Supporting Information

Weigelt et al. 10.1073/pnas.1306309110

SI Materials and Methods

We publish a data matrix of bioclimatic and physical characteristics, ordination and clustering results, and species richness predictions for the 17,883 islands >1 km^2 investigated in this article as comma-separated text file (Dataset S1, dx.doi.org/10.5061/dryad.fv94v). The table is sorted by IDs (*ID*) unique to each island. Each island refers to a polygon in the GADM database of Global Administrative Areas, version 1 (www.gadm.org/version1). Twelve islands identified in a previous study (1) to be missing from the GADM data or to be connected to continents erroneously (*IDs*: 85133, 85137, 85138, 85139, 85145, 85149, 85150, 100046, 100049, 100050) were drawn manually or clipped from continents. Longitude (*Long*) and Latitude (*Lat*) were calculated as polygon mass centroids. International Organization for Standardization country codes (*CountryISO*) and country names (*Country*) were adopted from GADM. In the case of multiple countries per island, country codes were amalgamated (up to five characters) and country names listed separated by semicolons. Where applicable, an archipelago name (*Archip*) was assigned.

For 11,546 islands, names (*Island*) were assigned using the NGA GEOnet Names Server (downloaded on March 29, 2012, from earth-info.nga.mil/gns/html/index.html; indicated as "gns" in column *Gazetteer*) for all regions but the United States, and the US Geological Survey Geographic Names Information System (downloaded on March 29, 2012, from geonames.usgs.gov/index.html; indicated as "gnis") for the United States. Only names classified as islands in these two resources were considered. Original ID (*Name_ID*) and geographic coordinates (*Name_long* and *Name_lat*) were adopted from the gazetteers. In total, 7,475 islands were assigned single names that fell inside their polygons (indicated as "inside" in column *Name_meth*). In 1,751 cases, more than just one name was located inside an island polygon (*No_names*), e.g., due to erroneously located names of closely adjacent islands or inland freshwater islands. For all islands with 10 or less names, the name located nearest (based on *Name_dist*) (in kilometers) to the island's mass centroid was chosen automatically (indicated as "insideclosest"). In this case, the alternative names are also given (*Name_alt*). For 73 islands with more than 10 names (large and well-known islands), the right name was chosen manually (indicated as "manually"). For 2,320 islands without a name within the polygon, names could be assigned because the island polygons were the closest features to island names not assigned previously and their mass centroids were not further than 10 km from the name's coordinates (indicated as "closest"). In this case, *No_names* indicates the number of names each polygon was the closest (based on *Name_dist*; in kilometers) feature for. If *No_names* was >1, the closest name was chosen. A quality check of 100 randomly drawn islands for each method found that about 93% of the names of method "inside," 84% of method "insideclosest," and 78% of method "closest" can be assumed to be correct. Hence, the island names may help to find data for certain islands in our dataset, but due to their insecure assignment they must not be used for automated matchups. The island coordinates should be used instead as a spatially explicit reference.

Island area (*Area*) (in square kilometers) was calculated for each GADM polygon in cylindrical equal area projection. As measures of island isolation, we provide the distance from an island's mass centroid to the nearest mainland coast (*Dist*) (in kilometers) measured in azimuthal equidistant projection using the "Near Table" tool in ArcGIS Desktop 9.31 (Esri) and the $\log_{10}$-transformed sum of the proportions of landmass within buffer distances of 100, 1,000, and 10,000 km around the island perimeter (*SLMP*) (1). Estimates of whether an island was connected to the mainland during the last glacial maximum (LGM) or not (*GMMC*) were based on global bathymetry data (2) assuming a sea level decrease of −122 m at 18,000 years before present (3). However, this metric does not account for regional differences in sea level fluctuations and plate tectonics. Maximum elevation above sea level (a.s.l.) of each island (*Elev*) (in meters) was extracted from the digital elevation model at 30-s resolution provided in WorldClim (4), which is based on SRTM (5) and GTOPO30 (6) using the "Zonal Statistics" tool in ArcGIS. For 1,891 small islands that did not fully enclose a 30-s WorldClim raster cell, we applied a 1-km buffer as indicated in column *Buffer*.

We extracted bioclimatic variables from WorldClim (BIO1, BIO7, BIO12, and BIO15) in a similar manner to *Elev*. When interpreting the climate patterns, one has to consider possible shortcomings of the WorldClim data. WorldClim interpolates climatic measurements between climate stations accounting for latitude, longitude, and elevation but disregards other important information like slope aspect or predominant wind directions (4). Especially for precipitation in mountainous tropical regions with few climate stations, the data might be imprecise (7).

Here, we provide maximum values per island polygon of annual mean temperature (*Temp*) (in degrees Celsius) and annual precipitation (*Prec*) (in millimeters), and minimum values of the annual temperature range (*varT*) (in degrees Celsius) and the coefficient of variation in monthly precipitation (*varP*). For a region of 129 islands >1 km^2 including parts of French Polynesia and the Pitcairn Islands that lack WorldClim temperature data, we modeled *Temp* and *varT* based on the relationships of sea surface temperature and its range with *Temp* and *varT* on neighboring islands. We extracted sea surface temperature data (8) for all islands of French Polynesia, the Cook Islands, the Pitcairn Islands, Kiribati, Wallis and Futuna, Fiji, American Samoa, Niue, Tokelau, Tonga, and Samoa. We then fitted linear models of the maximum values of annual mean temperature and minimum values of the temperature range from WorldClim for the islands covered by WorldClim (n = 255) and mean annual sea surface temperature (*ssTemp*) and range (*ssvarT*) and used the model to predict maximum mean annual temperature and minimum temperature range for the islands not covered ($Temp = -9.36 + 1.29 \times ssTemp$, $R^2 = 0.87$, $P < 0.001$; $varT = 4.96 + 1.39 \times ssvarT$, $R^2 = 0.93$, $P < 0.001$). Islands with modeled temperature data are marked in column *modeled_T*.

We calculated climate change velocity (*CCVT*) (in meters per year) since the LGM 21,000 y ago following refs. 9 and 10. Climate change velocity is the ratio between the temporal change in temperature (temporal gradient) and the contemporary spatial change in temperature (spatial gradient), and is expressed in distance units per time. We calculated the temporal gradient as the difference between the current annual mean temperature and the annual mean temperature at the LGM divided by 21,000 y. Current climate data were based on the 30-s WorldClim data and our model predictions for parts of French Polynesia and the Pitcairn Islands. Based on the predicted maximum annual mean temperature at sea level as intercepts, we modeled annual mean temperatures (*meanT*) for each WorldClim raster cell of the 129 missing islands. We used the mean slope of regressions between WorldClim annual mean temperature and elevation a.s.l. for the neighboring highly elevated volcanic islands Tahiti, Raiatea, Savaii, Upolu, and Kauai ($meanT = Temp + (-0.0056) \times Ele$-

vation; R^2 values of all *meanT* ~ *Elevation* models > 0.99, all *P* values < 0.001). Data from two past climate models (CCSM3 and MIROC3.2) were taken from the Paleoclimate Modeling Intercomparison Project Phase II (11). We used the mean of the two model predictions as LGM mean annual temperature according to ref. 10. For comparison with current climate, we downscaled the estimates to 30-s resolution and then calculated the spatial mean annual temperature gradient based on the contemporary climate data as the slope from each raster cell to its four nearest neighbors. To avoid dividing by zero, all values below 0.01 °C/km and values of cells with less than four direct neighbors were replaced with 0.01 °C/km. We then extracted mean values of climate change velocity for each island. If not stated otherwise above, GIS analyses were performed using R statistical software, version 2.14.2 (R Development Core Team; available at cran.r-project.org) and packages *sp* (12), *maptools* (13), raster (14), and *rgdal* (15).

Column names starting with "*PAM*" refer to results from nonhierarchical partitioning around medoids (PAM), and column names starting with "*UPGMA*" refer to results from the hierarchical unweighted pair-group method with arithmetic mean (UPGMA). Axis scores of principal component analyses (PCAs) are stored in columns starting with "*PCA*." Name suffixes refer to the set of bioclimatic and physical variables considered in each case (*nAE*, all variables except *Area* and *Elev*; *all*, all 10 variables; *cli*, contemporary bioclimatic variables; *geo*, physical variables). For all combinations of clustering method and variable subset, we present eight distinct groups. We used the Caliński and Harabasz index (16) to determine the optimal number of clusters. In the majority of cases for UPGMA, the optimum or local optimum was reached at eight clusters. However, for PAM, index values usually decreased with increasing number of clusters. We therefore adopted the number of eight clusters for all presented regionalizations because eight clusters were well suited for graphical presentation and conceptual discussion. This semiquantitative approach is in line with other studies that highlight the adequacy of choosing an arbitrary number of clusters (17). Ordination, cluster analyses, and evaluation were performed using the R packages *vegan* (18), *flashClust* (19), *cluster* (20), and *fpc* (21).

