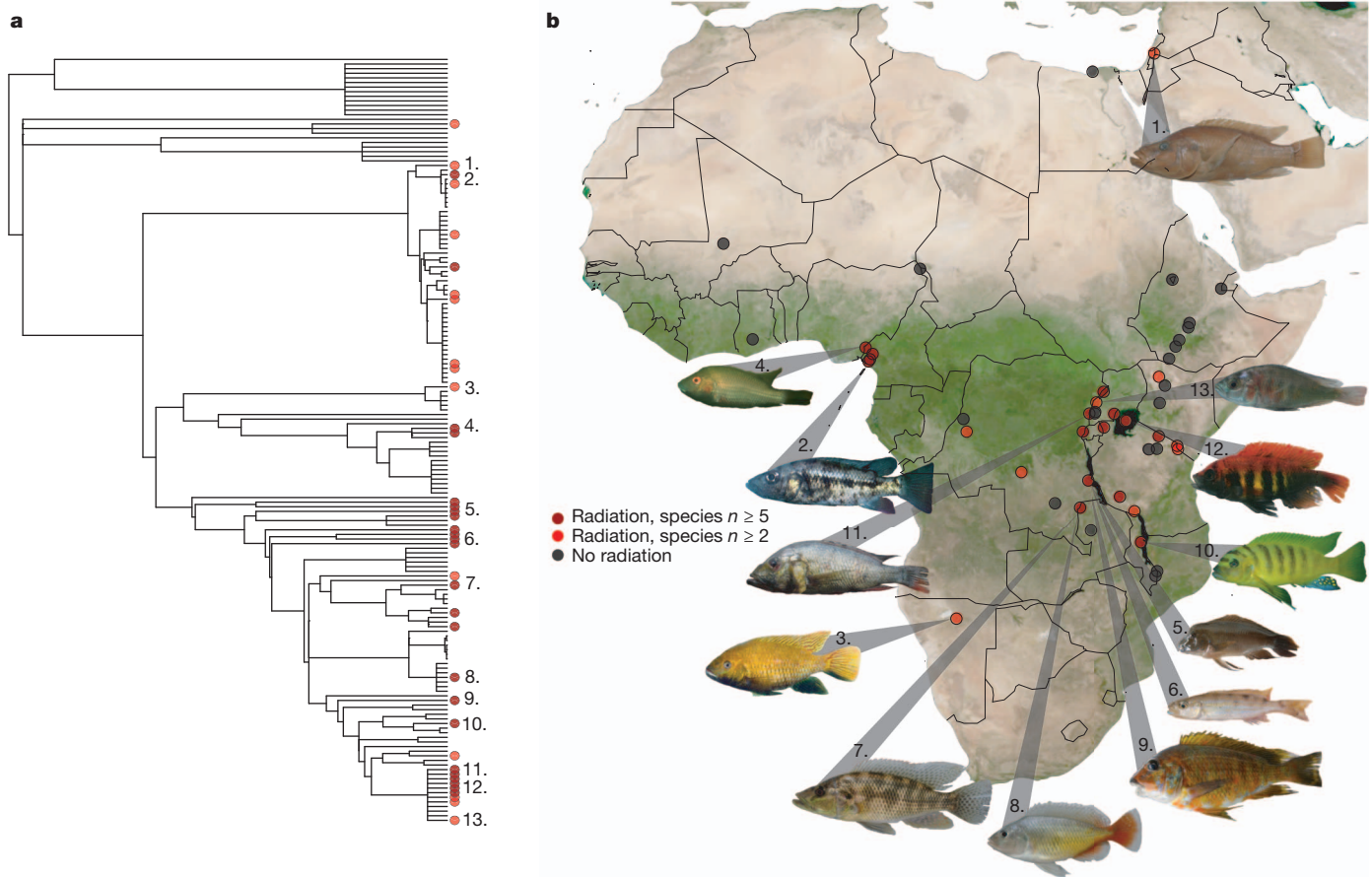


Figure 1 | Cichlid diversification is phylogenetically and geographically widespread. **a**, The distribution of intralacustrine adaptive radiation across the African cichlid phylogeny. Each tip represents one lineage in a lake; light red dots indicate at least one intralacustrine speciation event, dark red dots indicate radiation of five or more species. **b**, The geographic distribution of cichlid

