Mapping species richness of plant families in European vegetation

Martin Večeřa, Irena Axmanová, Josep Padullés Cubino, Zdeňka Lososová, Jan Divíšek, Ilona Knollová, Svetlana Aćić, Idoia Biurrun, Steffen Boch, Gianmaria Bonari, et al.

DR. MARTIN VEČEŘA (Orcid ID : 0000-0001-8507-791X)

MRS. IRENA AXMANOVÁ (Orcid ID : 0000-0001-9440-7976)

DR. ZDEŇKA LOSOSOVÁ (Orcid ID : 0000-0001-9152-7462)

DR. JAN DIVÍŠEK (Orcid ID : 0000-0002-5127-5130)

DR. IDOIA BIURRUN (Orcid ID : 0000-0002-1454-0433)

DR. STEFFEN BOCH (Orcid ID : 0000-0003-2814-5343)

DR. GIANMARIA BONARI (Orcid ID : 0000-0002-5574-6067)

DR. ANDRAŽ ČARNI (Orcid ID : 0000-0002-8909-4298)

DR. MARIA LAURA CARRANZA (Orcid ID : 0000-0001-5753-890X)

PROF. ALESSANDRO CHIARUCCI (Orcid ID : 0000-0003-1160-235X)

DR. JÜRGEN DENGLER (Orcid ID : 0000-0003-3221-660X)

DR. FLORIAN JANSEN (Orcid ID : 0000-0002-0331-5185)

DR. BORJA JIMÉNEZ-ALFARO (Orcid ID : 0000-0001-6601-9597)

DR. REMIGIUSZ PIELECH (Orcid ID : 0000-0001-8879-3305)

Article type : Special feature article

Francesco Maria Sabatini

Mapping species richness of plant families in European vegetation

Running title: Plant-family richness in European vegetation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/JVS.13035

Martin Večeřa[1]*, Irena Axmanová[1], Josep Padullés Cubino[1], Zdeňka Lososová[1], Jan Divíšek[1,2], Ilona Knollová[1], Svetlana Aćić[3], Idoia Biurrun[4], Steffen Boch[5], Gianmaria Bonari[6], Juan Antonio Campos[4], Andraž Čarni[7,8], Maria Laura Carranza[9], Laura Casella[10], Alessandro Chiarucci[11], Renata Ćušterevska[12], Pauline Delbosc[13], Jürgen Dengler[14,15,16], Federico Fernández-González[17], Jean-Claude Gégout[18], Ute Jandt[16,19], Florian Jansen[20], Anni Jašková[1], Borja Jiménez-Alfaro[21], Anna Kuzemko[22], Maria Lebedeva[23], Jonathan Lenoir[24], Tatiana Lysenko[25,26], Jesper Erenskjold Moeslund[27], Remigiusz Pielech[28,29], Eszter Ruprecht[30], Jozef Šibík[31], Urban Šilc[7], Željko Škvorc[32], Grzegorz Swacha[33], Irina Tatarenko[34], Kiril Vassilev[35], Thomas Wohlgemuth[36], Sergey Yamalov[23] & Milan Chytrý[1]

* Corresponding author: martinvec@seznam.cz

[1]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

[2]Department of Geography, Faculty of Science, Masaryk University, Brno, Czech Republic

[3]Department of Botany, Faculty of Agriculture, University of Belgrade, Belgrade-Zemun, Serbia

[4]Department of Plant Biology and Ecology, Faculty of Science and Technology, University of the Basque Country UPV/EHU, Leioa, Spain

[5]Biodiversity and Conservation Biology Research Unit, WSL Swiss Federal Research Institute, Birmensdorf, Switzerland

[6]Faculty of Science and Technology, Free University of Bozen-Bolzano, Bolzano, Italy

[7]Research Center of the Slovenian Academy of Sciences and Arts, Institute of Biology, Ljubljana, Slovenia

[8]University of Nova Gorica, School for Viticulture and Enology, Vipava, Slovenia.

[9]Department of Bioscience and Territory (Envixlab), University of Molise, Campobasso, Italy

[10]ISPRA - Italian National Institute for Environmental Protection and Research

[11]BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum - University of Bologna

[12]Institute of Biology, Faculty of Natural Sciences and Mathematics, University of Ss. Cyril and Methodius, Skopje, North Macedonia

[13]EA 7462 Géoarchitecture, Université de Bretagne Occidentale UFR Sciences et Techniques, Brest, France

[14]Vegetation Ecology, Institute of Natural Resource Sciences (IUNR), Zurich University of Applied Sciences (ZHAW), Wädenswil, Switzerland

[15]Plant Ecology, Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany

[16]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

[17]Institute of Environmental Sciences, University of Castilla-La Mancha, Toledo, Spain

[18]Université de Lorraine, AgroParisTech, INRAE, Silva, Nancy, France

[19]Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle, Germany

[20]Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany

[21]Research Unit of Biodiversity (CSIC, UO, PA), Oviedo University, Mieres, Spain
[22]Department of Geobotany and Ecology, M.G. Kholodny Institute of Botany, NAS of Ukraine, Kyiv, Ukraine
[23]South Ural Botanical Garden-Institute, Ufa, Russian Federation
[24]Ecologie et Dynamique des Systèmes Anthropisés (EDYSAN, UMR 7058 CNRS), Université de Picardie Jules Verne, Amiens, France
[25]Laboratory of Vegetation Science, Komarov Botanical Institute RAS, Saint-Petersburg, Russian Federation
[26] Laboratory of Phytodiversity Problems, Institute of Ecology of the Volga River Basin RAS - Branch of the Samara Scientific Center RAS, Togliatti, Russian Federation
[27]Department of Bioscience, Faculty of Technical Sciences, Aarhus University, Aarhus, Denmark
[28]Department of Forest Biodiversity, Faculty of Forestry, University of Agriculture in Kraków, Kraków, Poland
[29]Foundation for Biodiversity Research, Wrocław, Poland
[30]Hungarian Department of Biology and Ecology, Faculty of Biology and Geology, Babeş-Bolyai University, Cluj-Napoca, Romania
[31]Plant Science and Biodiversity Center, Slovak Academy of Sciences, Bratislava, Slovakia
[32]Faculty of Forestry, University of Zagreb, Zagreb, Croatia
[33]Botanical Garden, University of Wrocław, Wrocław, Poland
[34]School of Environment, Earth and Ecosystem Sciences, Faculty of STEM, Open University, UK
[35]Department of Plant and Fungal Diversity and Resources, Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, Bulgaria
[36]Forest Dynamics Research Unit, WSL Swiss Federal Research Institute, Birmensdorf, Switzerland

MV https://orcid.org/0000-0001-8507-791X
IA https://orcid.org/0000-0001-9440-7976
JPC https://orcid.org/0000-0002-2283-5004
ZL https://orcid.org/0000-0001-9152-7462
JDi https://orcid.org/0000-0002-5127-5130
IK https://orcid.org/0000-0003-4074-789X
SA https://orcid.org/0000-0001-6553-3797
IB https://orcid.org/0000-0002-1454-0433
SB https://orcid.org/0000-0003-2814-5343
GB https://orcid.org/0000-0002-5574-6067
JAC https://orcid.org/0000-0001-5992-2753
AČ https://orcid.org/0000-0002-8909-4298

MLC https://orcid.org/0000-0001-5753-890X

LC https://orcid.org/0000-0003-2550-3010

AC https://orcid.org/0000-0003-1160-235X

RĆ

PD https://orcid.org/0000-0002-9827-3415

JDe https://orcid.org/0000-0003-3221-660X

FF-G https://orcid.org/0000-0003-1234-4065

J-CG

UJ https://orcid.org/0000-0002-3177-3669

FJ https://orcid.org/0000-0002-0331-5185

AKJ

BJA

AK https://orcid.org/0000-0002-9425-2756

ML https://orcid.org/0000-0002-5020-527X

JL https://orcid.org/0000-0003-0638-9582

TL https://orcid.org/0000-0001-6688-1590

JEM https://orcid.org/0000-0001-8591-7149

RP http://orcid.org/0000-0001-8879-3305

ER https://orcid.org/0000-0003-0122-6282

JŠ https://orcid.org/0000-0002-5949-862X

UŠ https://orcid.org/0000-0002-3052-699X

ŽŠ https://orcid.org/0000-0002-2848-1454

GS https://orcid.org/0000-0002-6380-2954

IT https://orcid.org/0000-0001-6835-2465

KV https://orcid.org/0000-0003-4376-5575

TW https://orcid.org/0000-0002-4623-0894

SY https://orcid.org/0000-0002-7052-522X

MC https://orcid.org/0000-0002-8122-3075

Funding information

MV, IA, JPC, ZL, IK, AJ and MC were funded by the Czech Science Foundation, programme EXPRO (project no. 19-28491X); JDi by the Czech Science Foundation (18-02773S); IB by the Basque Government (IT936-16); AČ by the Slovenian Research Agency (ARRS, P1-0236); JŠ by the Slovak Research and Development Agency (APVV 16-0431); KV by the National Science Fund (Contract DCOST 01/7/19.10.2018).

Abstract

Aims: Biodiversity is traditionally studied mostly at the species level, but biogeographical and macroecological studies at higher taxonomic levels can provide valuable insights into the evolutionary processes at large spatial scales. Our aim was to assess the representation of vascular plant families within different vegetation formations across Europe.