As demonstration application of the presented data and multivariate framework in macroecology and biogeography, we used it to develop statistical predictions of the species richness of native vascular plants on all 17,883 islands >1 km^2. We built on existing richness data for vascular plants, including all 345 islands from ref. 1 that could be assigned to a single GADM polygon (22–94) and 130 islands for which data were available from published floras, checklists, and online databases (95–139). Following the rational of ref. 140, we used as predictors the 10 bioclimatic and physical variables presented here. As additional predictor, we included the species richness of the closest mainland grid cell derived from the cokriging based estimates provided by ref. 141 (column *SRML*). We allowed for first-order interactions among *Area* and *Temp*, *Dist* and *SRML*, *Temp* and *Prec*, as well as *Area* and *Dist*. For comparison, we fitted generalized linear models (GLMs) of the Gaussian and Poisson families, spatial simultaneous autoregressive lag models (SARs) accounting for spatial autocorrelation (142), and generalized additive models (GAMs) allowing nonlinear and spatial effects (143, 144). We preferred SARs of the lag type over SARs of the error type because the latter does not consider the spatial effect in predictions for new data (142). In GLMs and SARs, all variables were included as linear effects. For both, based on corrected Akaike information criterion (AICc)-based model comparisons and to reduce skewness, we $\log_{10}$-transformed the following predictors: *Area*, *SLMP* + 0.5, *Elev* + 1, *CCVT* + 1, *Prec* +1, and *SRML* (constants were added to avoid taking the logarithm of zero). Optimal lag distances for SARs were defined following ref. 145 evaluating model AICs and the improvement of Moran's I values of spatial autocorrelation in model residuals compared with nonspatial GLMs. In GAMs, each factor was added as penalized regression splines with up to three degrees of freedom (143, 144). Interactions were added as tensor product interactions with up to three degrees of freedom for each basis. In addition to the aforementioned interactions, GAMs included an isotropic smooth of *Lat* and *Long* on a sphere to account for spatial patterns in the response variable. All variables entered the GAMs untransformed except *Area*, which was $\log_{10}$-transformed after visual model inspection. For all model types, we ran a model selection procedure to identify the best among all possible candidate models and conducted multimodel inference by averaging all candidate models up to a sum of AICc weights of 0.95 (146). Although smooth terms in GAMs are already penalized to prevent overfitting, the minimum degrees of freedom is larger than zero (147), necessitating further model selection. Despite the "count" nature of the response variable, Gaussian GLMs with $\log_{10}$-transformed species richness as response variable performed better than Poisson GLMs of untransformed richness in terms of model fit and model diagnostics (pseudo-R^2 of best Poisson candidate model = 0.671 compared with pseudo-R^2 of best Gaussian candidate model = 0.734). Furthermore, the use of SARs did not improve model fit compared with GLMs (pseudo-R^2 of best SAR candidate model = 0.705). We therefore do not present results and predictions from Poisson GLMs and SARs. Model statistics and predictions from the best candidate models were very similar to those based on multimodel inference (e.g., pseudo-R^2 of best Gaussian GLM and pseudo-R^2 of averaged Gaussian GLMs, both = 0.734; pseudo-R^2 of best GAM = 0.937 compared with pseudo-R^2 of averaged GAMs = 0.936). However, we focus on predictions from multimodel inference here because for both GLMs and GAMs their prediction error (averaged mean error based on 10-fold cross-validation) was slightly smaller (GLM: best model prediction error = 0.137, averaged model prediction error = 0.127; GAM: best model prediction error = 0.044, averaged model prediction error = 0.031). Predicted species numbers together with their SEs can be found in the columns *SR_GLM* and *SR_SE_GLM* for GLM predictions and *SR_GAM* and *SR_SE_GAM* for GAM predictions. Both species richness and SEs were backtransformed [as $\log_{10}$(species richness + 1) was the modeled response variable] to represent actual species numbers. In the main results (Figs. 4 and 5), we focus on predictions based on GAMs because they are more flexible, account for spatial patterns, fit the data better (AIC best GAM = −167.8, AIC best GLM = 392.6), and yield more realistic predictions in regions where the other approaches strongly overestimate richness (e.g., on the western coasts of Africa and Canada; Fig. S7). Model averaging and multimodel inference, generalized additive models, spatial simultaneous autoregressive models, and k-fold cross-validation were applied using the R packages *MuMIn* (148), *mgcv* (143, 144), *spdep* (142), and *boot* (149).

1. Weigelt P, Kreft H (2013) Quantifying island isolation—insights from global patterns of insular plant species richness. *Ecography* 36(4):417–429.
2. Amante C, Eakins BW (2009) ETOPO1—1 arc-minute global relief model: Procedures, data sources and analysis. *NOAA Technical Memorandum NESDIS NGDC-24* (National Geophysical Data Center, Boulder, CO).
3. Miller KG, et al. (2005) The Phanerozoic record of global sea-level change. *Science* 310(5752):1293–1298.
4. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25(15):1965–1978.
5. Jarvis A, Reuter HI, Nelson A, Guevara E (2008) *Hole-Filled Seamless SRTM Data V4* [International Centre for Tropical Agriculture (CIAT), Cali, Colombia]. Available at http://srtm.csi.cgiar.org. Accessed March 15, 2013.
6. US Geological Survey (1996) *GTOPO30—Global 30 Arc Second Elevation Data Set* (US Geological Survey Earth Resources Observation and Science Center, Sioux Falls,

SD). Available at http://www1.gsi.go.jp/geowww/globalmap-gsi/gtopo30/gtopo30.html. Accessed March 15, 2013.