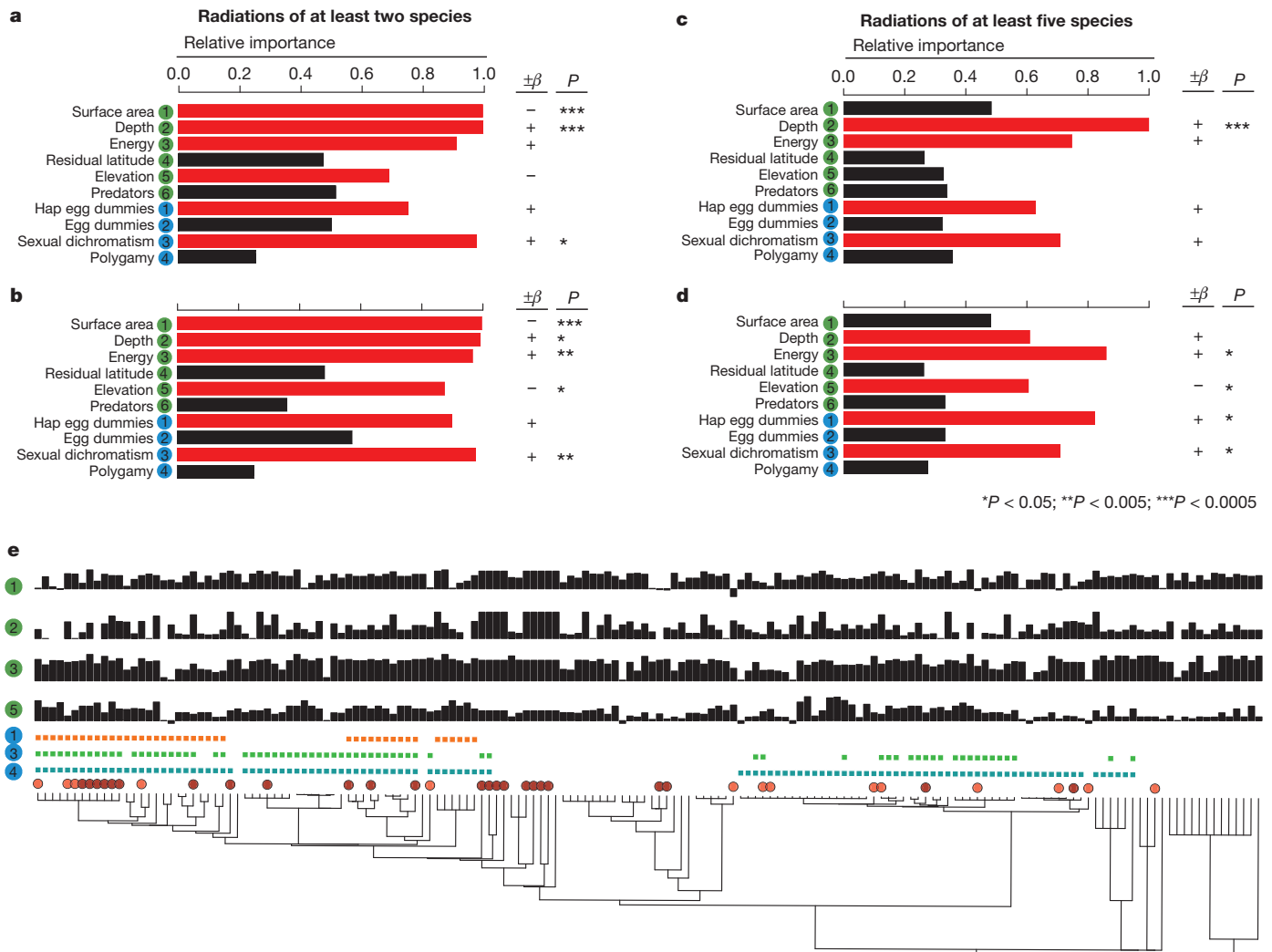
include both environmental variables and lineage-specific traits (Fig. 2). Lake depth, energy and sexual dichromatism are the most consistently well-supported predictor variables: all were positively related to diversification with high relative importance scores for both of our radiation thresholds (that is, ≥ 2 or ≥ 5) in logistic regressions, and were top predictors in the binary portion of hurdle Poisson regressions (Supplementary Fig. 6). There was equivocal support for diversification to be more likely in lakes with small surface area, because this variable only had a high relative importance score at the lower threshold of radiation in logistic regression analyses. As a conservative test, we conducted these analyses excluding Lake Tanganyika, an outlier in both depth and age. In these tests, the same environmental variables were the strongest extrinsic factors associated with diversification, and a negative effect of elevation emerged as a predictor of diversification for both thresholds. Sexual dichromatism remained the most consistent intrinsic trait predictor of diversification (Fig. 2 and Supplementary Fig. 6).