Location: Europe.

Methods: We used a dataset of 816,005 vegetation plots from the European Vegetation Archive (EVA). For each plot, we calculated the relative species richness of each plant family as the number of species belonging to that family divided by the total number of species. We mapped the relative species richness, averaged across all plots in 50 km × 50 km grid cells, for each family and broad habitat groups: forests, grasslands, scrub and wetlands. We also calculated the absolute species richness and the Shannon diversity index for each family.

Results: We produced 522 maps of mean relative species richness for a total of 152 vascular plant families occurring in forests, grasslands, scrub and wetlands. We found distinct spatial patterns for many combinations of families and habitat groups. The resulting series of 522 maps has been made freely available, both as images and GIS layers.

Conclusions: The distinct spatial patterns revealed in the maps suggest that the relative species richness of plant families at the community level reflects the evolutionary history of individual families. We believe that the maps and associated data can inspire further biogeographical and macroecological studies and strengthen the ongoing integration of phylogenetic, functional and taxonomic diversity concepts.

Keywords

biodiversity, Europe, European Vegetation Archive (EVA), family-level taxonomy, macroecology, map, plant community, relative species richness, vascular plant family, vegetation

Introduction

The recent accumulation of biodiversity data (e.g. Dornelas et al., 2018; Bruelheide et al., 2019; König et al., 2019) and the development of new analytical techniques (e.g. Webb et al., 2008; Rabosky et al., 2014; Chytrý et al., 2020) has drawn considerable attention from scientists and practitioners to other taxonomic and phylogenetic levels of biodiversity beyond the traditional species level. For instance, the evolutionary relationships among species within ecological communities (e.g. Webb et al., 2002; Lososová et al., 2015, 2021) or the historical biogeography of phylogenetic lineages (e.g. Chacón et al., 2017; Jin et al., 2020) have become prominent topics in current biodiversity research. The recently assembled vegetation-plot databases with continental or global coverage (Chytrý et al., 2016; Dengler et al., 2018; Bruelheide et al., 2019) provide excellent opportunities to link high-quality community-ecology data with phylogenies and functional traits (e.g. Bruelheide et al., 2018). Integrating community ecology, phylogenetics, historical biogeography and global-change research has been suggested as a promising way to (i) increase our understanding of community assembly rules (Webb et al., 2002; Cavender-Bares et al., 2009); (ii) forecast changes in diversity and dynamics of communities (e.g. Manning et al., 2015); and (iii) explain broad-scale biodiversity patterns (Donoghue, 2008; Beck et al., 2012). This knowledge may be of critical importance also for policy-making and setting conservation priorities, especially in the context of the current global biodiversity crisis (IPBES, 2019; European Commission, 2019).

Families, more than other taxonomic levels in the tree of life of plants, have been traditionally used by botanists and ecologists to group taxa based on their common characteristics (e.g. Cronquist, 1968). Recently, the family level received a considerable interest in ecological and biogeographical research (Laiolo et al., 2020). First, studies exploring the phylogenetic structure of communities across large areas or multiple habitat types often rely on family-level phylogenies as a backbone (e.g. Power et al., 2017; Qian et al., 2017). This is mainly due to the lack of gene sequences for many species, which does not allow producing genus-level phylogenies for extensive species lists (Qian & Zhang, 2014). Second, various biogeographical and ecological aspects of individual plant families have been studied comprehensively, often emphasizing the dispersal and migration pathways resulting in the current distribution patterns (e.g. Chen et al., 2014; Chacón et al., 2017; Sainge et al., 2017; Jin et al., 2020). Third, higher-level taxonomy including the family level provides a relevant interpretation framework for the studies of plant functional traits (e.g. plant height, specific leaf area or seed mass). Such studies often focus on the phylogenetic community structure, exploring the extent to which these traits are phylogenetically conserved or convergent (Cavender-Bares et al., 2009).

The aforementioned studies indicate an increasing need to understand the distribution of plant taxa above the species level. One way to accommodate this is by examining patterns of species richness within plant

families in different habitat types across large spatial extents (e.g. Pausas & Lamont, 2018). For Europe, no distribution maps of vascular plant families have yet been published (except for Malyshev, 1972 covering only the European part of the former USSR). The maps in Atlas Florae Europaeae (AFE; Jalas & Suominen, 1972–1994; Jalas et al., 1996, 1999; Kurtto et al., 2004, 2007, 2010, 2013, 2018) currently cover 4,664 species and subspecies of vascular plants from 79 families, which is only about 25% of the total European flora. The most complete occurrence data for families can be obtained by the aggregation of lower-level taxa from the Global Biodiversity Information Facility (GBIF, 2020). However, this database is highly spatially biased (Meyer et al., 2016). Moreover, neither AFE nor GBIF provides co-occurrence data at the level of local plant communities that would allow separate analyses for individual habitat or vegetation types.

The information on the proportion of species of individual plant families in a given community (i.e., the relative species richness per family) may be essential, for example, for the interpretation of the patterns of phylogenetic diversity of plant communities explored over large spatial extents (Padullés Cubino et al., 2021). The patterns of species richness in the geographical space differ among habitat types with different ecology, land use and biogeographical history (Divíšek & Chytrý, 2018; Večeřa et al., 2019) which may translate into the patterns of species richness within and across families. Consequently, we expect the patterns of relative species richness of individual families in the European vegetation to reflect biogeographical processes. In particular, assuming niche conservatism (Wiens & Donoghue, 2004; Wiens et al., 2010), a family may be better represented in habitats or geographical regions characterized by environmental conditions similar to those under which their ancestors evolved and persist.

In this study, we (i) provide the first maps of relative species richness of vascular plant families across Europe, (ii) identify regions with high or low relative species richness of individual families, and (iii) compare the family-richness patterns among four major terrestrial vegetation formations: forests, grasslands, scrub and wetlands (hereafter referred to as 'habitat groups'). We believe these maps and supporting statistics can be useful for a variety of purposes, such as identifying family-specific diversity hotspots, providing information for biodiversity conservation, and inspiring and setting a context for macroecological, biogeographical and biodiversity studies.

Materials and methods

Data preparation and filtering

We obtained 1,159,856 georeferenced vegetation plots from the European Vegetation Archive (EVA; Chytrý et al., 2016, project number 110 – 2020-09-20, data export from 2020-10-22) covering all the major

types of European vegetation. The geographic scope of the dataset is Europe, including Georgia but excluding Svalbard. Some regions, particularly the northern and eastern parts of the continent, have sparser data coverage, especially in some habitat groups (see Appendix S1, Figure S1.1 in the Supplementary material).

Using the EUNIS-ESy expert system v.2020-06-08 (Chytrý et al., 2020), we classified the plots based on their species composition and species cover into habitat types of EUNIS (European Nature Information System, 3rd hierarchical level) and merged them into broad habitat groups of forest, grassland, scrub and wetland vegetation. We excluded 261,841 plots representing vegetation with non-vascular plants only, forestry plantations, habitat types not classified into the above-mentioned four habitat groups, and transitions between habitat types.

We restricted our selection to include only plots ranging in size from 100 to 1000 m^2 for forest vegetation, and from 1 to 100 m^2 for grassland, scrub and wetland vegetation (35,105 plots removed). These ranges reflect the plot sizes traditionally used by the European phytosociology (Westhoff & van der Maarel, 1978). However, we kept plots with unknown sampling size (175,887 plots), since an important part of the geographical coverage (e.g. most of France) would otherwise be lost. These plots came in major part from phytosociological surveys, thus the size of most of them likely falls within the selected range. The relationship between species richness and plot size for individual habitat groups was generally weak (Spearman correlation ~ from -0.12 to 0.15, $p < 0.001$), though it was slightly stronger in wetlands (0.40, $p < 0.001$).

Vegetation plots were assigned to the UTM grid cells of 50 km × 50 km that were subsequently used for mapping. Some vegetation plots in the EVA database are not accurately georeferenced, as indicated by their "location uncertainty". We excluded plots with a location uncertainty higher than the distance from the reported location of the plot to the nearest edge of a 50 km × 50 km grid cell into which the plot belonged (46,905 plots removed). Thus, we kept only plots unequivocally assigned to a unique grid cell. The final dataset comprised 816,005 plots: 230,197 forest, 351,107 grassland, 86,370 scrub and 148,331 wetland plots. An overview of the contributing databases and numbers of plots assigned to particular EUNIS habitat types is available in Appendix S1, Tables S1.1 and S1.2.

Nomenclature

The plant nomenclature was unified according to the Euro+Med PlantBase (2006–2020). For the family concept, we followed World Plants (Hassler, 2020), which extends the concept of Angiosperm Phylogeny Group (APG IV; see Angiosperm Phylogeny Website, 2001–2017). The only nomenclatural difference for European flora is the name of the Adoxaceae family, which can be found under the conserved name Viburnaceae in the APG system, but the same species are included. While the last updates of APG were

published in 2017, the World Plants database is continuously updated and includes more, especially later described species.

All infraspecific taxa were merged at the species level. The taxa determined at the genus or family level were retained, while higher ranks were excluded. We assigned the native or alien status to each species based on their origin within or outside Europe (Euro+Med PlantBase, 2006–2020; POWO, 2017–2020). We identified 293 alien species of extra-European origin, which were excluded from our analyses as we focused on the native European flora. Although several species of European origin might also be considered alien (archaeophytes or neophytes) in some countries outside their native range, these were retained. The list of all taxa recorded in the final dataset with their frequencies in habitat groups is in Appendix S1, Table S1.3.