7. Soria-Auza RW, et al. (2010) Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: A case study from Bolivia. *Ecol Modell* 221(8):1221–1229.
8. Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. *J Clim* 15(13):1609–1625.
9. Loarie SR, et al. (2009) The velocity of climate change. *Nature* 462(7276):1052–1055.
10. Sandel B, et al. (2011) The influence of Late Quaternary climate-change velocity on species endemism. *Science* 334(6056):660–664.
11. Braconnot P, et al. (2007) Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum—Part 1: Experiments and large-scale features. *Clim Past* 3(2):261–277.
12. Pebesma E, Bivand R, Rowlingson B, Gomez-Rubio V (2012) sp: Classes and Methods for Spatial Data, 1.0-5 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/sp/index.html. Accessed December 21, 2012.
13. Bivand R, et al. (2013) maptools: Tools for Reading and Handling Spatial Objects, 0.8-23 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/maptools/index.html. Accessed February 10, 2013.
14. Hijmans RJ, van Etten J (2012) raster: Geographic Data Analysis and Modeling, 2.0-41 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/raster/index.html. Accessed December 24, 2012.
15. Bivand R, et al. (2013) rgdal: Bindings for the Geospatial Data Abstraction Library, 0.8-5 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/rgdal/index.html. Accessed February 6, 2012.
16. Caliński T, Harabasz J (1974) A dendrite method for cluster analysis. *Commun Stat-Theor M* 3(1):1–27.
17. Metzger MJ, et al. (2013) A high-resolution bioclimate map of the world: A unifying framework for global biodiversity research and monitoring. *Glob Ecol Biogeogr* 22(5): 630–638.
18. Oksanen J, et al. (2013) vegan: Community Ecology Package, 2.0-6 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/vegan/index.html. Accessed February 11, 2012.
19. Murtagh F, R Development Team, Langfelder P (2012) flashClust: Implementation of Optimal Hierarchical Clustering, 1.01-2 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/flashClust/index.html. Accessed August 21, 2012.
20. Maechler M (2012) cluster: Cluster Analysis Extended Rousseeuw et al., 1.14.3 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/cluster/index.html. Accessed October 14, 2012.
21. Hennig C (2013) fpc: Flexible Procedures for Clustering, 2.1-5 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/fpc/index.html. Accessed January 21, 2013.
22. Abbott I (1978) Factors determining the number of land bird species on islands around south-western Australia. *Oecologia* 33(2):221–233.
23. Baldini RM (1998) Flora vascolare dell'isola del Giglio (Arcipelago Toscano): Revisione tassonomica ed aggiornamento. *Webbia* 52(2):307–404.
24. Baldini RM (2000) Flora vascolare dell'isola di Pianosa (Arcipelago Toscano): Revisione tassonomica ed aggiornamento. *Webbia* 55(1):107–189.
25. Baldini RM (2001) Flora vascolare dell'isola di Giannutri (Arcipelago toscano). *Webbia* 56(1):69–125.
26. Barkalov VYU (2000) Phytogeography of the Kurile Islands. *Nat Hist Res (Chiba)* 7: 1–14.
27. Batianoff GN, Dillewaard HA (1996) Floristic analysis of the Great Barrier Reef continental islands. *State of the Great Barrier Reef World Heritage Area*, eds Wachenfeld D, Oliver J, Davis K (Great Barrier Reef Marine Park Authority, Townsville, QLD, Australia), pp 300–322.
28. Bocchieri E (1988) L'isola Asinara (Sardegna nord-occidentale) e la sua flora. *Webbia* 42(2):227–268.
29. Bocchieri E (1992) The flora of the island of Reulino (Sardinia, Italy). *Willdenowia* 22(1):55–63.
30. Borhidi A (1991) *Phytogeography and Vegetation Ecology of Cuba* (Akadémiai Kiadó, Budapest).
31. Borkowsky O (1994) Übersicht der Flora von Korfu. *Braunschweiger Geobotanische Arbeiten* 3:1–202.
32. Brodie J, Sheehy Skeffington M (1990) Inishbofin: A re-survey of the flora. *The Irish Naturalists' Journal* 23(8):293–298.
33. Brullo S, Minissale P, Spampinato G (1995) Considerazioni fitogeografiche sulla flora della Sicilia. *Ecol Mediterr* 21(1-2):99–117.
34. Buckley RC (1983) The flora and vegetation of Barrow Island, Western Australia. *J R Soc West Aust* 66(3):91–94.
35. Case TJ, Cody ML (1983) *Island Biogeography in the Sea of Cortéz* (Univ of California Press, Berkeley, CA).
36. Chown SL, Gremmen NJM, Gaston KJ (1998) Ecological biogeography of southern ocean islands: Species-area relationships, human impacts, and conservation. *Am Nat* 152(4):562–575.
37. Christodoulakis D (1996) The flora of Ikaria (Greece, E. Aegean Islands). *Phyton* 36(1): 63–91.
38. Cronk QCB (1980) Extinction and survival in the endemic vascular flora of Ascension Island. *Biol Conserv* 17(3):207–219.
39. Cronk QCB (1997) Islands: Stability, diversity, conservation. *Biodivers Conserv* 6(3): 477–493.
40. d'Antonio CM, Dudley TL (1995) Biological invasions as agents of change on islands versus mainlands. *Islands—Biological Diversity and Ecosystem Function*, eds Vitousek PM, Loope LL, Adsersen H (Springer, Berlin), Vol 115, pp 103–121.
41. Dahl AL (1991) *Island Directory. UNEP Regional Seas Directories and Bibliographies* (UNEP, Nairobi, Kenya).
42. Davis SD, et al. (1986) *Plants in Danger. What Do We know?* (International Union for Conservation of Nature and Natural Resources, Gland, Switzerland, and Cambridge).
43. Davis SD, Heywood VH, Hamilton AC (1994) *Centres of Plant Diversity. A Guide and Strategy for Their Conservation. Europe, Africa and the Middle East* (IUCN Publications Unit, Cambridge), Vol 1.
44. Davis SD, Heywood VH, Hamilton AC (1995) *Centres of Plant Diversity. A Guide and Strategy for Their Conservation. Asia, Australia and the Pacific* (IUCN Publications Unit, Cambridge), Vol 2.
45. Davis SD, Heywood VH, Herrera-MacBryde O, Villa-Lobos JL, Hamilton AC (1997) *Centres of Plant Diversity. A Guide and Strategy for Their Conservation. The Americas* (IUCN Publications Unit, Cambridge), Vol 3.
46. de Leonardis W, Zizza A (1994) Flora di interesse apistico dell'isola di Salina (arcipelago Eoliano). Prospettive e Potenzialità. *Apicoltura* 9:73–101.
47. de Lange PJ, Rolfe JR (2010) *New Zealand Indigenous Vascular Plant Checklist* (New Zealand Plant Conservation Network, Wellington, New Zealand).
48. Ferro G, Furnari F (1968) Flora e vegetazione di Stromboli (Isole Eolie) *Arch Bot Biogeogr Ital* 44(1–2):21–45; (3):59–85.
49. Ferro G, Furnari F (1970) *Flora e Vegetazione di Vulcano (Isole Eolie)* (Pubblicazioni dell'Istituto di Botanica dell'Universita di Catania, Catania, Italy).
50. Florence J, Lorence DH (1997) Introduction to the flora and vegetation of the Marquesas Islands. *Allertonia* 7(4):226–237.
51. Florence J, Waldren S, Chepstow Lusty A (1995) The flora of the Pitcairn Islands: A review. *Biol J Linn Soc Lond* 56(1-2):79–119.
52. Frodin DG (2001) *Guide to Standard Floras of the World* (Cambridge Univ Press, Cambridge), 2nd Ed.
53. Gabrielsen GW, Brekke B, Alsos IG, Hansen JR (1997) *Natur-og Kulturmiljøet på Jan Mayen* (Norsk Polarinstituttt, Oslo).
54. Gamisans J, Jeanmonod D (1995) La flore de Corse: Bilan des connaissances, intérêt patrimonial et état de conservation. *Ecol Mediterr* 21(1-2):135–148.
55. Groombridge B (1992) *Global Biodiversity. Status of the Earth's Living Resources* (Chapman & Hall, London).
56. Hansen A (1980) Eine Liste der Flora der Inseln Kos, Laymnos, Pserimos, Telendos und Nachbar-Inselchen (Ostägäis, Griechenland). *Biol Gallo-Hell* 9(1):3–105.
57. Harvey LE (1994) Spatial patterns of inter-island plant and bird species movements in the Galápagos Islands. *J R Soc N Z* 24(1):45–63.
58. Hnatiuk RJ (1993) Subantarctic Islands: Species lists. *Flora of Australia Online: Oceanic Islands Excluding Norfolk and Lord Howe Islands* (Department of Sustainability, Environment, Water, Population and Communities, Australian Government, Canberra, ACT, Australia). Available at http://www.anbg.gov.au/abrs/online-resources/flora/stddisplay.xsql?pnid=54674. Accessed April 7, 2011.
59. Hobohm C (2000) Plant species diversity and endemism on islands and archipelagos, with special reference to the Macaronesian Islands. *Flora* 195(1):9–24.
60. Hoffmann A, Teillier S (1991) La flora de la isla San Felix (Archipielago de las Desventuradas, Chile). *Gayana Bot* 48(1–4):89–99.
61. Jahn R, Schönfelder P (1995) *Exkursionsflora für Kreta* [Ulmer (Eugen), Stuttgart].
62. Johnson M, Mason L, Raven P (1968) Ecological parameters and plant species diversity. *Am Nat* 102(926):297–306.
63. Johnson MP, Simberloff DS (1974) Environmental determinants of island species numbers in the British Isles. *J Biogeogr* 1(3):149–154.
64. Lawesson JE, Adsersen H, Bentley P (1987) *An Updated and Annotated Check List of the Vascular Plants of the Galápagos Islands* (University of Aarhus Press, Aarhus, Denmark).
65. Lawesson JE, Skov F (2002) The phytogeography of Denmark revisited. *Plant Ecol* 158(1):113–122.
66. Levin GA, Moran R (1989) The vascular flora of Socorro, Mexico. *Memoirs of the San Diego Society of Natural History* 16:1–71.
67. Lowry PP, II (1996) Diversity, endemism, and extinction in the flora of New Caledonia: A review. *Rare, Threatened, and Endangered Floras of Asia and the Pacific Rim*. Academica Sinica Monograph Series, eds Peng CI, Lowry PP, II (Taipei Institute of Botany, Taipei, Taiwan), Vol 16, pp 181–206.
68. MacDonald IAW, Cooper J (1995) *Insular Lessons for Global Biodiversity Conservation with Particular Reference to Alien Invasions. Islands: Biological Diversity and Ecosystem Function, Ecological Studies*, eds Vitousek PM, Loope LL, Adsersen H (Springer, Berlin), Vol 115, pp 189–203.
69. Malyshev LI (1994) Prognosis of spatial diversity and degree of knowledge of the Siberian flora. *Bioraznoobrazie: Stepeń Taksonomicheskoi Izuchennosti*, ed Sokolov VEE (Nauka, Moscow), pp 42–52.
70. McMaster RT (2005) Factors influencing vascular plant diversity on 22 islands off the coast of eastern North America. *J Biogeogr* 32(3):475–492.
71. Médail F, Quézel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. *Ann Mo Bot Gard* 84(1):112–127.
72. Medail F, Verlaque R (1997) Ecological characteristics and rarity of endemic plants from southeast France and Corsica: Implications for biodiversity conservation. *Biol Conserv* 80(3):269–281.
73. Médail F, Vidal E (1998) Organisation de la richesse et de la composition floristique d'îles de la Méditerranée occidentale (sud-est de la France). *Can J Bot* 76(2): 321–331.
74. Meyer J-Y (2004) Threat of invasive alien plants to native flora and forest vegetation of eastern Polynesia. *Pac Sci* 58(3):357–375.
75. Moody A (2000) Analysis of plant species diversity with respect to island characteristics on the Channel Islands, California. *J Biogeogr* 27(3):711–723.
76. Panitsa M, Tzanoudakis D (2001) A floristic investigation of the islet groups Arki and Lipsi (East Aegean Area, Greece). *Folia Geobot* 36(3):265–279.