All regression models show a strong association between lake depth and cichlid diversification, which is consistent with depth being an important axis of niche differentiation in intralacustrine speciation in fishes. Lake depth and age are typically highly correlated (Supplementary Information 3), and deeper lakes might additionally have greater environmental stability and/or greater persistence times, both of which would allow lineages more opportunity for diversification. Results showing a relationship between cichlid radiation and time for diversification support this idea (Supplementary Information 4). However, analyses on a subset of the data wherein time and depth were uncorrelated show a better fit for lake depth than time for diversification in predicting radiation (Supplementary Information 3), suggesting a

role for depth apart from time in cichlid radiation. Depth partitioning of resources and reproduction is important in many cases of speciation in fishes²⁰, and case studies indicate that depth-specific divergence in mating traits and preferences and depth-specific ecological adaptation can be key factors in cichlid speciation²¹. Furthermore, increased depth increases habitat area for fishes, and the resultant larger population sizes may influence speciation and extinction rates.

Net solar radiation emerges as a second strong predictor of cichlid diversification in multiple regression models. Links between energy and evolutionary diversification have been frequently proposed in the context of latitudinal gradients in species richness^{8,9}, although only rarely has this relationship been tested. Increased energy input might increase carrying capacities, leading to larger total population sizes and increased rates of speciation and/or lower rates of extinction. Alternatively, high inputs of energy may lead to shortened generation times and/or increased mutation rates, resulting in increased rates of population differentiation and speciation^{8,9}.

In contrast to diversification in terrestrial systems³, we find that increased lake surface area does not increase the likelihood that colonizing lineages will undergo intralacustrine speciation. Ascertainment bias could influence this result: data on the presence of species in very small lakes are rarer than in large lakes, and many of the small lakes included in our data set are known because they harbour endemic cichlids (Supplementary Information 5.2). However, in other systems such as *Anolis* lizards, very small islands never host adaptive radiations²². Regardless of potential size-related sampling bias, this finding demonstrates a marked contrast between cichlids and terrestrial taxa in that speciation is apparently not constrained by surface area in cichlids.



* $P < 0.05$; ** $P < 0.005$; *** $P < 0.0005$

Figure 2 | Multiple logistic regression shows that environment and lineage-specific traits together best explain cichlid diversification in African lakes. a–d, Bar length is proportional to relative importance values, with red bars indicating relative importance values greater than 0.6 in multiple regression models. a, c, Results for the full data set. b, d, Results excluding Lake Tanganyika. Plus and minus symbols indicate the sign of multiple logistic regression coefficient estimates (β). Asterisks indicate the significance of a term in phylogenetic multiple logistic regression analyses. Green and blue circle labels represent environmental variables and species traits, respectively. Among environmental variables, there are positive associations between diversification and lake depth and environmental energy. Lake surface area is a negative predictor of diversification when radiations are considered to consist of two or more endemic species (a, b), but for larger radiations (five or more species) the

significance of this negative size effect disappears (c, d); extremely species-rich radiations only occur in large lakes. Among lineage-specific traits, the presence of sexual dichromatism is a consistent predictor of diversification. Haplochromine egg dummies are consistently associated with diversification, but in most cases the significance of this effect disappears when phylogeny is accounted for. e, High relative importance variables, plus the low relative importance variable polygamous mating system, plotted on the African cichlid phylogeny (dots indicate radiation, as in Fig. 1). All cichlid lineages with sexual dichromatism or haplochromine egg dummies have polygamous mating systems. That the evolution of sexual dichromatism and egg dummies only occurs in lineages with polygamous mating systems suggests that polygamy is a prerequisite to strong sexual selection in cichlids.