Data analyses

For each plot, we calculated the relative species richness of each plant family as the number of species belonging to that family divided by the total number of species recorded in the plot. We used the relative species richness, not the absolute species number, to minimize the effect of changes in species richness related to varying plot sizes across the dataset. For mapping, we considered only families recorded in at least ten plots within at least one of the four habitat groups. For each habitat group, we calculated the mean relative species richness of a particular family in a 50 km × 50 km grid as the average relative species richness of that family from all individual plots found in the focal grid cell. Subsequently, we mapped the mean relative species richness of families separately for each habitat group, but only for grid cells containing at least five plots of a particular habitat group. For each map, we derived the colour scale of the mean relative species richness of a family from the data using the *k*-means algorithm. Thus, the scale differs among maps, allowing to explore the relative importance of each family in local communities regardless of its number of species.

The patterns of relative species richness of plant families may depend on the number of species and the distribution of relative frequencies of individual species within the family. For example, the geographic patterns for the families with few species or one or few species much more frequent than the others may reflect idiosyncratic patterns of such species. Therefore, we suggest the geographic patterns for families to be interpreted in conjunction with the Shannon diversity index (Shannon, 1948) which was calculated for each combination of family and habitat group as:

$$H' = -\Sigma\, p_i * \ln p_i$$

where p_i is the relative frequency of species i within a family, i.e. the number of occurrences of species i in a set of vegetation plots (within a habitat group) divided by the number of occurrences of all species belonging to the family in this set of vegetation plots.

Higher values of the index indicate that a given family contains a high number of species that are relatively evenly frequent in the plots of a given habitat group, while lower values indicate that the family contains a small number of species and/or the frequency of species of the family is highly uneven within the habitat group. Consequently, the red areas in the maps of mean relative species richness of families do not always represent plant communities containing more species of the given family than those in blue areas. Instead, they may represent species-poor communities in which the family is represented by very few species. As additional information facilitating the interpretation of mean relative species richness of families, we prepared maps of the mean (absolute) number of families and the mean species:family ratio per grid cell for each habitat group (both averaged from vegetation plots).

To further explore the distribution and biogeographical affinity of plant families in European vegetation, we assigned vegetation plots to biogeographical regions, roughly corresponding to different biomes or latitudinal zones, and compared relative species richness of selected families among them. Based on the map by European Environment Agency (2016), we delimited five biogeographical regions: (i) Mediterranean region (roughly corresponding to Mediterranean biome); (ii) Steppic region (temperate steppe); (iii) Temperate region (temperate broadleaf forest); (iv) Boreal region (taiga); and (v) Arcto-alpine region (tundra). For details on the delimitation of biogeographical regions, see Appendix S4, Figure S4.1. For each habitat group and 11 selected families, which are among the top ten most species-rich families or top ten most frequent families in Europe, we plotted relative species richness for the biogeographical regions using boxplots, and summarized the Shannon diversity index, the number of plots and the number of species.

To consider the uncertainty of the patterns of mean relative species richness of families across Europe, we prepared two additional sets of maps. First, the mean relative species richness of a family in a grid cell may be the result of highly variable numbers of species of the communities therein. For example, this can happen if an extremely species-poor community is found in the same grid cell with an extremely species-rich community. To show areas in which the patterns might be affected more likely by such contrasts, we mapped the standard error of the mean (SEM) of the relative species richness. Second, the patterns of mean relative species richness of families might be affected by varying sampling intensity. To show this, we mapped the number of plots in a grid cell for each habitat group.

The data were processed in R v.3.6.0 (R Core Team, 2019). We used functions from the R packages 'magrittr' (Bache et al., 2020), 'stringr' (Wickham, 2019) and 'tidyr' (Wickham, 2020) to process the vegetation data; 'berryFunctions' (Boessenkool, 2020), 'classInt' (Bivand et al., 2020a), 'raster' (Hijmans et al., 2020), 'rgdal' (Bivand et al., 2020b) and 'spatialEco' (Evans et al., 2020) for the geoprocessing of the data and creating maps; and 'vegan' (Oksanen et al., 2019) to calculate the Shannon diversity index.

Results

The final dataset contained 10,758 species from 158 vascular plant families, with the number of genera per family ranging from 1 to 158 (mean = 8.2, median = 2), and the number of species ranging from 1 to 1,596 (mean = 68.1, median = 7). Appendix S3, Table S3.1 provides the basic statistics for all 158 families (of which 152 were mapped subsequently), summarizing, for example, the number of vegetation plots with an occurrence of a particular family, or the number of genera or species for each combination of family and habitat group.

We found 19 families recorded in more than 20% of the plots (i.e. >180,000 plots). The five most frequent families were Poaceae (occurring in 89.6% of the plots), Asteraceae (64.7%), Rosaceae (63.2%), Cyperaceae (56.1%) and Fabaceae (47.9%) for which the maps of the mean relative species richness are shown in Figure 1. For the basic statistics for these and the further 15 most frequent families, see Table 1. Among the most frequent families, Fabaceae had the highest Shannon diversity index in forests (H' = 4.20), Asteraceae in both grasslands and scrub vegetation (H' = 4.68 and 5.28, respectively) and Cyperaceae in wetlands (H' = 3.69). The five families with the highest number of species were Asteraceae (1,596), Poaceae (815), Fabaceae (767), Caryophyllaceae (619) and Rosaceae (560) (see Appendix S3, Table S3.1 for the number of species within other families). The numbers of species for these and the further 15 most species-rich families for each habitat group are shown in Figure 2.

However, most families were represented by a few genera and species. These families were often rare and with a restricted geographic range within Europe. There were 34 families represented by a single species.

The mean number of families in the 50 km × 50 km UTM grid cells ranged from 1 to 23 in wetlands and from 4 up to 34 in forests. For forests, grasslands and scrub, we found that, on average, more families occurred in Central, Southeastern and partly Eastern Europe, as opposed to Northwestern Europe and coastal areas of Southern Europe (Figure 3a). For wetlands, the mean number of families was distributed more randomly across Europe (Figure 3a). The mean species:family ratio exhibited a more patchy distribution with specific patterns for individual habitat groups. The highest species:family ratios were found (i) within forests in some areas of Southeastern and Northern Europe; (ii) within grasslands in

Southeastern, Eastern and partly Central Europe; (iii) within scrub scattered throughout the continent; and (iv) within wetlands in mountain areas of Central Europe, the British Isles, and in Northern Europe (Figure 3b).

The 522 maps showing the mean relative species richness for each of the four habitat groups were prepared for 152 families that were sufficiently represented in the data (recorded in at least ten plots within at least one of the four habitat groups) (Appendix S2, Figure S2.1). The maps have been made freely available online, both as images and GIS layers (https://doi.org/10.5281/zenodo.4688660).

Spatial patterns of the mean relative species richness of families across Europe highly varied among families and also among habitat groups for a given family. The distribution of Poaceae (high mean relative richness in Northwestern Europe) and Fabaceae (high mean relative richness in Southern/Southeastern Europe) are examples of families that showed a consistent pattern in different habitat groups, except for wetlands, where the distribution was more scattered. In contrast, Asteraceae and Rosaceae are examples of inconsistent distribution patterns of mean relative species richness between habitat groups. In forests and scrub, relative species richness of Asteraceae was spatially structured with higher values in mountain areas of Central and Southeastern Europe, while in grasslands, they were much more evenly distributed, with high values occurring especially in Eastern Europe. Similarly, the high relative richness of Rosaceae in forests was concentrated especially in Western Europe, in wetlands in Northern Europe, while in grasslands and scrub, the areas with high values were scattered throughout the continent. The families represented by few species (e.g. Capparaceae, Coriariaceae, Cytinaceae, Moraceae, Myrtaceae and Styracaceae) usually reached their highest values of mean relative species richness in Southern and Southeastern Europe. Other small families were concentrated elsewhere, for instance, Diapensiaceae in northern Scandinavia, Eriocaulaceae in western Ireland and Scotland, or Mazaceae and Nitrariaceae in the eastern part of European Russia.

There were also medium-size families with high relative species richness largely limited to certain regions or habitat groups, for instance, Ericaceae in Northern and Northwestern Europe, Saxifragaceae in mountains of southern Central Europe and forests of Western Europe or Betulaceae in Northern and partly Northwestern Europe.

The geographic patterns of the relative richness of families across Europe partly reflected contrasts between different biomes. Our boxplots show that some families, most notably Fabaceae and Lamiaceae, were associated primarily with the Mediterranean region and secondarily with the Steppic region (Appendix S4, Figure S4.2). In some but not all habitat groups, the same association was also observed for Apiaceae, Brassicaceae and Caryophyllaceae. Other families, especially Poaceae and Rosaceae, tended to be better represented in the Temperate, Steppic and partly also in the Boreal region. Cyperaceae were also well-

represented in these regions. However, of the two species-rich graminoid families, Poaceae were clearly better represented in the Temperate region, while Cyperaceae were better represented in the Boreal region. The spatial distribution of plant families across biogeographical regions showed consistent patterns between forests, shrublands and grasslands, but not so much for wetlands. For comparison of the Shannon diversity index, the number of plots and the number of species for the selected families among biogeographical regions, see Appendix S4, Table S4.1.