77. Pietsch TW, et al. (2003) Biodiversity and biogeography of the islands of the Kuril Archipelago. *J Biogeogr* 30(9):1297–1310.
78. Price JP (2004) Floristic biogeography of the Hawaiian Islands: Influences of area, environment and paleogeography. *J Biogeogr* 31(3):487–500.
79. Rannie WF (1986) Summer air temperature and number of vascular species in Arctic Canada. *Arctic* 39(2):133–137.
80. Renvoize S (1975) A floristic analysis of the western Indian Ocean coral islands. *Kew Bull* 30(1):133–152.
81. Roos MC, Keßler PJA, Gradstein RS, Baas P (2004) Species diversity and endemism of five major Malesian islands: Diversity-area relationships. *J Biogeogr* 31:1893–1908.
82. Sachet M-H (1962) Flora and vegetation of Clipperton Island. *Proc Calif Acad Sci* 31(10):249–307.
83. Simberloff DS (1970) Taxonomic diversity of island biotas. *Evolution* 24(1):23–47.
84. Snogerup S, Snogerup B, Phitos D, Kamari G, Anagnostopoulos A (1991) Flora and vegetation of Kira Panagia, N. Sporades, Greece. *Botanika Chronika* 10:547–566.
85. Sosa V, Dávila P (1994) Una evaluación del conocimiento florístico de México. *Ann Mo Bot Gard* 81(4):749–757.
86. Stuessy TF, Crawford DJ, Marticorena C, Rodriguez R (1998) Island biogeography of angiosperms of the Juan Fernandez archipelago. *Evolution and Speciation of Island Plants*, eds Stuessy TF, Ono M (Cambridge Univ Press, Cambridge), pp 121–138.
87. Sun B-Y, Stuessy TF (1998) Prelimary observations on the evolution of endemic angiosperms of Ullung Island, Korea. *Evolution and Speciation of Island Plants*, eds Stuessy TF, Ono M (Cambridge Univ Press, Cambridge), pp 181–202.
88. Sykes WR (1981) The vegetation of Late Island, Tonga. *Allertonia* 2(6):323–353.
89. Thaman RR (1992) Vegetation of Nauru and the Gilbert islands: Case studies of poverty, degradation, disturbance, and displacement. *Pac Sci* 46(2):128–158.
90. Turland NJ, Chiiton L, Press JR (1993) *Flora of the Cretan Area. Annotated Checklist and Atlas* (The Natural History Museum, London).
91. Whistler WA (1983) Vegetation and flora of the Aleipata Islands, Western Samoa. *Pac Sci* 37(3):227–249.
92. Wright DH (1983) Species-energy theory: An extension of species-area theory. *Oikos* 41(3):496–506.
93. Young SB (1971) The vascular flora of Saint Lawrence Island, with special reference to floristic zonation in the Arctic regions. Contributions from the Gray Herbarium of Harvard University 201:11–115.
94. Zanoni TA, Buck WR (1999) Navassa Island and its flora. 2. Checklist of the vascular plants. *Brittonia* 51(4):389–394.
95. Christmas Island National Park (2002) *Third Christmas Island National Park Management Plan* (Parks Australia North, Christmas Island, Australia).
96. Du Puy DJ (1993) Christmas Island: Species lists. *Flora of Australia Online: Oceanic Islands Excluding Norfolk and Lord Howe Islands* (Department of Sustainability, Environment, Water, Population and Communities, Australian Government, Canberra, ACT, Australia). Available at www.anbg.gov.au/abrs/online-resources/flora/stddisplay.xsql?pnid=54674. Accessed April 6, 2011.
97. McCrea J (2003) *Inventory of the Land Conservation Values of the Houtman Abrolhos Islands* (Department of Fisheries, Government of Western Australia, Perth, WA, Australia).
98. Junak S, Philbrick R, Chaney S, Clark R (1997) *A Checklist of Vascular Plants of Channel Islands National Park* (Southwest Parks and Monuments Association, Tucson, AZ).
99. Jaramillo Díaz P, Guézou A (2011) CDF checklist of Galapagos vascular plants. *Charles Darwin Foundation Galapagos Species Checklist*, eds Bungartz F, et al. (Charles Darwin Foundation, Puerto Ayora, Galapagos, Ecuador). Available at www.darwinfoundation.org/datazone/checklists/vascular-plants/. Accessed February 25, 2011.
100. Wagner WL, Herbst DR, Lorence DH (2005) Flora of the Hawaiian Islands Web site (Smithsonian Institution, Washington, DC). Available at http://botany.si.edu/pacificislandbiodiversity/hawaiianflora/index.htm. Accessed October 16, 2010.
101. Robinson A, Canty P, Fotheringham D (2008) Investigator Group expedition 2006: Flora and vegetation. *Trans R Soc S Aust* 132(2):173–220.
102. Sandbakk BE, Alsos IG, Arnesen G, Elven R (1996) The Flora of Svalbard (University of Tromsø, Tromsø, Norway). Available at http://svalbardflora.net/. Accessed March 16, 2011.
103. Moran R (1996) *The Flora of Guadalupe Island, Mexico* (California Academy of Sciences, San Francisco).
104. CARMABI (2009) Dutch Caribbean Biodiversity Explorer (CARMABI Foundation, Willemstad, Curacao, Netherlands Antilles). Available at www.dcbiodata.net/explorer/home. Accessed June 24, 2011.
105. Greene S, Walton D (1975) An annotated check list of the sub-Antarctic and Antarctic vascular flora. *Polar Rec (Gr Brit)* 17(110):473–484.
106. Wellington Botanical Society (2008) *Native Vascular Plants of Great Barrier Island* (Wellington Botanical Society, Wellington, New Zealand).
107. Arechavaleta M, Zurita N, Marrero MC, Martín JL (2005) *Lista Preliminar de Especies Silvestres de Cabo Verde (Hongos, Plantas y Animales Terrestres)* (Consejería de Medio Ambiente y Ordenación Territorial, Gobierno de Canarias, Santa Cruz de Tenerife, Canary Islands, Spain).
108. Stalter R, Lamont EE (2006) The historical and extant flora of Sable Island, Nova Scotia, Canada. *J Torrey Bot Soc* 133(2):362–374.
109. Florence J, Chevillotte H, Ollier C, Meyer J-Y (2007) *Base de Données Botaniques Nadeaud de l'Herbier de la Polynésie Française (PAP)* (Centre IRD de Tahiti et Antenne IRD, Pape'ete, Tahiti, French Polynesia). Available at www.herbier-tahiti.pf. Accessed July 1, 2011.
110. Convey P, Lewis Smith R, Hodgson D, Peat H (2000) The flora of the South Sandwich Islands, with particular reference to the influence of geothermal heating. *J Biogeogr* 27(6):1279–1295.
111. Universitat de les Illes Balears (2007) *Herbario Virtual del Mediterráneo Occidental* (Àrea de Botànica, Departament de Biologia, Universitat de les Illes Balears, Palma, Illes Balears, Spain). Available at http://herbarivirtual.uib.es/cas-med/index.html. Accessed August 7, 2012.
112. Whistler WA (1998) *A Study of the Rare Plants of American Samoa* (US Fish and Wildlife Service, Honolulu, HI).
113. Gage S, Joneson SL, Barkalov VYU, Eremenko NA, Takahashi H (2006) A newly compiled checklist of the vascular plants of the Habomais, the Little Kurils. *Bull Hokkaido Univ Mus* 3:67–91.
114. Takahashi H, et al. (2006) A floristic study of the vascular plants of Kharimkotan, Kuril Islands. *Bull Hokkaido Univ Mus* 3:41–66.
115. Franklin J, Keppel G, Whistler WA (2008) The vegetation and flora of Lakeba, Nayau and Aiwa islands, central Lau Group, Fiji. *Micronesica* 40(1):169–225.
116. Takahashi H, et al. (2002) A floristic study of the vascular plants of Raikoke, Kuril Islands. *Acta Phytotaxon Geobot* 53(1):17–33.
117. Hill MJ, ed (2002) *Biodiversity Surveys and Conservation Potential of Inner Seychelles Islands* (Smithsonian Institution, Washington, DC).
118. Robinson AC, Canty PD, Wace NM, Barker RM (2003) The Encounter 2002 expedition to the Isles of St Francis, South Australia: Flora and vegetation. *Trans R Soc S Aust* 127(2):107–128.
119. Searle J, Madden S (2006) *Flora Assessment of South Stradbroke Island* (Gold Coast City Council, Gold Coast City, QLD, Australia).
120. de Lange P, Cameron E (1999) The vascular flora of Aorangi Island, Poor Knights Islands, northern New Zealand. *NZ J Bot* 37(3):433–468.
121. Kelly L (2006) The vascular flora of Huggins Island, Onslow County, North Carolina. *Castanea* 71(4):295–311.
122. Hill SR (1986) An annotated checklist of the vascular flora of Assateague Island (Maryland and Virginia). *Castanea* 51(4):265–305.
123. Dowhan JJ, Ron R (1989) Flora of Fire Island, Suffolk County, New York. *Bull Torrey Bot Club* 116(3):265–282.
124. Esler AE (1978) Botanical features of the Mokohinau Islands. *TANE* 24:187–197.
125. Butler BJ, Barclay JS, Fisher JP (1999) Plant communities and flora of Robins Island (Long Island), New York. *J Torrey Bot Soc* 126(1):63–76.
126. Brofas G, Karetsos G, Panitsa M, Theocharopoulos M (2001) The flora and vegetation of Gyali Island, SE Aegean, Greece. *Willdenowia* 31(1):51–70.
127. Byrd GV (1984) Vascular vegetation of Buldir Island, Aleutian Islands, Alaska, compared to another Aleutian Island. *Arctic* 37(1):37–48.
128. D'Arcy WG (1971) The island of Anegada and its flora. *Atoll Res Bull* 139:1–21.
129. Proctor GR (1980) Checklist of the plants of Little Cayman. Geography and ecology of Little Cayman. *Atoll Res Bull* 241(7):71–80.
130. Marquand ED (1901) *Flora of Guernsey and the Lesser Channel Islands: Namely Alderney, Sark, Herm, Jethou, and the Adjacent Islets* (Dulau, London).
131. Lester-Garland LV (1903) *A Flora of the Islands of Jersey: With a List of the Plants of the Channel Islands in General, and Remarks upon Their Distribution and Geographical Affinities* (West, Newman, London).
132. Egorova EM (1964) Flora of Shiashkotan Island. *Bull Main Bot Gard* 54:114–120.
133. Gerlach J (2003) The biodiversity of the granitic islands of Seychelles. *Phelsuma* 11(Supplement):1–47.
134. University of Kent (2012) Cook Islands Biodiversity and Ethnobiology Database (University of Kent, Kent, UK). Available at http://cookislands.pacificbiodiversity.net. Accessed April 12, 2012.
135. Burton RM (1991) A check-list and evaluation of the flora of Nisyros (Dodecanese, Greece). *Willdenowia* 20(1):15–38.
136. Raulerson L (2006) Checklist of plants of the Mariana Islands. *University of Guam Herbarium Contribution* 37:1–69.
137. Kamari G, Phitos D, Snogerup B, Snogerup S (1988) Flora and vegetation of Yioura, N Sporades, Greece. *Willdenowia* 17(1):59–85.
138. Case TJ, Cody ML, Ezcurra E (2002) *A New Island Biogeography of the Sea of Cortés* (Oxford Univ Press, New York).
139. Shaw JD, Spear D, Greve M, Chown SL (2010) Taxonomic homogenization and differentiation across Southern Ocean Islands differ among insects and vascular plants. *J Biogeogr* 37(2):217–228.
140. Kreft H, Jetz W, Mutke J, Kier G, Barthlott W (2008) Global diversity of island floras from a macroecological perspective. *Ecol Lett* 11(2):116–127.
141. Kreft H, Jetz W (2007) Global patterns and determinants of vascular plant diversity. *Proc Natl Acad Sci USA* 104(14):5925–5930.
142. Bivand R, et al. (2011) spdep: Spatial Dependence: Weighting Schemes, Statistics and Models, 0.5-56 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/spdep/index.html. Accessed January 11, 2013.
143. Wood SN (2006) *Generalized Additive Models: An Introduction with R* (Chapman & Hall/CRC, Boca Raton, FL).
144. Wood SN (2003) Thin plate regression splines. *J R Stat Soc B* 65(1):95–114.
145. Kissling WD, Carl G (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Glob Ecol Biogeogr* 17(1):59–71.
146. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York), 2nd Ed.
147. Wood SN, Augustin NH (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecol Modell* 157(2–3):157–177.
148. Barton K (2013) MuMIn: Multi-Model Inference, 1.9.0 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/MuMIn/index.html. Accessed January 24, 2013.
149. Canty A, Ripley B (2012) boot: Bootstrap R (S-Plus) Functions, 1.3-7 (Package for R Statistical Software). Available at http://cran.r-project.org/web/packages/boot/index.html. Accessed October 12, 2012.