The positive association between sexual dichromatism and diversification in all our models (Fig. 2 and Supplementary Information 4 and 5) suggests that the intensity of sexual selection may be a key influence on the probability that lineages radiate. Sexual dichromatism is a common proxy for strong sexual selection in studies of diversification¹⁰. Variation among and within populations in traits under sexual selection, and in associated preferences, can readily lead to pre-mating isolation among populations, and thereby facilitate speciation^{23,24}. Sexual selection is known to be important in cichlid speciation from case studies, but here we show an association between sexual selection and diversification in cichlids at macroevolutionary scales. Examination of the co-occurrence between dichromatism and mating system shows that sexual dichromatism only evolves in lineages that have polygamous mating systems (Fig. 2), a pattern predicted if mating system determines opportunity for sexual selection²⁵. Yet, mating

system does not emerge as an important predictor of radiation in our models. This result suggests that dichromatism is a more direct indicator of the actual strength of sexual selection than is mating system, a pattern that has been suggested in meta-analysis of findings from other taxa¹⁰ but which has never been tested in cichlids.

Although African cichlid fishes are an iconic example of adaptive radiation, our analysis shows great heterogeneity in the occurrence of adaptive radiation across this clade: most lineages present in lakes do not diversify. However, some lineage traits significantly predict whether radiations happen when a suitable environment is colonized. This result makes clear that the propensity for high diversification is not an intrinsic property of all cichlids, but one that has evolved in some branches of the cichlid tree. Although other unmeasured variables undoubtedly explain additional variation in the occurrence of adaptive radiation across the African cichlid phylogeny, we show here,

that cichlid adaptive radiations are not a simple function of any one predictor variable, but instead are best predicted by variables representing both extrinsic environmental effects and intrinsic, lineage-specific traits. For cichlids, it is the combined effects of the intensity of sexual selection and environmental opportunity, in the form of lake depth, energy availability and lake age, that best predict whether adaptive radiation will occur. More generally, the finding that propensity for adaptive radiation is underlain by several factors helps to explain why only some taxa radiate, even in environmental settings—such as islands and lakes—that are home to some of evolution's classic cases of adaptive radiation. Thus it is possible that adaptive radiation is predictable, but only when traits and environmental factors are jointly considered.

METHODS SUMMARY

We built maximum likelihood phylogenies in RAxML²⁶ using nine genes and sequences from 656 African cichlid species, and used PATHd8 (ref. 27), three geological dates and one fossil date to time-calibrate these trees (see Supplementary Information 1).

We compiled information about the presence and species richness of cichlid lineages in lakes across Africa. Because most colonizing lineages do not diversify, this data set is zero-inflated, and thus for analysis in a logistic regression framework we coded lineages in each lake as either 'non-diversifying' or 'diversifying' using one of two thresholds (Supplementary Information 2.3–4). We then compiled information about character states for traits potentially linked to cichlid diversification, and environmental variables for all lakes (Supplementary Information 2). We calculated maximum time for diversification for lineages using either the midpoint of geological age estimates for the lake or the mean stem age of the radiating group estimated from our calibrated molecular phylogenies.

To use the tree to account for phylogeny, we trimmed it to include only lineages that occur in lakes, and a single taxon to represent each diversifying lineage. For lineages present in several lakes, we added a tip to the tree for each instance the lineage had independently colonized a lake (see Supplementary Information 1.3), thereby accounting for each 'opportunity' for diversification.

We used phylogenetic logistic regression²⁸ to assess the relationship between each predictor variable and cichlid diversification state. Then, after assessing collinearity between predictor variables (Supplementary Information 3), we used multiple logistic regression to assess the combined influence of predictor variables on diversification state using a two-stage approach (Supplementary Information 5). First, we used AICc-based model averaging¹⁹ to assess the relative importance of predictor variables. Second, we included predictor variables with relative importance values above 0.6 in phylogenetic logistic and phylogenetic hurdle Poisson regression models to attain phylogenetically corrected regression parameter estimates.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 January; accepted 16 April 2012.