The SEM of relative species richness was low in most families throughout the study area, with median values ranging from 0.01% to 0.81% (Appendix S5, Figure S5.1). This result suggests low uncertainty in the variability of the relative species richness of plots within a grid cell. We also used the number of plots per grid cell as an uncertainty measure. Here, we found that the best-surveyed areas with the highest number of plots and the lowest uncertainty were Denmark, parts of Central Europe and the Netherlands. The median number of plots per grid cell ranged from 21 in scrub to 49 in grasslands (Appendix S5, Figure S5.2).

Discussion

This study is the first to explore the diversity of vascular plant families at the community level across Europe. Previous studies showed either differences in the phylogenetic structure across habitat types (Gerhold et al., 2015; Lososová et al., 2015; 2021) or the distribution maps of a few selected families across all habitat types merged (Malyshev, 1972; Kurtto et al., 2018). Our maps revealed considerable variation in the relative species richness of families across the continent and within the four major habitat groups. These patterns suggest that diversification of individual plant families was influenced by different biogeographical processes and resulted in different numbers of species colonizing European habitats.

Patterns of relative species richness of selected families

Our results indicate that many families show distinct spatial patterns in their mean relative species richness across the continent. There are also clear patterns when the relative species richness of families is combined across broad geographical areas belonging to the same biome, indicating that the macroevolution of the European and Holarctic flora was probably related to major biogeographical processes. This suggests that niche conservatism (Wiens & Donoghue, 2004; Wiens et al., 2010) likely drives the current distribution of plant families in Europe.

We found that the most frequent families in European vegetation are Poaceae, Asteraceae, Rosaceae, Cyperaceae and Fabaceae. These families have broad distributions (GBIF, 2020) and occur in a wide range of habitats.

Poaceae are supposed to have originated in tropical forest edges and diversified owing to their acquired ability to tolerate drought, fire and grazing (Bouchenak-Khelladi et al., 2010; Clayton, 1981). Our maps show that Poaceae exhibit high mean relative species richness (i) in forests, scrub, and grasslands of Northwestern Europe, and (ii) in forests and grasslands of some Mediterranean areas. The first pattern may result from long-term anthropogenic pressure on communities or biogeographic history (Pleistocene glaciations and subsequent recolonization) of this area, both potentially resulting in communities dominated by grasses and containing fewer families compared with, for example, mountain areas of Central Europe (Figure 1 and Figure 3). The second pattern is likely related to the occurrence of annual Poaceae species that are more frequent in the Mediterranean than in temperate areas (Linder & Rudall, 2005).

Species of Asteraceae are dominant and highly represented in many European habitats. While Poaceae are more frequent in vegetation plots than Asteraceae, the latter are more species-rich, which is in agreement with the globally known numbers (12,244 species in 828 genera of Poaceae vs 34,989 species in 1,730 genera of Asteraceae; Hassler, 2020). Similarly to Poaceae and Cyperaceae, Asteraceae also experienced a dramatic diversification and range expansion in the Oligocene to newly arisen open habitats (Ford et al., 2006; Panero & Crozier, 2016), which was further supported by their effective wind dispersal. They tend to prevail in open treeless vegetation, especially in the Steppic region, or in the open-canopy montane forests of the Arcto-alpine region. Asteraceae are highly represented in the native flora, even among endemics.

In contrast to Asteraceae, Rosaceae (5,581 species in 123 genera worldwide; Hassler, 2020) include many species of trees and shrubs in Europe. While this family is widely distributed across Europe, our maps suggest a strong affinity to forest and scrub vegetation, where Rosaceae species have the highest relative species richness. This pattern is in accordance with the assumption that the diversification burst of Rosaceae corresponds with the rise of deciduous forests (Wang et al., 2009) and supports the niche conservatism hypothesis (Wiens et al., 2010).

Cyperaceae can be found in various habitats from dry steppes to wetlands. They often dominate in the latter, having the highest relative species richness and likely an evolutionary centre in biomes rich in wetlands, such as the current boreal taiga biome. Although Cyperaceae is one of the most frequent families in Europe, it has a small number of species, corresponding to the global pattern (5,680 species in 92 genera, but 1,997 species belong to the genus *Carex*; Hassler, 2020). On the other hand, species from Cyperaceae often have large distribution ranges (GBIF, 2020).

Fabaceae, ranked fifth in overall frequency, are globally the third most species-rich family (Christenhusz & Byng, 2016; 22,249 species in 772 genera; Hassler, 2020). This family likely evolved in (semi-)arid regions of the tropics and subtropics. Currently, its species often dominate deciduous (semi-)arid woody vegetation (Lewis et al., 2005; Schrire et al., 2005). Fabaceae are characterized by strong niche conservatism and experienced many intercontinental dispersal events while tracking favourable environmental conditions (Schrire et al., 2005). Their infrequent spread between biomes has been associated with changes in their growth form (Gagnon et al., 2019). These characteristics are well reflected in our results. The mean relative species richness of Fabaceae in Europe is particularly high in forests and scrub of the Mediterranean region (with a higher representation of drought-tolerant shrubby legumes) and low in temperate areas (with predominantly herbaceous species).

Families with the lowest species numbers comprised (i) globally small families containing just few genera (e.g. Diapensiaceae, Staphyleaceae, Taxaceae) or only one genus (e.g. Coriariaceae, Platanaceae, Scheuchzeriaceae), and (ii) medium-size families with the distribution centre outside Europe, usually in the tropics and subtropics, represented in Europe marginally by a small number of genera and species (e.g. Lauraceae, Loranthaceae, Myrtaceae). In this regard, Orchidaceae, the second most diversified family worldwide (30,113 species in 803 genera; Hassler, 2020) with the highest representation in the tropics (cf. POWO, 2017–2020), were represented by only 140 species in our dataset. Most small families had high relative species richness or restricted overall range in Southern and especially Southeastern Europe, with a strong affinity to forest and scrub vegetation. These are, for instance, families with well-known relict species, such as Gesneriaceae (*Haberlea rhodopensis*, *Jancaea heldreichii* and *Ramonda* spp.), Lauraceae (*Laurus nobilis*), Platanaceae (*Platanus orientalis*), Staphyleaceae (*Staphylea pinnata*) or Styracaceae (*Styrax officinalis*). This pattern reflects the shift of thermophilous taxa to refugial areas during the Quaternary climate cooling events and supports the idea that the related extinction was phylogenetically biased (Svenning, 2003; Eiserhardt et al., 2015).

Examples of families with markedly different geographical patterns are Betulaceae, Diapensiaceae and Ericaceae. These families are prevalent in Northwestern and Northern Europe because they are ones of the few families tolerating cold conditions combined with nutrient-poor soils in arctic, boreal and alpine regions (Schwery et al., 2015). Similarly, Saxifragaceae exhibited higher mean relative species richness in the European mountain areas, where this family could have experienced high speciation rates that resulted in a high number of endemics (Loidi et al., 2015).

Interpretation issues and limitations of the data used

Our maps of the mean relative species richness of families were designed to visualize spatial patterns individually for each family. However, these patterns must be interpreted with caution and in combination with the Shannon diversity index and the species:family ratio.

The most intuitive way of reading the maps of the mean relative species richness of families is that (A) the grid cells coloured in red host plant communities containing more species of a particular family than those coloured in blue. This is the prevailing case for the combinations of family and habitat group characterized by relatively high values of the Shannon diversity index, i.e. harbouring many species that are relatively evenly distributed across vegetation plots. This also applies to the interpretation of the mean relative species richness within a family in areas with a high number of families and, at the same time, a high species:family ratio (Figure 3).

Another case might be that (B) the number of species within a family is rather constant across regions but the total species richness across all families varies substantially. This may cause that the highest values (visualized in red in the maps) are based on many species-poor plots in which the family is represented by a few or even one species. Such species are/is, however, of higher importance in these species-poor than in species-rich communities. This may be the prevailing case for the combinations of family and habitat group characterized by lower values of the Shannon diversity index. This also applies to the interpretation of the mean relative species richness within a family in areas containing a small number of families and families that are species-poor (Figure 3).

Both the above-mentioned cases (A and B) may appear within a single map. For instance, a high mean relative species richness of Poaceae in European grasslands can be found in two different contexts. In the Mediterranean, Poaceae are represented by many annual species (case A), while in Northwestern Europe, Poaceae do not have very high absolute species richness, but occur in relatively species-poor communities with low representation of species from other families (especially forbs).

Vegetation data stored in the EVA database were originally collected for various purposes, but mainly for phytosociological surveys. This implies some limitations that may have affected the spatial patterns of mean relative species richness of vascular plant families. The sampling density and representativeness of the regional vegetation diversity vary across regions. While in some regions of Central, Western and Southern Europe, the available data are more representative, datasets from other regions have considerable gaps. In some regions, especially in extensive parts of European Russia, there are no vegetation plots available at all. In other regions (e.g. Norway and Sweden), there are data for some habitat types but not for others. To alleviate this variation in data quality, we considered the patterns at the level of broadly defined habitat groups and averaged them in 50 km × 50 km grid cells. Thus, the observed patterns are less prone to

fine-scale differences in sampling density among local communities. Further, we mapped the number of plots in grid cells (Appendix S5, Figure S5.2) to show the areas with higher or lower sampling density.