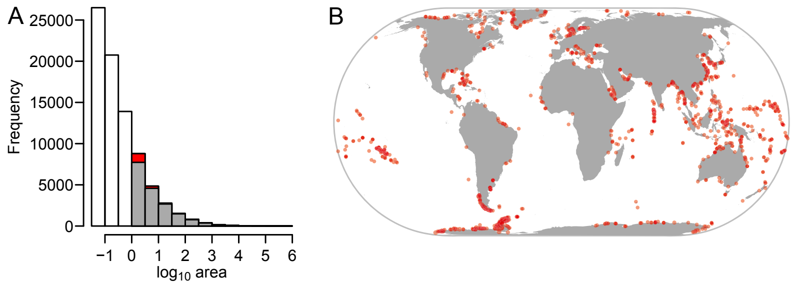

Fig. S1. (*A*) Histogram of area size of all 80,604 islands $>10^{-1.5}$ km^2 included in the GADM dataset (www.gadm.org/version1). The 17,883 >1-km^2 islands considered in the bioclimatic and physical characterization are shown in gray. The 1,509 islands >1 km^2 that were not included due to lacking climate data are colored red and mapped in *B*. These encompass mainly islands only slightly larger than 1 km^2 distributed more or less evenly across island rich regions of the globe, and include also all islands south of −60°, where no WorldClim climate data coverage is available.

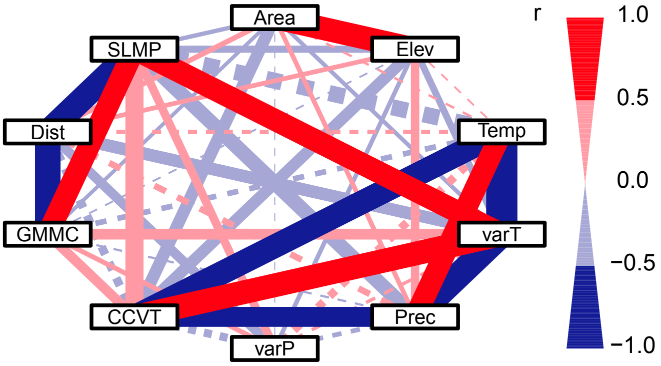

Fig. S2. Correlations among bioclimatic and physical variables for 17,883 islands >1 km^2 worldwide. Coefficients and *P* values were corrected for spatial autocorrelation. Solid lines denote significant relationships at $P < 0.05$, whereas dashed lines are nonsignificant. Abbreviations follow *SI Materials and Methods*.

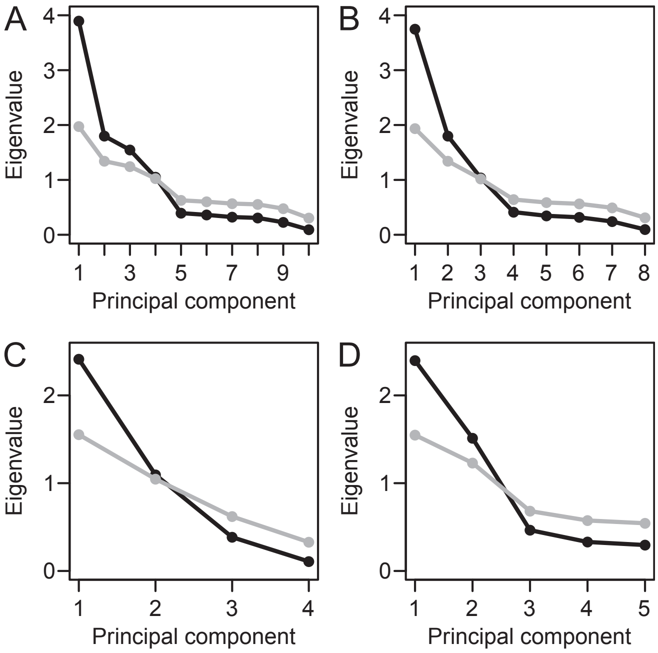

Fig. S3. Scree plots of eigenvalues (black) of principal components. PCAs were conducted for 17,883 islands >1 km^2 worldwide including (*A*) all 10 bioclimatic and physical variables used in the bioclimatic and physical characterization of the world's islands, (*B*) all variables but *Area* and *Elev*, (*C*) contemporary bioclimatic variables only (*Temp*, *varT*, *Prec*, *varP*), and (*D*) physical variables only (*Area*, *Elev*, *Dist*, *SLMP*, *GMMC*). Abbreviations follow *SI Materials and Methods*. Gray dots and lines indicate square roots of eigenvalues used for weighting in cluster analyses.