Published online 10 June 2012.

1. Simpson, G. G. *The Major Features of Evolution* (Columbia Univ. Press, 1953).
2. Losos, J. B. Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* **175**, 623–639 (2010).
3. Kisel, Y. & Barraclough, T. G. Speciation has a spatial scale that depends on levels of gene flow. *Am. Nat.* **175**, 316–334 (2010).
4. Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
5. Vamosi, S. M. The presence of other fish species affects speciation in threespine sticklebacks. *Evol. Ecol. Res.* **5**, 717–730 (2003).
6. MacArthur, R. H. & Wilson, E. O. Equilibrium-theory of insular zoogeography. *Evolution* **17**, 373– (1963).
7. Price, S. A., Holzman, R., Near, T. J. & Wainwright, P. C. Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. *Ecol. Lett.* **14**, 462–469 (2011).
8. Mittelbach, G. G. *et al.* Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).

9. Evans, K. L., Warren, P. H. & Gaston, K. J. Species-energy relationships at the macroecological scale: a review of the mechanisms. *Biol. Rev. Camb. Philos. Soc.* **80**, 1–25 (2005).
10. Kraaijeveld, K., Kraaijeveld-Smit, F. J. L. & Maan, M. E. Sexual selection and speciation: the comparative evidence revisited. *Biol. Rev. Camb. Philos. Soc.* **86**, 366–377 (2010).
11. Farrell, B. D. "Inordinate fondness" explained: why are there so many beetles? *Science* **281**, 555–559 (1998).
12. Liem, K. F. Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. *Syst. Zool.* **22**, 425–441 (1973).
13. Ricklefs, R. E. History and diversity: explorations at the intersection of ecology and evolution. *Am. Nat.* **170**, S56–S70 (2007).
14. Fryer, G. Some aspects of evolution in Lake Nyasa. *Evolution* **13**, 440–451 (1959).
15. Sturmbauer, C., Baric, S., Salzburger, W., Ruber, L. & Verheyen, E. Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Mol. Biol. Evol.* **18**, 144–154 (2001).
16. Seehausen, O. & van Alphen, J. M. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.* **2**, 262–271 (1999).
17. Seehausen, O. Evolution and ecological theory—chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. *Heredity* **99**, 361–363 (2007).
18. Greenwood, P. H. Towards a phyletic classification of the 'genus' *Haplochromis* (Pisces, Cichlidae) and related taxa. Part 1. *Bull. Br. Mus. Nat. Hist. Zool.* **35**, 265–322 (1979).
19. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference* (Springer, 2002).
20. Ingram, T. Speciation along a depth gradient in a marine adaptive radiation. *Proc. R. Soc. B* **278**, 613–618 (2011).
21. Seehausen, O. *et al.* Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008).
22. Losos, J. B. & Schluter, D. Analysis of an evolutionary species-area relationship. *Nature* **408**, 847–850 (2000).
23. Lande, R. Rapid origin of sexual isolation and character divergence in a cline. *Evolution* **36**, 1–12 (1982).
24. Maan, M. E. & Seehausen, O. Ecology, sexual selection and speciation. *Ecol. Lett.* **14**, 591–602 (2011).
25. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) 136–179 (Aldine, 1972).
26. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
27. Britton, T., Anderson, C. L., Jacquet, D., Lundqvist, S. & Bremer, K. Estimating divergence times in large phylogenetic trees. *Syst. Biol.* **56**, 741–752 (2007).
28. Ives, A. R. & Garland, T. Phylogenetic logistic regression for binary dependent variables. *Syst. Biol.* **59**, 9–26 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Ives, T. J. Davies and A. Mooers for analytical advice, P. McIntyre and C. Reidy Liermann for access to global energy data, S. Mwaiko, R. B. Stelkens, C. Katongo and U. Schliwen for unpublished DNA sequences, U. Schliwen, J. Jensen and O. Rittner for photographs, and A. McCune, D. Rabosky, R. Harrison, I. Lovette, E. Michel, G. Mittelbach, C. Melian, J. Brodersen, M. Maan, T. Ingram, B. Matthews, B. Dalziel, M. Pennell, J. Eastman, the Harmon laboratory group, the McCune laboratory group and the Seehausen laboratory group for discussions and comments on the manuscript. Bioinformatics facilities were supported by grants from the National Center for Research Resources (5P20RR016448-10) and the National Institute of General Medical Sciences (8P20 GM103397-10) from the National Institutes of Health. This work was supported by the Swiss National Science Foundation project 31003A-118293 (to O.S.) and US National Science Foundation grant DEB 0919499 (to L.J.H.).