Although we focused on (semi-)natural vegetation and excluded species of extra-European origin from the dataset, the alien species of European origin (i.e. species native to some parts of Europe and considered aliens in other parts) may confound the patterns of relative species richness of species-poor families. Examples include Juglandaceae (represented in the dataset by a single species, *Juglans regia*), Hydrangeaceae (*Philadelphus coronarius*) and Moraceae (*Ficus carica*). These species are native to Southeastern Europe, but the overall patterns in our maps are affected by human-assisted introductions, especially escapes from cultivation.

Future outlook and potential applications

Our maps of the mean relative species richness of vascular plant families across Europe can inspire macroecological and biodiversity studies and strengthen the integration of taxonomic, phylogenetic and functional approaches in biodiversity research. They can also help develop new hypotheses about the mechanisms behind biodiversity patterns at different spatial scales and inform biodiversity conservation planning. For example, although several studies assessed the ecological and evolutionary drivers of species richness in different habitats across Europe (e.g., Araújo et al., 2005; Divíšek & Chytrý, 2018; Jiménez-Alfaro et al., 2018; Večeřa et al., 2019), it is unclear whether the effect of these drivers holds across higher taxa (i.e. those located deeper in the phylogeny). Because the species within each family share a large fraction of their evolutionary history, they are also more likely to have inherited similar traits and occupy similar distribution ranges as their ancestor taxa. Examining the variation in the effect of diversity drivers across different levels of taxonomy could provide new insights into the mechanisms determining the assembly of natural communities (Laiolo et al., 2020). Furthermore, it could also point out the processes that lead to phylogenetic niche conservatism or niche differentiation in different regions in Europe.

Further analyses could clarify whether the number of species in each family is lower or higher in specific habitat types or regions than under random expectation. These results would allow the identification of regions where particular families are over- or under-represented and inform conservation strategies at the continental scale. This goal is important because entire families are sometimes considered as relevant units for biological conservation (e.g. Orchidaceae; CITES, 2019). Defining regional conservation targets for specific families (e.g. rare ones, those with a distribution limit in Europe or endangered by land-use changes) may be also useful to preserve certain lineages and their evolutionary contexts that may be otherwise neglected by a conservation perspective focused solely on maintaining species richness or preserving particular endemic species.

Conclusions

Our maps of the relative species richness of vascular plant families in four vegetation formations provide novel information on the structure of European plant diversity. They demonstrate that the spatial structure of European plant diversity includes an important phylogenetic component, which could reflect the evolutionary and biogeographical history of the continent. However, it can also reflect recent trends associated with land-use changes and other processes of global change. Further, our study shows that the relative species richness of individual families within communities varies among habitat groups across Europe, indicating their specific evolutionary history. We believe these maps can support macroecological and biogeographical research, which, in turn, would contribute to a deeper understanding of these patterns.

Acknowledgements

We thank Jiří Danihelka for his help with unifying taxonomic nomenclature in the vegetation data; Stephan Hennekens for managing the European Vegetation Archive; Sylvain Abdulhak, Alicia Acosta, Emiliano Agrillo, Iva Apostolova, Erwin Bergmeier, Jan-Bernard Bouzillé, Henry Brisse, János Csiky, Mirjana Ćuk, Els De Bie, Gianpietro Giusso del Galdo, Olga Demina, Iris de Ronde, Michele De Sanctis, Panayotis Dimopoulos, Tetiana Dziuba, Úna FitzPatrick, Xavier Font, Rosario G. Gavilán, Valentin Golub, Friedemann Goral, Adrian Indreica, Maike Isermann, Jan Jansen, John Janssen, Veronika Kalníková, Ali Kavgacı, Larisa Khanina, Filip Küzmič, Flavia Landucci, Igor Lavrinenko, Armin Macanović, Corrado Marcenò, Aleksander Marinšek, Vladimir Onipchenko, Viktor Onyshchenko, Aaron Pérez-Haase, Tomáš Peterka, Vadim Prokhorov, Valerijus Rašomavičius, Maria Pilar Rodríguez-Rojo, John S. Rodwell, Solvita Rūsiņa, Joachim Schrautzer, Pavel Shirokikh, Zvjezdana Stančić, Angela Stanisci, Ioannis Tsiripidis, Emin Uğurlu, Roberto Venanzoni, Denys Vynokurov, Wolfgang Willner, French National Forest Inventory and Swedish National Forest Inventory for data from vegetation-plot databases they manage; and last but not least, all the people who collected the vegetation data used in this study.

Author contributions

MV, IA, JPC, ZL, JDi and MC conceived the ideas; SA, IB, SB, GB, JAC, AČ, LC, RĆ, JDe, FF-G, J-CG, FJ, AJ, BJA, AK, JL, JEM, RP, ER, ŽŠ, KV, TW and MC contributed the vegetation-plot data and helped with developing a pilot version of this study; MLC, AC, PD, UJ, ML, TL, JŠ, UŠ, GS, IT and SY contributed the vegetation-plot data; MV, IA and IK prepared and processed the data; MV, IA and JDi prepared the maps and tabular outputs; MV and IA wrote the manuscript with significant contributions of JPC, ZL, JDi and MC. All authors commented on the manuscript drafts. MV and IA contributed equally to this study.

Data availability statement

The vegetation-plot data used in this study are stored in the European Vegetation Archive (EVA; http://euroveg.org/eva-database) under project number 110 – 2020-09-20. Maps of the mean relative species richness of vascular plant families in European vegetation can be downloaded from Zenodo repository (https://doi.org/10.5281/zenodo.4688660).

References

Angiosperm Phylogeny Website (2001–2017) *Angiosperm Phylogeny Website, version 14*. Available at http://www.mobot.org/ MOBOT/research/APweb [accessed 22 October 2020]

Araújo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: Implications for conservation planning. *Global Ecology and Biogeography*, 14, 17–30. https://doi.org/10.1111/j.1466-822X.2004.00128.x

Bache, S.M., Wickham, H. & Henry, L. (2020) *magrittr: A Forward-Pipe Operator for R. Version 2.0.1*. Available at https://cran.r-project.org/web/packages/magrittr/index.html

Beck, J., Ballesteros-Mejia, L., Buchmann, C.M., Dengler, J., Fritz, S.A., Gruber, B. *et al.* (2012) What's on the horizon for macroecology? *Ecography*, 35, 673–683. https://doi.org/10.1111/j.1600-0587.2012.07364.x

Bivand, R., Ono, H., Dunlap, R., Stigler, M., Denney, B. & Hernangómez, D. (2020a) *classInt: Choose Univariate Class Intervals*. *Version 0.4-3*. Available at https://cran.r-project.org/web/packages/classInt/index.html

Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R.J., *et al.* (2020b) *rgdal: Bindings for the 'Geospatial' Data Abstraction Library. Version 1.5-12*. Available at https://cran.r-project.org/web/packages/rJava/index.html

Boessenkool, B. (2020) *berryFunctions: Function Collection Related to Plotting and Hydrology. Version 1.19.1*. Available at https://cran.r-project.org/web/packages/berryFunctions/index.html

Bouchenak-Khelladi, Y., Verboom, G.A., Savolainen, V. & Hodkinson, T.R. (2010) Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Botanical Journal of the Linnean Society*, 162, 543–557. https://doi.org/10.1111/j.1095-8339.2010.01041.x

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M. *et al.* (2018) Global trait–environment relationships of plant communities. *Nature Ecology & Evolution*, 2, 1906–1917. https://*doi*.org/10.1038/s41559-018-0699-8

Bruelheide, H., Dengler, J., Jiménez-Alfaro, B., Purschke, O., Hennekens, S.M., Chytrý, M. *et al.* (2019) sPlot – A new tool for global vegetation analyses. *Journal of Vegetation Science*, 30, 161–186. https://doi.org/10.1111/jvs.12710

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12, 693–715. https://doi.org/10.1111/j.1461-0248.2009.01314.x

Chacón, J., Luebert, F. & Weigend, M. (2017) Biogeographic events are not correlated with diaspore dispersal modes in Boraginaceae. *Frontiers in Ecology and Evolution*, 5, 26. https://doi.org/10.3389/fevo.2017.00026

Chen, L.-Y., Zhao, S.-Y., Mao, K.-S., Les, D.H., Wanga, Q.-F. & Moody, M.L. (2014) Historical biogeography of Haloragaceae: An out-of-Australia hypothesis with multiple intercontinental dispersals. *Molecular Phylogenetics and Evolution*, 78, 87–95. https://doi.org/10.1016/j.ympev.2014.04.030

Christenhusz, M.J.M & Byng, J.W. (2016) The number of known plant species in the world and its annual increase. *Phytotaxa*, 261, 201–217. http://dx.doi.org/10.11646/phytotaxa.261.3.1

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F. *et al.* (2016) European Vegetation Archive (EVA): An integrated database of European vegetation plots. *Applied Vegetation Science*, 19, 173–180. https://doi.org/10.1111/avsc.12191

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M., Rodwell, J. *et al.* (2020) EUNIS Habitat Classification: expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, 23, 648–675. https://doi.org/10.1111/avsc.12519

CITES (2019) *CITES Appendix II*. Available at https://cites.org/sites/default/files/eng/app/2019/E-Appendices-2019-11-26.pdf [accessed 10 December 2020]

Clayton, W.D. (1981) Evolution and distribution of grasses. *Annals of the Missouri Botanical Garden*, 68, 5–14. https://doi.org/10.2307/2398808

Cronquist, A. (1968) *The Evolution and Classification of Flowering Plants*. Boston, MA, US: Houghton Mifflin.