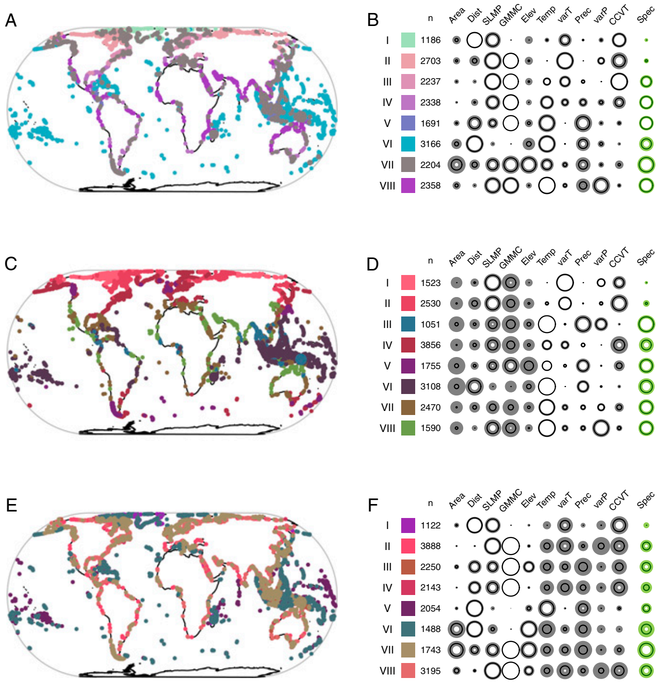

Fig. S4. PAM clustering using weighted PCA axes (Euclidean distance) based on (*A* and *B*) all 10 variables, (*C* and *D*) contemporary bioclimatic variables only (*Temp*, *varT*, *Prec*, *varP*), and (*E* and *F*) physical variables only (*Area*, *Dist*, *SLMP*, *GMMC*, *Elev*). Colors are calculated as mean RGB values of all constituent islands of each cluster based on the corresponding PCA colors in Fig. 2. Points were plotted in decreasing order of *Area*. Circles in *B*, *D*, and *F* indicate variable characteristics within clusters: circle, arithmetic mean; shaded ring, SD. Abbreviations follow *SI Materials and Methods*. *Spec* indicates predicted vascular plant species richness.

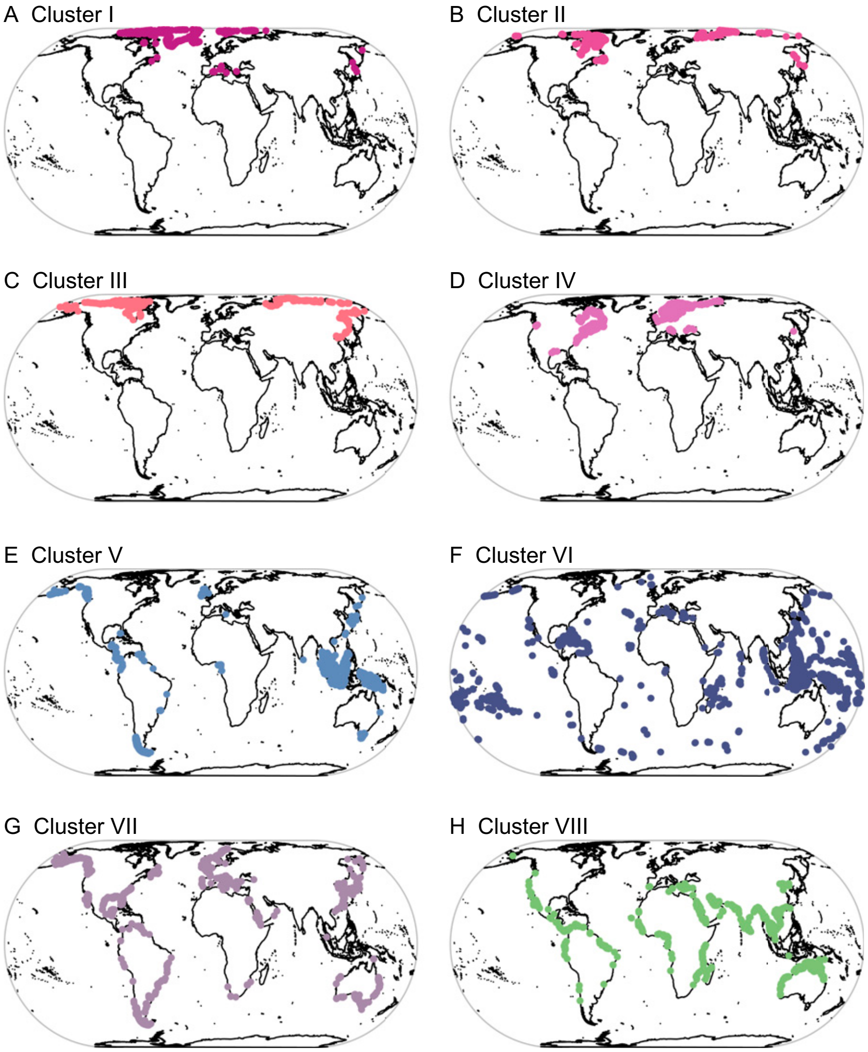

Fig. S5. Ecoregions derived from PAM clustering using weighted PCA axes (Euclidean distance) calculated for 17,883 islands >1 km^2 worldwide. Each map in *A*–*H* refers to one cluster (I-VIII) in Fig. 4. PCA was based on eight environmental variables (*Dist*, *SLMP*, *GMMC*, *Temp*, *varT*, *CCVT*, *Prec*, *varP*), excluding *Area* and *Elev*. Abbreviations follow *SI Materials and Methods*. Colors are calculated as mean red–green–blue (RGB) values of all constituent islands of each cluster based on the PCA colors in Fig. 2*E*.

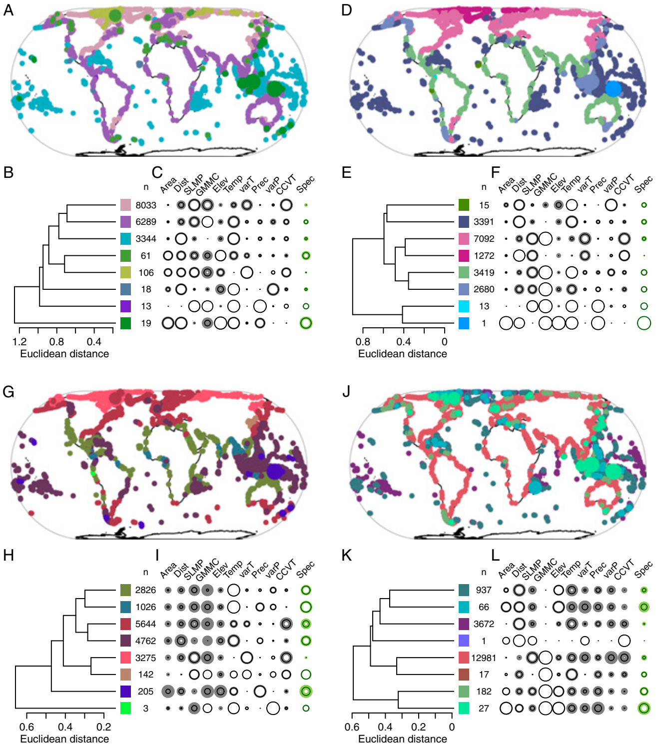

Fig. S6. UPGMA clustering using weighted PCA axes (Euclidean distance) based on (*A*–*C*) all 10 variables, (*D*–*F*) all variables but *Area* and *Elev*, (*G*–*I*) contemporary bioclimatic variables only (*Temp*, *varT*, *Prec*, *varP*), and (*J*–*L*) physical variables only (*Area*, *Dist*, *SLMP*, *GMMC*, *Elev*). Colors are calculated as mean RGB values of all constituent islands of each cluster based on the corresponding PCA colors in Fig. 2. Points were plotted in decreasing order of *Area*. Circles in *C*, *F*, *I*, and *L* indicate variable characteristics within clusters: circle, arithmetic mean; shaded ring, SD. Abbreviations follow *SI Materials and Methods*. *Spec* indicates predicted vascular plant species richness.