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. wrote the paper.

Author Information Sequence data are deposited in the GenBank database under accession numbers listed in Supplementary Table 1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.E.W. (catherine.wagner@eawag.ch) or O.S. (ole.seehausen@eawag.ch).

METHODS

Phylogenetic framework. We compiled sequence data for nine genes and 656 African cichlid species, with the goal of phylogenetically placing all African cichlid lineages present in lakes (Supplementary Information 1). The aligned, concatenated data set included a total of 6,947 base pairs.

We used a maximum likelihood approach in RAxML for phylogenetic analyses²⁶ (Supplementary Information 1.2). To account for phylogenetic uncertainty, we used 100 replicates of the rapid bootstrap algorithm in RAxML and estimated branch lengths for each of these bootstrap replicate topologies. To ultrametricize and time-calibrate this set of trees, we used PATHd8 (ref. 27). For time-calibration we used three geological dates and one fossil date: two dates associated with the breakup of Gondwana (the Africa–Madagascar split and the Madagascar–India split), the age of the earliest known fossil *Oreochromis*, and the age of Lake Nabugabo (Supplementary Information 1.2). We then drew 95% confidence intervals on node ages from the distribution of branching times estimated from this set of calibrated ultrametric trees.

Cichlid radiation data, ecological variables and species traits. We compiled information about presence of cichlid lineages in lakes across Africa, and the endemic diversity of the lineages present in each of these lakes (Supplementary Information 2). Because most colonizing lineages do not diversify, this data set is zero-inflated; the processes influencing radiation therefore should be analytically considered separately from the processes influencing the species richness of radiating lineages (see Supplementary Information 2.4). We thus coded each lineage in each lake as either ‘diversifying’ or ‘non-diversifying’ at two diversity thresholds. First, radiations with one or more intralacustrine speciation events (any lineage that had at least one endemic species in a lake co-occurring with its sister taxon, be it either a widespread species (in three cases) or a lake endemic itself); second, as a more conservative threshold for radiation, diversification events producing at least five endemic species. Analyses conducted at other thresholds produced qualitatively identical results to those at these threshold values. Single endemic species not co-occurring in the same lake with a sister taxon were not considered to be radiating lineages.

We compiled information about lineage-level character states for traits potentially linked to cichlid speciation, and environmental variables, and then used phylogenetic logistic regression²⁸ to analyse the association between these factors and diversification state. Tested lineage traits included the presence of a polygamous mating system, of mouthbrooding, of generalized egg dummies and specialized haplochromine-type egg dummies on the anal fin of male fish, and the presence of strong sexual dichromatism (Supplementary Information 2). Many of these traits have been proposed to be linked to sexual selection, and mouthbrooding has additionally been proposed as an ecological key innovation because it liberates cichlids from substrate-related habitat requirements for attachment and guarding of eggs²⁹. These traits are rarely polymorphic within cichlid lineages. These few instances were coded by majority state, or as missing data (in one case, presence/absence of egg dummies in *Thoracochromis* of Lake Fwa, where majority state was ambiguous).

We compiled information on physical and environmental variables for all lakes in the data set. These included surface area, maximum depth, latitude, net solar radiation (the difference between the influx of solar energy and that reflected back into the atmosphere at a given geographic location, referred to simply as ‘energy’) and elevation (Supplementary Information 2). We chose these variables as the main factors correlating with lake type, habitat availability and climate that were available for many lakes. As a further environmental variable, we included the presence of large predatory fish (genera *Lates*, *Hydrocynus*, *Hepsetus*) because of their hypothesized influence on cichlid diversification^{14,30}.