Dengler, J., Wagner, V., Dembicz, I., García-Mijangos, I., Naqinezhad, A., Boch, S. *et al.* (2018) GrassPlot – a database of multi-scale plant diversity in Palaearctic grasslands. *Phytocoenologia*, 48, 331–347. https://doi.org/10.1127/phyto/2018/0267

Divíšek, J. & Chytrý, M. (2018) High-resolution and large-extent mapping of plant species richness using vegetation-plot databases. *Ecological Indicators*, 89, 840–851. https://doi.org/10.1016/j.ecolind.2017.11.005

Donoghue, M.J. (2008) A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 11549–11555. https://doi.org/10.1073/pnas.0801962105

Dornelas, M., Antão, L.H., Moyes, F., Bates, A.E., Magurran, A.E., Adam, D. *et al.* (2018) BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography*, 27, 760–786. https://doi.org/10.1111/geb.12729

Eiserhardt, W.L., Borchsenius, F., Plum, C.M., Ordonez, A. & Svenning, J.-C. (2015) Climate-driven extinctions shape the phylogenetic structure of temperate tree floras. *Ecology Letters*, 18, 263–272. https://doi.org/10.1111/ele.12409

European Commission (2019) *European Green Deal*. Available at https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en [accessed 8 December 2020]

European Environment Agency (2016) *Biogeographical regions*. Available at https://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3 [accessed 23 February 2021]

Euro+Med (2006–2020) *Euro+Med PlantBase – the information resource for Euro-Mediterranean plant diversity*. Available at http://ww2.bgbm.org/EuroPlusMed/ [accessed 09 May 2020]

Evans, J.S., Murphy, M.A. & Ram, K. (2020) *spatialEco: Spatial Analysis and Modelling Utilities.Version 1.3-2*. Available at https://cran.r-project.org/web/packages/spatialEco/index.html

Ford, B.A., Iranpour, M., Naczi, R.F.C., Starr, J.R. & Jerome, C.A. (2006) Phylogeny of *Carex* subg. *Vignea* (Cyperaceae) based on non-coding nrDNA sequence data. *Systematic Botany*, 31, 70–82. https://doi.org/10.1600/036364406775971813

Gagnon, E., Ringelberg, J.J., Bruneau, A., Lewis, G.P. & Hughes, C.E. (2019) Global Succulent Biome conservatism across the pantropical Caesalpinia group (Leguminosae). *New Phytologist*, 222, 1994–2008. https://doi.org/10.1111/nph.15633

GBIF (2020) *GBIF Home Page*. Available at https://www.gbif.org [accessed 05 May 2020]

Gerhold, P., Cahill, J.F. Jr, Winter, M., Bartish, I.V. & Prinzing, A. (2015) Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, 29, 600–614. https://doi.org/10.1111/1365-2435.12425

Hassler, M. (2020) *World Plants: Synonymic Checklist and Distribution of the World Flora. Version 11.0.* Available at https://www.worldplants.de [accessed 07 December 2020]

Hijmans, R.J., van Etten, J., Sumner, M., Cheng, J., Baston, D., Bevan, A. *et al.* (2020) *raster: Geographic Data Analysis and Modeling. Version 3.3-13*. Available at https://cran.r-project.org/web/packages/raster/index.html

IPBES (2019) *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Bonn, Germany: IPBES secretariat.

Jalas, J. & Suominen, J. [Eds.] (1972–1994) *Atlas Florae Europaeae. Volume 1* (1972), *Vol. 2* (1973), *Vol. 3* (1976), *Vol. 4* (1979), *Vol. 5* (1980), *Vol. 6* (1983), *Vol. 7* (1986), *Vol. 8* (1989), *Vol. 9* (1991), *Vol. 10* (1994). Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Jalas, J., Suominen, J. & Lampinen, R. [Eds.] (1996) *Atlas Florae Europaeae*. *Volume 11*. Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Jalas, J., Suominen, J., Lampinen, R. & Kurtto, A. [Eds.] (1999) *Atlas Florae Europaeae*. *Volume 12*. Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Jin, J.-J., Yang, M.-Q., Fritsch, P.W., van Velzen, R., Li, D.-Z. & Yi, T.-S. (2020) Born migrators: Historical biogeography of the cosmopolitan family Cannabaceae. *Journal of Systematics and Evolution*, 58, 461–473. https://doi.org/10.1111/jse.12552

Jiménez-Alfaro, B., Girardello, M., Chytrý, M., Svenning, J.-C., Willner, W., Gégout, J.-C. *et al.* (2018) History and environment shape species pools and community diversity in European beech forests. *Nature Ecology and Evolution*, 2, 483–490. https://doi.org/10.1038/s41559-017-0462-6

König, C., Weigelt, P., Schrader, J., Taylor, A., Kattge, J. & Kreft, H. (2019) Biodiversity data integration—the significance of data resolution and domain. *PLoS Biology* 17, e3000183. https://doi.org/10.1371/journal.pbio.3000183

Kurtto, A., Fröhner, S.E. & Lampinen, R. [Eds.] (2007) *Atlas Florae Europaeae*. *Volume 14*. Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Kurtto, A., Lampinen, R. & Junikka, L. [Eds.] (2004) *Atlas Florae Europaeae*. *Volume 13*. Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Kurtto, A., Sennikov, A.N. & Lampinen, R. [Eds.] (2013, 2018) *Atlas Florae Europaeae*. *Volume 16* (2013), *Vol. 17* (2018). Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Kurtto, A., Weber, H.E., Lampinen, R. & Sennikov, A.N. [Eds.] (2010) *Atlas Florae Europaeae*. *Volume 15*. Helsinki, FI: The committee for mapping the flora of Europe & Societas biologica Fennica Vanamo.

Laiolo, P., Pato, J., Jiménez-Alfaro, B. & Obeso, J.R. (2020) Evolutionary conservation of within-family biodiversity patterns. *Nature Communications*, 11, 882. https://doi.org/10.1038/s41467-020-14720-3

Lewis, G., Schrire, B., Mackinder, B. & Lock, M. [Eds.] (2005) *Legumes of the world*. Kew, UK: Royal Botanic Gardens.

Linder, H.P. & Rudall, P.J. (2005) Evolutionary history of Poales. *Annual Review of Ecology, Evolution, and Systematics*, 36, 107–124. https://doi.org/10.1146/annurev.ecolsys.36.102403.135635

Loidi, J., Campos, J.A., Herrera, M., Biurrun, I., García-Mijangos, I. & García-Baquero, G. (2015) Eco-geographical factors affecting richness and phylogenetic diversity patterns of high-mountain flora in the Iberian Peninsula. *Alpine Botany,* 125, 137–146.

Lososová, Z., Šmarda, P., Chytrý, M., Purschke, O., Pyšek, P., Sádlo, J. *et al.* (2015) Phylogenetic structure of plant species pools reflects habitat age on the geological time scale. *Journal of Vegetation Science*, 26, 1080–1089. https://doi.org/10.1111/jvs.12308

Lososová, Z., Divíšek, J., Chytrý, M., Götzenberger, L., Těšitel, J. & Mucina, L. (2021) Macroevolutionary patterns in European vegetation. *Journal of Vegetation Science*, 32, e12942. https://doi.org/10.1111/jvs.12942

Malyshev, L.I. (1972) *Floristic spectra of the Soviet Union*. In: Chrshanovsky, V.G., Lavrenko, E.M., Linczevsky, L.A., Nikitin, V.V., Vassilczenko, I.T. & Zakirov, K.Z. (Eds.), *The history of the flora and vegetation of Eurasia* [in Russian]. St. Petersburg, RU: Nauka, pp. 17–40.

Manning, P., Gossner, M.M., Bossdorf, O., Allan, E., Zhang, Y.-Y., Prati, D. *et al.* (2015) Grassland management intensification weakens the associations among the diversities of multiple plant and animal taxa. *Ecology*, 96, 1492–1501. https://doi.org/10.1890/14-1307.1

Meyer, C., Weigelt, P. & Kreft, H. (2016) Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters*, 19, 992–1006. https://doi.org/10.1111/ele.12624

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D. *et al.* (2019) *vegan: Community Ecology Package*. *Version 2.5-6*. Available at https://cran.r-project.org/web/packages/vegan/index.html

Padullés Cubino, J., Lososová, Z., Bonari, G., Agrillo, E., Attorre, F., Bergmeier, E. *et al*. (2021) Phylogenetic structure of European forest vegetation. *Journal of Biogeography*. https://doi.org/10.1111/JBI.14046

Panero, J.L. & Crozier, B.S. (2016) Macroevolutionary dynamics in the early diversification of Asteraceae. *Molecular Phylogenetics and Evolution*, 99, 116–132. https://doi.org/10.1016/j.ympev.2016.03.007

Pausas, J.G. & Lamont, B.B. (2018) Ecology and biogeography in 3D: The case of the Australian Proteaceae. *Journal of Biogeography*, 45, 1469–1477. https://doi.org/10.1111/jbi.13348

Power, S.C., Verboom, G.A., Bond, W.J. & Cramer, M.D. (2017) Environmental correlates of biome-level floristic turnover in South Africa. *Journal of Biogeography*, 44, 1745–1757. https://doi.org/10.1111/jbi.12971

POWO (2017–2020) *Plants of the World Online*. Facilitated by the Royal Botanic Gardens, Kew. Available at http://www.plantsoftheworldonline.org/ [accessed 14 December 2020]

Rabosky, D.L., Grundler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W. *et al.* (2014). BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5, 701–707. https://doi.org/10.1111%2F2041-210X.12199

Qian, H., Jin, Y. & Ricklefs, R.E. (2017) Phylogenetic diversity anomaly in angiosperms between eastern Asia and eastern North America. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 11452–11457. https://doi.org/10.1146/annurev.ecolsys.30.1.421

Qian, H. & Zhang, J. (2014) Using an updated time-calibrated family-level phylogeny of seed plants to test for non-random patterns of life forms across the phylogeny. *Journal of Systematics and Evolution*, 52, 423–430. https://doi.org/10.1111/jse.12027

R Core Team (2019) *R: A language and environment for statistical computing*. Vienna, AT: R Foundation for Statistical Computing.

