Data accompanying the publication "Weigelt, P., Jetz, W. & Kreft, H. (2013) Bioclimatic and physical characterization of the world's islands. Proceedings of the National Academy of Sciences of the United States of America, doi: 10.1073/pnas.1306309110."

The table contains bioclimatic and physical characteristics, ordination and clustering results, as well as species richness predictions for the 17,883 islands > 1 km² investigated in Weigelt *et al.* (2013) as comma separated text file. The table is sorted by IDs (column *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). Longitude (*Long*) and Latitude (*Lat*) are given for the polygon mass centroids. ISO 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 5 characters) and country names listed separated by semicolons. Where applicable, an archipelago name (*Archip*) was assigned.

Island names (*Island*) were assigned using the NGA GEOnet Names Server (earth-info.nga.mil/gns/html/index.html; indicated as "gns" in column *Gazetteer*) for all regions but the United States, and the USGS Geographic Names Information System (geonames.usgs.gov/index.html; indicated as "gnis") for the United States. Original ID (*Name_ID*) and geographic coordinates (*Name_long* and *Name_lat*) were adopted from the gazetteer. Assignment method as explained in *SI Materials and Methods* in Weigelt *et al.* (2013) is indicated in column *Name_meth.* Alternative names (*Name_alt*) and the number of names (*No_names*) are given. The island names may help to find data for certain islands in our data set but due to their insecure assignment they must not be used for automated match-ups. The island coordinates should be used instead as a spatially explicit reference.

We give island area (*Area*; km²) calculated for each GADM polygon. As measures of island isolation, we provide the distance from an islands mass centroid to the nearest mainland coast (*Dist*, km) 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; Weigelt & Kreft, 2013). Estimates of whether an island was connected to the mainland during the LGM or not (*GMMC*) were based on global bathymetry data (Amante & Eakins, 2009) assuming a sea level decrease of -122 m at 18,000 a BP (Miller *et al.*, 2005). Maximum elevation a.s.l. of each island (*Elev*; m) was extracted from the digital elevation model at 30 sec. resolution provided in WorldClim (Hijmans *et al.*, 2005). For islands that did not fully enclose a 30 sec. WorldClim raster cell, we applied a 1 km buffer as indicated in column *Buffer*.

Bioclimatic variables were extracted from WorldClim (BIO1, BIO7, BIO12, and BIO15) in a similar manner to *Elev*. We provide maximum values per island polygon of annual mean temperature (*Temp*; °C) and annual precipitation (*Prec*; mm) and minimum values of the annual temperature range (*varT*; °C) and the coefficient of variation in monthly precipitation (*varP*). For a region of 129 islands > 1 km² 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. Islands with modeled temperature data are marked in column *modeled_T*. We calculated climate change velocity (*CCVT*; m/y) since the LGM 21,000 years ago following Loarie *et al.* (2009) and Sandel *et al.* (2011). Please see the detailed explanation in *SI Materials and Methods* in Weigelt *et al.* (2013).

Column names starting with "*PAM*" refer to clustering results from non-hierarchical 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 (PCA) 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 ten variables; *cli* = contemporary bioclimatic variables; *geo* = physical variables).

We developed statistical predictions of the species richness of native vascular plants on all 17,883 islands > 1 km². We built on existing richness data for vascular plants, including all 345 islands from Weigelt & Kreft (2013) that could be assigned to a single GADM polygon and 130 islands for which data was available from published floras, checklists and online databases. Following the rational of Kreft *et al.* (2008) we used as predictors the ten bioclimatic and physical variables presented in Weigelt *et al.* (2013). As additional predictor we included the species richness of the closest mainland grid cell (column *SRML*) derived from the co-kriging based estimates provided by Kreft & Jetz (2007). Predicted species numbers together with their standard errors can be found in the columns *SR_GLM* and *SR_SE_GLM* for predictions from generalized linear models and *SR_GAM* and *SR_SE_GAM* for predictions from generalized additive models. Both species richness and standard errors were back-transformed (as $\log_{10}$ (species richness + 1) was the modeled response variable) to represent actual species numbers.

References

Amante, C. & Eakins, B.W. (2009) ETOPO1 - 1 arc-minute global relief model: procedures, data sources and analysis. In: *NOAA Technical Memorandum NESDIS NGDC-24*, p. 19. National Geophysical Data Center, Boulder, Colorado.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.

Kreft, H. & Jetz, W. (2007) Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 5925-5930.

Kreft, H., Jetz, W., Mutke, J., Kier, G. & Barthlott, W. (2008) Global diversity of island floras from a macroecological perspective. *Ecology Letters*, **11**, 116-127.

Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009) The velocity of climate change. *Nature*, **462**, 1052-1055.

Miller, K.G., Kominz, M.A., Browning, J.V., Wright, J.D., Mountain, G.S., Katz, M.E., Sugarman, P.J., Cramer, B.S., Christie-Blick, N. & Pekar, S.F. (2005) The Phanerozoic record of global sea-level change. *Science*, **310**, 1293-1298.

Sandel, B., Arge, L., Dalsgaard, B., Davies, R., Gaston, K., Sutherland, W. & Svenning, J.C. (2011) The influence of Late Quaternary climate-change velocity on species endemism. *Science*, **334**, 660-664.

Weigelt, P. & Kreft, H. (2013) Quantifying island isolation – insights from global patterns of insular plant species richness. *Ecography*, **36**, 417-429.

Weigelt, P., Jetz, W. & Kreft, H. (2013) Bioclimatic and physical characterization of the world's islands. *Proceedings of the National Academy of Sciences of the United States of America*, **doi: 10.1073/pnas.1306309110**