We calculated maximum time for diversification for lineages using either the midpoint of geological age estimates for the lake (either basin age or most recent desiccation age) or the mean stem age of the radiating group estimated from our calibrated molecular phylogenies. We also conducted analyses using only geological lake ages, and these produced very similar results (Supplementary Tables 3 and 4). **Regression models.** To account for phylogeny in regression models, we trimmed the best maximum likelihood topology to include only lineages that occur in lakes

and a single taxon for each within-lake radiation. For lineages present in several lakes, we added a tip to the tree for each instance where the lineage is found in a unique lake, such that each lineage found in several lakes is represented as a polytomy with a tip corresponding to each lake where it occurs. We set branch lengths on these added tips to have a total length of that expected under a pure birth model (Supplementary Information 1.3). Using this approach, our trimmed and manipulated phylogenies had a branch for each ‘opportunity’ to diversify; that is, each instance a lineage entered a new lake.

We used phylogenetic logistic regression²⁸ to assess the relationship between single predictor variables and diversification state. To assess the combined influence of our predictor variables on cichlid diversification state, we used multiple logistic regression models. Before including the predictor variables in multiple regression models, we checked for collinearity between both continuous and binary predictor variables. We calculated Pearson correlation coefficients (r^2) for all pairs of continuous predictor variables. For binary predictor variables, we used the r^2 equivalent³¹, r^2_{L1} , as an assessment of collinearity (Supplementary Information 3). We removed one variable from each pair of predictor variables with r^2 (or r^2_{L1}) greater than 0.7 after preliminary models including variables with higher correlations caused analytical problems (inflations of standard error, a diagnostic of collinearity problems in logistic regression³²).

Because we discovered a strong correlation between lake depth and time for diversification during collinearity tests ($r^2 = 0.76$), we conducted further tests to determine the relative effects of time and depth. We excluded taxa from lakes greater than 150 m in depth ($n = 3$ of 46), leaving the remaining data subset uncorrelated in time and depth ($r^2 = 0.25$). We compared AIC values among models incorporating time, time + depth, and depth as predictors of cichlid diversification in this data set (Supplementary Information 3).

We examined the combined influence of predictor variables on diversification state in multiple regression models. Because likelihood-based phylogenetic logistic regression methods are not available, we used the following two-step approach. First, we used AICc-based model averaging¹⁹ to evaluate the parameter estimates and the relative importance of predictor variables in a likelihood-based framework. We calculated model-averaged parameter estimates and standard errors for each predictor variable using relative AICc weights of models in which the variables appeared. We calculated the relative importance of each predictor variable as the sum of the AICc weights of all models that included this variable. Second, we included predictor variables with relative importance values above 0.6 in phylogenetic multiple logistic regression²⁸ models to attain phylogenetically corrected regression parameter estimates.

As an additional test of our results, we performed phylogenetic hurdle Poisson regression using the R package MCMCglmm³³, using the number of speciation events within each colonizing lineage as the response variable. This approach models two latent variables associated with the data: one associated with a binary process, the other modelling the non-zero response values in the data set as a Poisson process (Supplementary Information 5.3). We repeated these analyses over a set of 100 bootstrap replicate trees to account for phylogenetic uncertainty. Results were qualitatively identical to those from analyses modelling the binary process alone (Supplementary Information 5.3). Also using this modelling framework, we did post hoc tests for interaction effects between environmental and lineage-specific variables; these produced some evidence for interaction effects between lake depth and sexual dichromatism in predicting cichlid adaptive radiation (Supplementary Information 5.4).

29. Kuwamura, T. in *Fish communities in Lake Tanganyika* (eds Kawanabe, H., Hori, M. & Nagoshi, M.) 59–86 (Kyoto Univ. Press, 1997).
30. Worthington, E. B. & Ricardo, C. K. The fish of Lake Tanganyika (other than Cichlidae). *Proc. Zool. Soc. Lond.* **1936**, 1061–1112 (1936).
31. Menard, S. Coefficients of determination for multiple logistic regression analysis. *Am. Stat.* **54**, 17–24 (2000).
32. Quinn, G. P. & Keough, M. J. *Experimental Design and Data Analysis for Biologists* (Cambridge Univ. Press, 2002).
33. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**, 1–22 (2010).