Sainge, M.N., Chuyong, G.B. & Peterson, A.T. (2017) Endemism and geographic distribution of African Thismiaceae. *Plant Ecology and Evolution*, 150, 304–312. https://doi.org/10.5091/plecevo.2017.1196

Schrire, B.D., Lavin, M. & Lewis, G.P. (2005) Global distribution patterns of the Leguminosae: Insights from recent phylogenies. *Biologiske Skrifter*, 55, 375–422.

Schwery, O., Onstein, R.E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R.J. & Linder, H.P. (2015) As old as the mountains: the radiations of the Ericaceae. *New Phytologist*, 207, 355–367. https://doi.org/10.1111/nph.13234

Shannon, C.E. (1948) A mathematical theory of communications. *Bell Systems Technological Journal*, 27, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

Svenning, J.-C. (2003) Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters*, 6, 646–653. https://doi.org/10.1046/j.1461-0248.2003.00477.x

Večeřa, M., Divíšek, J., Lenoir, J., Jiménez-Alfaro, B., Biurrun, I., Knollová, I. *et al.* (2019) Alpha diversity of vascular plants in European forests. *Journal of Biogeography*, 46, 1919–1935. https://doi.org/10.1111/jbi.13624

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R. *et al.* (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 3853–3858. https://doi.org/10.1073/pnas.0813376106

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33, 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Webb, C.O., Ackerly, D.D. & Kembel, S.W. (2008) Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, 24, 2098–2100. https://doi.org/10.1093/bioinformatics/btn358

Westhoff, V. & van der Maarel, E. (1978) *The Braun-Blanquet approach*. In: Whittaker, R.H. (Ed.), *Classification of plant communities,* 2nd edition. The Hague, NL: Junk, pp. 619–729.

Wickham, H. (2019) *stringr: Simple, Consistent Wrappers for Common String Operations*. *Version 1.4.0*. Available at https://cran.r-project.org/web/packages/stringr/index.html

Wickham, H. (2020) *tidyr: Tidy Messy Data. Version 1.1.2*. Available at https://cran.r-project.org/web/packages/tidyr/index.html

Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, 19, 639–644. https://doi.org/10.1016/j.tree.2004.09.011

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V. *et al.* (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, 13, 1310–1324. https://doi.org/10.1111/j.1461-0248.2010.01515.x

List of appendices

Appendix S1 Supplementary materials and methods

Appendix S2 Maps of the mean relative species richness of vascular plant families

Appendix S3 Summary statistics

Appendix S4 Comparison of the relative richness of families among biogeographical regions

Appendix S5 Uncertainty maps

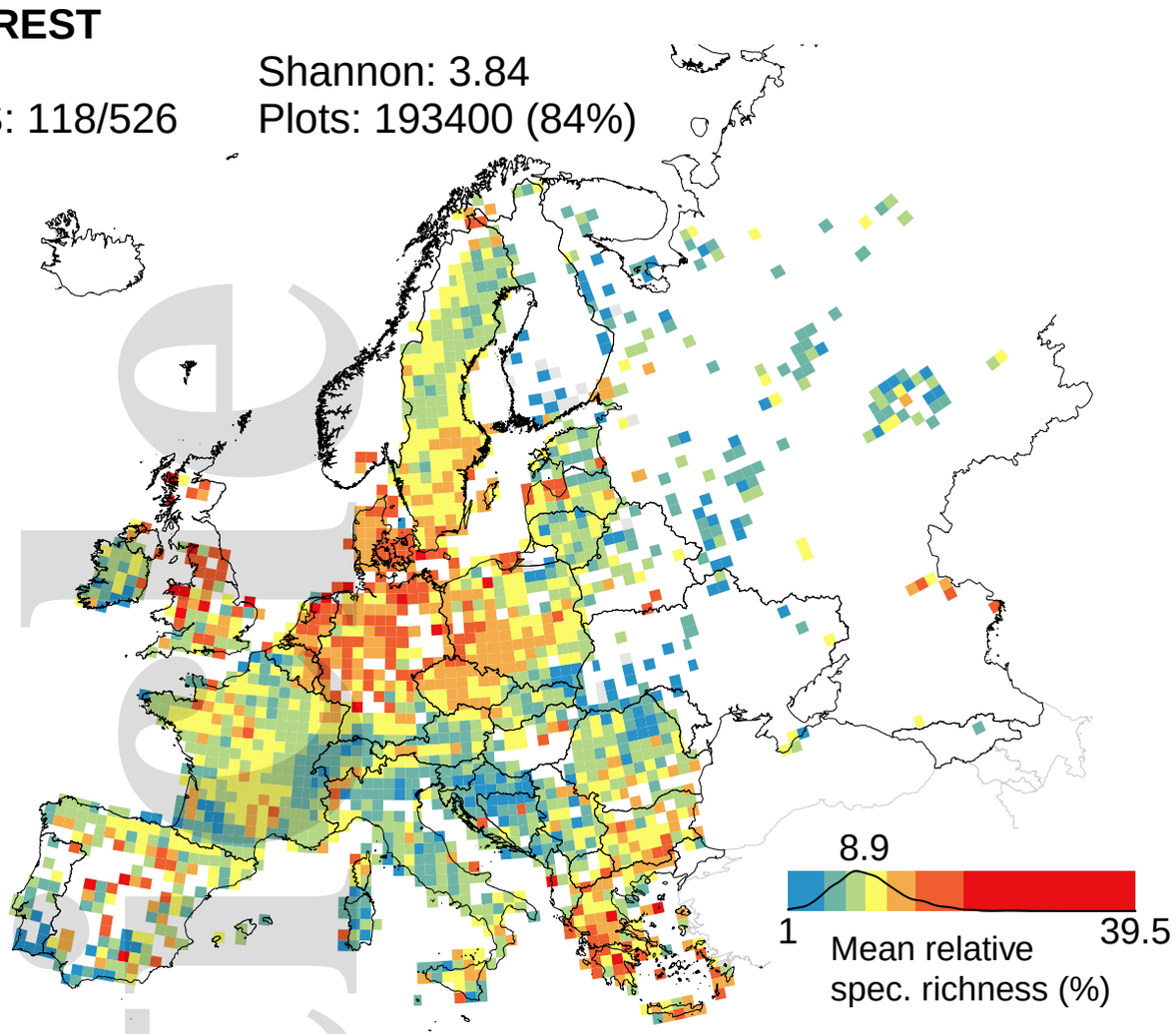
Table 1 Summary statistics for the twenty most frequent plant families. Families are sorted by their relative frequency in the whole dataset (N = 816,005). Shannon diversity index describes the taxonomic diversity of the family and its representation in European vegetation (see *Materials and methods* for details). Relative mean species number was calculated across plots and multiplied by 100; the values were weighted relative to all taxa recorded in a plot. See Appendix S3, Table S3.1 for more details and information for all families. *For.* – forests, *Grass.* – grasslands, *Wet.* – wetlands.