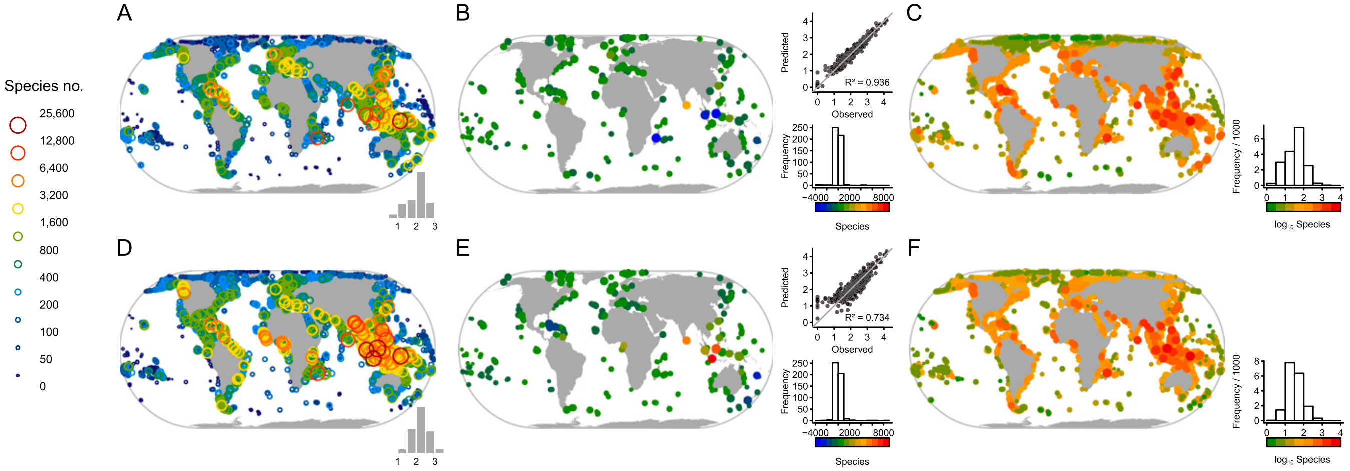

Fig. S7. Predicted pattern of species richness for vascular plants on 17,883 islands >1 km^2 worldwide based on model averaging of generalized additive models (*A*–*C*), and generalized linear models (*D*–*F*). *A* and *D* show the predicted species richness values, with circles plotted in order of increasing species richness and embedded histograms providing an impression of the distribution of predicted richness on a logarithmic scale [log$_{10}$(species richness + 1)]. *B* and *E* show residual species richness for the islands included in the training dataset (side plots provide biplots of observed vs. predicted values and corresponding pseudo-R^2 values). *C* and *F* show SEs of the richness predictions. In *B*, *C*, *E*, and *F*, values are plotted in order of decreasing frequency to show rare values on top of frequent values if points overlap. Both residuals and SEs were backtransformed [as log$_{10}$ (species richness + 1) was the modeled response variable] to represent actual species numbers.

Table S1. Summary statistics of 10 bioclimatic and physical variables for 17,883 islands >1 km^2 worldwide (untransformed)

Statistic	*Area*	*Dist*	*SLMP*	*GMMC*	*Elev*	*Temp*	*varT*	*Prec*	*varP*	*CCVT*
Min	1.00	0.00	0.12	0.00	0.00	−210.00	53.00	0.00	0.00	0.00
Mean	428.99	441.64	1.11	0.74	98.24	113.56	229.98	1,446.61	46.77	35.71
Median	4.08	50.84	1.17	1.00	13.00	111.00	196.00	1,191.00	40.00	15.71
Max	773,633.97	6,067.08	2.17	1.00	4,613.00	314.00	613.00	7,628.00	193.00	168.57
SD	11,372.76	811.98	0.41	0.44	251.19	146.57	127.35	1,152.59	29.26	37.68
Moran’s I	0.00	0.94	0.70	0.58	0.10	0.99	0.92	0.75	0.78	0.82
Unit	km^2	km	—	Yes/No	m	°C	°C	mm	—	m/y

Abbreviations follow *SI Materials and Methods*. All Moran’s I values are significant at $P < 0.001$ except for area ($P = 0.433$).

Table S2. Matrix of Pearson correlation coefficients among 10 bioclimatic and physical variables for 17,883 islands >1 km^2

	Area	*Elev*	*Temp*	*varT*	*Prec*	*varP*	*CCVT*	*GMMC*	*Dist*	*SLMP*
Elev	0.618***									
Temp	0.064	0.048								
varT	−0.111**	−0.182*	−0.835**							
Prec	0.126***	0.226***	0.583*	−0.692**						
varP	−0.027	−0.119**	0.219	0.090	−0.126					
CCVT	−0.326***	−0.312***	−0.633*	0.635*	−0.533*	−0.216				
GMMC	−0.101***	−0.048	−0.150	0.288**	−0.059	0.131*	0.264**			
Dist	0.141***	0.126**	0.120	−0.334*	0.170	−0.259**	−0.230	−0.685***		
SLMP	−0.115***	−0.191***	−0.431	0.590**	−0.393*	0.225*	0.486*	0.583***	−0.658***	
Age	0.049	−0.248*	0.195*	−0.209*	−0.182	0.203*	0.140	—	−0.416***	0.381***

Correlations with geologic age could only be calculated for a subset of 102 volcanic islands. A correlation coefficient between island age and *GMMC* is not given because age was only assessed for islands not connected to the mainland during the last glacial maximum. Correlation coefficients and significances were corrected for spatial autocorrelation: ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$. Abbreviations follow *SI Materials and Methods*.

Table S3. Axis scores of variables used in PCAs calculated for 17,883 islands >1 km^2 worldwide and axis eigenvalues, based on (A) all 10 bioclimatic and physical variables, (B) all variables but *Area* and *Elev*, (C) contemporary bioclimatic variables only, and (D) physical variables only

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
A										
Area	−0.165	0.014	−0.664	0.160	−0.598	0.205	−0.021	0.239	0.211	0.062
Elev	−0.185	0.013	−0.660	−0.059	0.501	−0.351	−0.232	−0.288	−0.087	−0.087
Temp	−0.374	−0.372	0.211	0.027	−0.315	−0.330	−0.251	−0.014	−0.018	−0.636
varT	0.441	0.177	−0.152	0.187	0.207	0.245	0.207	0.089	0.195	−0.724
Prec	−0.354	−0.247	0.025	−0.416	0.224	0.675	−0.128	−0.155	0.298	−0.072
varP	0.051	−0.423	0.072	0.744	0.261	0.134	−0.314	0.060	0.184	0.192
GMMC	0.268	−0.458	−0.142	−0.361	0.168	−0.066	−0.085	0.699	−0.199	0.022
Dist	−0.294	0.498	0.084	0.150	0.109	0.284	−0.417	0.350	−0.484	−0.115
SLMP	0.403	−0.265	−0.125	−0.021	−0.266	0.302	−0.201	−0.461	−0.579	−0.042
CCVT	0.394	0.248	0.083	−0.234	−0.139	−0.129	−0.710	−0.039	0.423	0.046
Eigenvalue	3.895	1.800	1.547	1.050	0.395	0.363	0.323	0.307	0.227	0.093
B										
Temp	−0.397	0.368	−0.044	0.546	−0.032	−0.006	0.096	0.630		
varT	0.460	−0.174	−0.176	−0.373	0.018	−0.053	−0.214	0.735		
Prec	−0.362	0.246	0.420	−0.403	0.580	0.156	−0.318	0.085		
varP	0.046	0.422	−0.762	−0.031	0.363	−0.186	−0.175	−0.200		
GMMC	0.282	0.461	0.372	−0.118	−0.037	−0.697	0.263	−0.011		
Dist	−0.303	−0.500	−0.153	−0.098	0.443	−0.335	0.552	0.110		
SLMP	0.418	0.268	0.036	0.019	0.299	0.566	0.585	0.024		
CCVT	0.390	−0.249	0.215	0.613	0.493	−0.139	−0.313	−0.054		
Eigenvalue	3.748	1.800	1.037	0.413	0.346	0.318	0.242	0.097		
C										
Temp	−0.583	0.261	−0.391	−0.663						
varT	0.608	0.076	0.347	−0.710						
Prec	−0.539	−0.202	0.813	−0.085						
varP	−0.003	0.941	0.255	0.223						
Eigenvalue	2.412	1.095	0.385	0.108						
D										
Area	−0.250	0.657	0.477	0.519	−0.098					
Elev	−0.252	0.661	−0.400	−0.578	0.074					
GMMC	0.524	0.257	−0.483	0.459	0.465					
Dist	−0.558	−0.202	0.099	0.046	0.798					
SLMP	0.537	0.156	0.608	−0.430	0.364					
Eigenvalue	2.397	1.512	0.465	0.330	0.295					

Abbreviations follow *SI Materials and Methods.*

Table S4. Summary statistics for clusters from PAM clustering using weighted PCA axes (Euclidean distance) based on (A) all 10 bioclimatic and physical variables, (B) all variables but *Area* and *Elev*, (C) contemporary bioclimatic variables only (*Temp*, *varT*, *Prec*, *varP*), and (D) physical variables only (*Area*, *Elev*, *Dist*, *SLMP*, *GMMC*)

Variable	I	II	III	IV	V	VI	VII	VIII
A								
Area	41 ± 217	44 ± 419	8 ± 17	4 ± 11	11 ± 28	124 ± 779	3,181 ± 32,250	24 ± 81
Dist	1,047 ± 328	132 ± 176	27 ± 73	73 ± 162	546 ± 473	1,486 ± 1,314	161 ± 297	35 ± 130
SLMP	1.23 ± 0.27	1.42 ± 0.27	1.44 ± 0.22	1.18 ± 0.26	0.86 ± 0.24	0.51 ± 0.2	1.08 ± 0.3	1.33 ± 0.27
GMMC	0 ± 0	0.99 ± 0.08	1 ± 0	0.99 ± 0.09	1 ± 0	0 ± 0	0.91 ± 0.29	0.96 ± 0.19
Elev	65 ± 139	40 ± 81	28 ± 49	20 ± 35	34 ± 56	145 ± 319	404 ± 466	22 ± 43
Temp	−7.5 ± 7.2	−11.2 ± 3.8	3.8 ± 4.5	14.8 ± 6.8	22.5 ± 8	23.8 ± 6.4	11 ± 8.9	25.9 ± 3.5
varT	30.7 ± 8.9	44.3 ± 5.9	30.6 ± 6.3	22.6 ± 7	10.4 ± 2.6	11.1 ± 3.4	18.2 ± 7.2	17.3 ± 5.5
Prec	446 ± 398	260 ± 128	830 ± 354	1,374 ± 677	2,791 ± 965	2,036 ± 973	2,003 ± 1,138	1,692 ± 1,290
varP	34.3 ± 18	55.8 ± 13.2	24.4 ± 8.7	42.7 ± 19.7	31.3 ± 21.6	40.3 ± 21.6	34.9 ± 21	98.8 ± 23.9
CCVT	57.3 ± 21.12	62.61 ± 38.84	97.55 ± 27.38	34.63 ± 23.63	11.13 ± 7.27	7.73 ± 3.89	16.27 ± 16.95	9.76 ± 4.54
Spec	18 ± 26	25 ± 29	147 ± 107	212 ± 120	231 ± 168	188 ± 231	469 ± 690	287 ± 227
B								
Area	438 ± 6197	649 ± 14,247	199 ± 5775	30 ± 302	1,334 ± 24,159	656 ± 11,894	90 ± 2,306	37 ± 174
Dist	1,036 ± 335	256 ± 183	34 ± 77	17 ± 48	418 ± 462	1,445 ± 1,309	86 ± 177	23 ± 81
SLMP	1.23 ± 0.27	1.23 ± 0.2	1.55 ± 0.23	1.44 ± 0.22	0.88 ± 0.25	0.52 ± 0.22	1.16 ± 0.28	1.35 ± 0.26
GMMC	0 ± 0	1 ± 0	1 ± 0.04	1 ± 0	1 ± 0	0 ± 0	0.99 ± 0.1	0.97 ± 0.17
Elev	113 ± 259	67 ± 151	29 ± 65	38 ± 89	162 ± 346	175 ± 392	98 ± 181	37 ± 96
Temp	−7.3 ± 7.4	−8.4 ± 5	−11.9 ± 5	3.6 ± 4.9	20 ± 9.2	23.6 ± 6.6	13.8 ± 7	25.9 ± 3.1
varT	30.6 ± 9	39 ± 4.7	47.8 ± 4.8	30.7 ± 6.7	11.1 ± 2.9	11.3 ± 3.7	21.5 ± 6.4	17.3 ± 5.5
Prec	468 ± 418	382 ± 243	225 ± 153	834 ± 382	2,855 ± 971	2,011 ± 994	1,428 ± 702	1,730 ± 1,277
varP	33.9 ± 18.1	45.8 ± 13.3	62.8 ± 13.9	25.2 ± 9.1	31.3 ± 22.3	41.1 ± 22.6	40.6 ± 18.1	98.8 ± 23.7
CCVT	54.44 ± 22.75	87.35 ± 32.12	35.71 ± 21.59	97.01 ± 29.19	9.33 ± 5.47	7.65 ± 3.94	29.75 ± 21.37	9.57 ± 4.59
Spec	23 ± 39	31 ± 39	35 ± 51	151 ± 129	334 ± 550	216 ± 378	261 ± 189	314 ± 260
C								
Area	321 ± 5939	431 ± 11,207	823 ± 23,872	169 ± 2743	711 ± 9,603	995 ± 19,228	72 ± 778	36 ± 187
Dist	144 ± 251	364 ± 467	310 ± 546	325 ± 602	161 ± 454	1,323 ± 1,298	251 ± 516	105 ± 291
SLMP	1.47 ± 0.24	1.36 ± 0.28	1.12 ± 0.39	1.19 ± 0.37	1.03 ± 0.31	0.68 ± 0.35	0.99 ± 0.37	1.27 ± 0.34
GMMC	0.87 ± 0.34	0.78 ± 0.41	0.82 ± 0.39	0.78 ± 0.41	0.93 ± 0.26	0.44 ± 0.5	0.67 ± 0.47	0.86 ± 0.34
Elev	35 ± 79	75 ± 180	81 ± 253	95 ± 232	219 ± 324	130 ± 353	83 ± 222	42 ± 130
Temp	−14.2 ± 2.5	−7.5 ± 4.6	26.7 ± 0.8	6.4 ± 5.8	8.2 ± 2.7	26.6 ± 1.1	21.9 ± 5.2	25 ± 4.4
varT	48.2 ± 4	38.1 ± 4.7	12.5 ± 3.5	24.5 ± 6.6	15.3 ± 3.9	9.3 ± 1.9	19.3 ± 6.9	19.3 ± 6.8
Prec	159 ± 64	377 ± 172	3,508 ± 957	978 ± 360	2,754 ± 849	2,559 ± 696	1,112 ± 527	1,057 ± 648
varP	66.2 ± 10	40.1 ± 11.3	83.9 ± 20	23.2 ± 7.9	25.2 ± 12.8	29.8 ± 11.7	59.4 ± 12.4	108.6 ± 19.7
CCVT	38.86 ± 22.77	79.63 ± 35.52	6.54 ± 3.13	61.72 ± 41.83	24.73 ± 25.94	7.46 ± 3.45	18.97 ± 16.39	12.3 ± 8.79
Spec	19 ± 20	36 ± 41	411 ± 506	177 ± 178	247 ± 279	294 ± 519	254 ± 210	254 ± 207
D								
Area	9 ± 42	3 ± 4	7 ± 7	3 ± 4	5 ± 6	1,817 ± 18,540	2,787 ± 32,024	17 ± 26
Dist	978 ± 403	5 ± 7	316 ± 359	332 ± 365	1,575 ± 1,437	1,191 ± 946	179 ± 296	11 ± 17
SLMP	1.22 ± 0.24	1.41 ± 0.25	0.98 ± 0.26	1.07 ± 0.3	0.49 ± 0.18	0.68 ± 0.36	1.1 ± 0.3	1.43 ± 0.24
GMMC	0 ± 0	1 ± 0	1 ± 0.06	1 ± 0	0 ± 0.02	0 ± 0	1 ± 0.02	1 ± 0.04
Elev	5 ± 15	5 ± 5	80 ± 82	4 ± 4	17 ± 26	466 ± 512	369 ± 410	54 ± 71
Temp	−0.4 ± 14.3	10.2 ± 14.5	11.7 ± 13.4	8.6 ± 15.9	23.5 ± 7.6	15.1 ± 15.1	10.9 ± 12.6	9.2 ± 13.7
varT	27.1 ± 10.2	28.7 ± 12.5	20.1 ± 11.4	24.5 ± 13.9	11.3 ± 4.2	16.7 ± 10.3	20.9 ± 11.1	27.3 ± 12.2
Prec	676 ± 656	1,074 ± 991	1,858 ± 1,250	1,428 ± 1,201	1,915 ± 953	1,749 ± 1,222	1,842 ± 1,275	1,237 ± 1,024
varP	41.3 ± 24.9	56.1 ± 34	41.6 ± 25.6	42.5 ± 25.4	40 ± 21.4	39.6 ± 24.1	42 ± 28	54.1 ± 31.8
CCVT	50.61 ± 25.42	48.21 ± 43.15	34.94 ± 36.06	47.91 ± 42.15	10.89 ± 12.97	11.4 ± 15.28	19.36 ± 22.16	43.79 ± 40.19
Spec	43 ± 65	144 ± 112	177 ± 136	110 ± 92	102 ± 61	335 ± 536	512 ± 675	241 ± 205

Abbreviations follow *SI Materials and Methods*. *Spec* indicates predicted vascular plant species richness. Note that cluster numbers in A to D do not correspond to each other but refer to Fig. 4 and Fig. S5 for B and to Fig. S4 for A, C, and D.

Table S5. Variable importance of all 10 bioclimatic and physical variables, mainland plant species richness and interaction terms for vascular plant species richness on 475 islands >1 km^2 worldwide

Variable	GAM	GLM
Area	1.00	1.00
Dist	0.83	1.00
SLMP	1.00	1.00
GMMC	0.74	0.39
Elev	0.56	0.30
Temp	1.00	1.00
varT	0.30	0.26
CCVT	0.31	0.39
Prec	1.00	1.00
varP	0.27	0.27
SRML	0.33	1.00
Area:*Temp*	0.97	0.96
Prec:*Temp*	0.29	0.87
Dist:*Area*	0.70	0.67
Dist:*SRML*	0.13	1.00

Variable importance was assessed as cumulative AICc weights based on multimodel inference for generalized additive models (GAMs) and generalized linear models (GLMs). In addition to the here listed variables, all candidate GAMs included an isotropic smooth of *Lat* and *Long* on a sphere to account for spatial patterns. Abbreviations follow *SI Materials and Methods*.

Other Supporting Information Files

Dataset S1 (CSV)