Family	All plots where recorded	Relative frequency (%) in the whole dataset	Shannon index				Relative mean species number per plot *100			
			For.	*Grass.*	*Scrub*	*Wet.*	*For.*	*Grass.*	*Scrub*	*Wet.*
Poaceae	731010	89.6	3.84	4.23	4.38	3.31	997.96	2445.21	1630.30	1500.78
Asteraceae	527859	64.7	4.16	4.68	5.28	3.03	465.39	1275.81	659.50	333.45
Rosaceae	515732	63.2	3.55	3.69	3.86	2.28	908.09	492.47	688.42	391.06
Cyperaceae	457814	56.1	3.38	3.80	3.49	3.69	339.46	429.77	641.81	1924.23
Fabaceae	390574	47.9	4.20	3.79	4.81	2.12	212.13	763.13	461.08	122.41
Rubiaceae	372914	45.7	2.74	2.87	3.04	1.35	239.56	270.94	264.38	255.79
Lamiaceae	371513	45.5	3.50	3.86	4.36	2.06	365.31	351.42	302.21	322.43
Ranunculaceae	350264	42.9	3.27	2.71	3.87	2.09	258.75	354.50	112.50	258.17
Apiaceae	319006	39.1	3.57	3.72	4.40	2.45	152.50	332.21	160.20	270.91
Caryophyllaceae	314997	38.6	2.87	3.74	4.55	2.04	156.50	409.90	162.35	99.27
Plantaginaceae	286853	35.2	2.63	2.64	3.58	2.62	95.69	371.33	111.32	91.72
Juncaceae	278906	34.2	2.06	2.51	2.88	2.21	199.54	252.66	187.20	363.04
Polygonaceae	248725	30.5	2.68	2.20	2.37	2.33	37.84	312.89	122.71	283.37
Fagaceae	212393	26.0	1.82	1.04	1.92	0.67	658.57	23.49	152.42	20.81
Caprifoliaceae	211014	25.9	2.71	2.71	3.32	1.47	171.66	115.22	109.23	68.71
Betulaceae	188647	23.1	1.88	1.77	1.84	1.27	432.88	19.95	174.27	166.98
Ericaceae	184823	22.6	2.48	2.20	2.61	2.18	434.19	72.00	1430.91	585.12
Brassicaceae	183395	22.5	3.24	4.01	4.85	1.94	102.05	172.93	72.99	129.73
Primulaceae	180137	22.1	2.54	2.67	3.10	1.56	122.08	100.40	100.27	190.81
Violaceae	143907	17.6	1.88	2.53	2.77	0.35	132.01	51.12	58.14	50.35

Poaceae

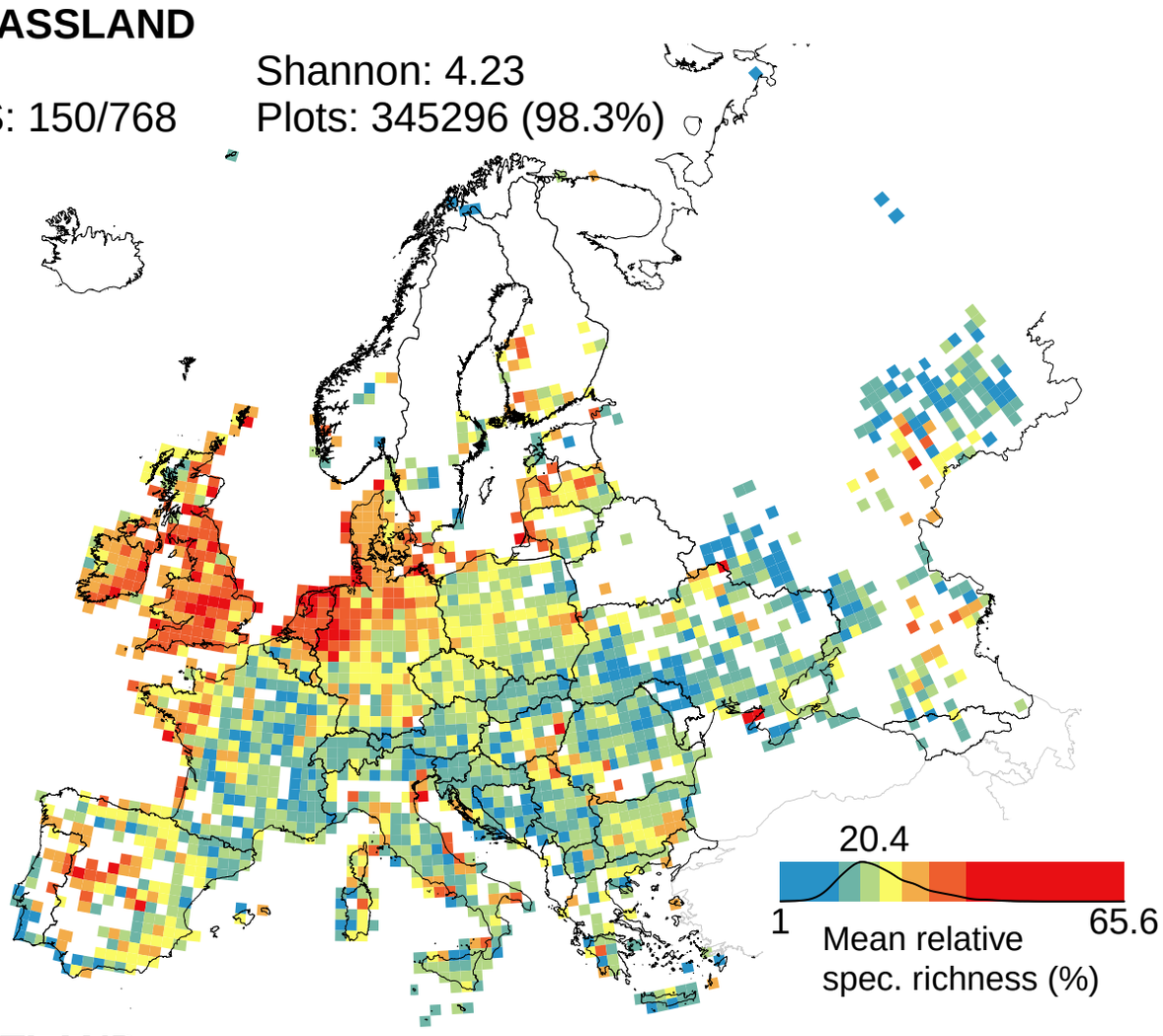
FOREST

G/S: 118/526
Shannon: 3.84
Plots: 193400 (84%)



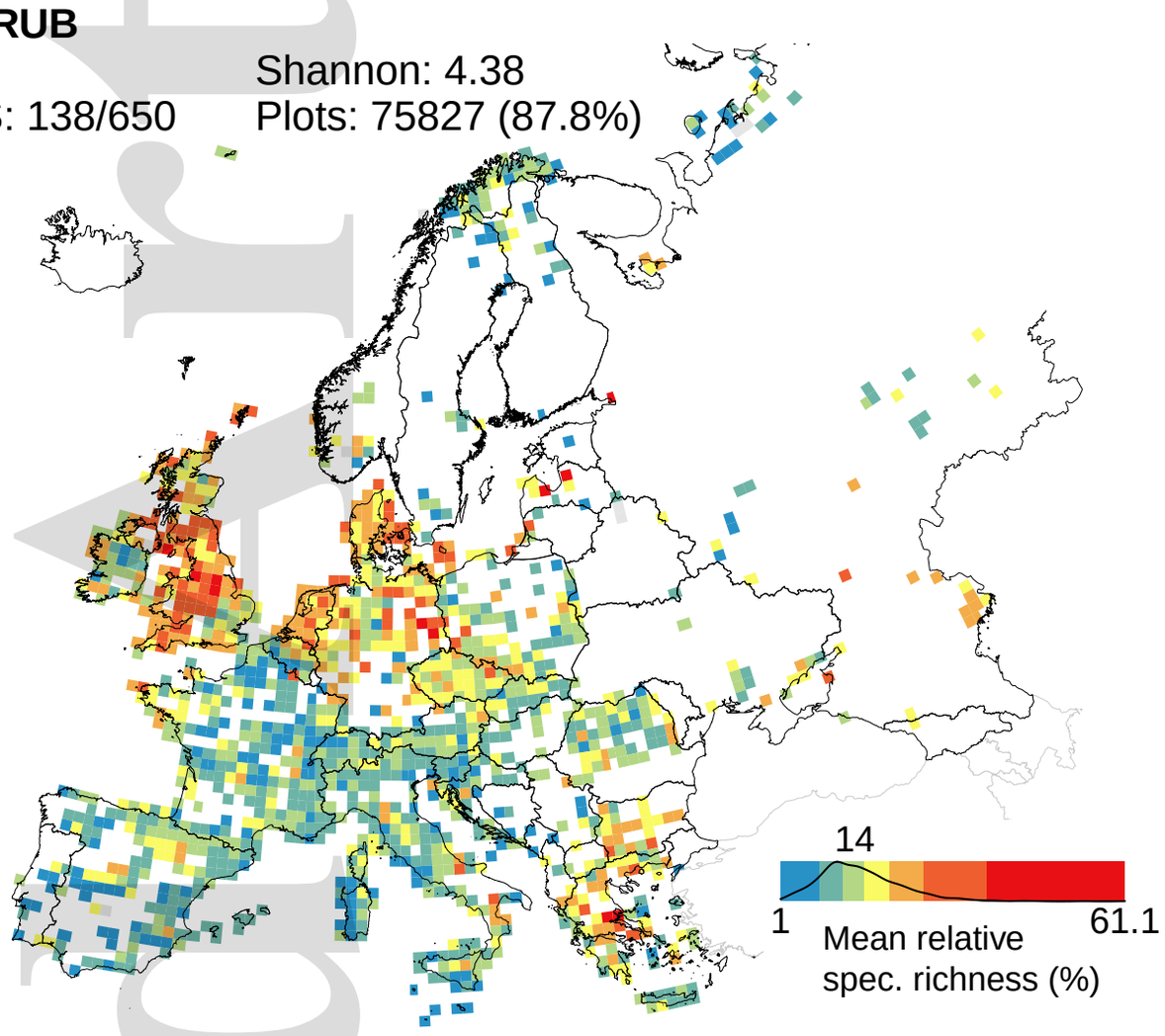
GRASSLAND

G/S: 150/768
Shannon: 4.23
Plots: 345296 (98.3%)



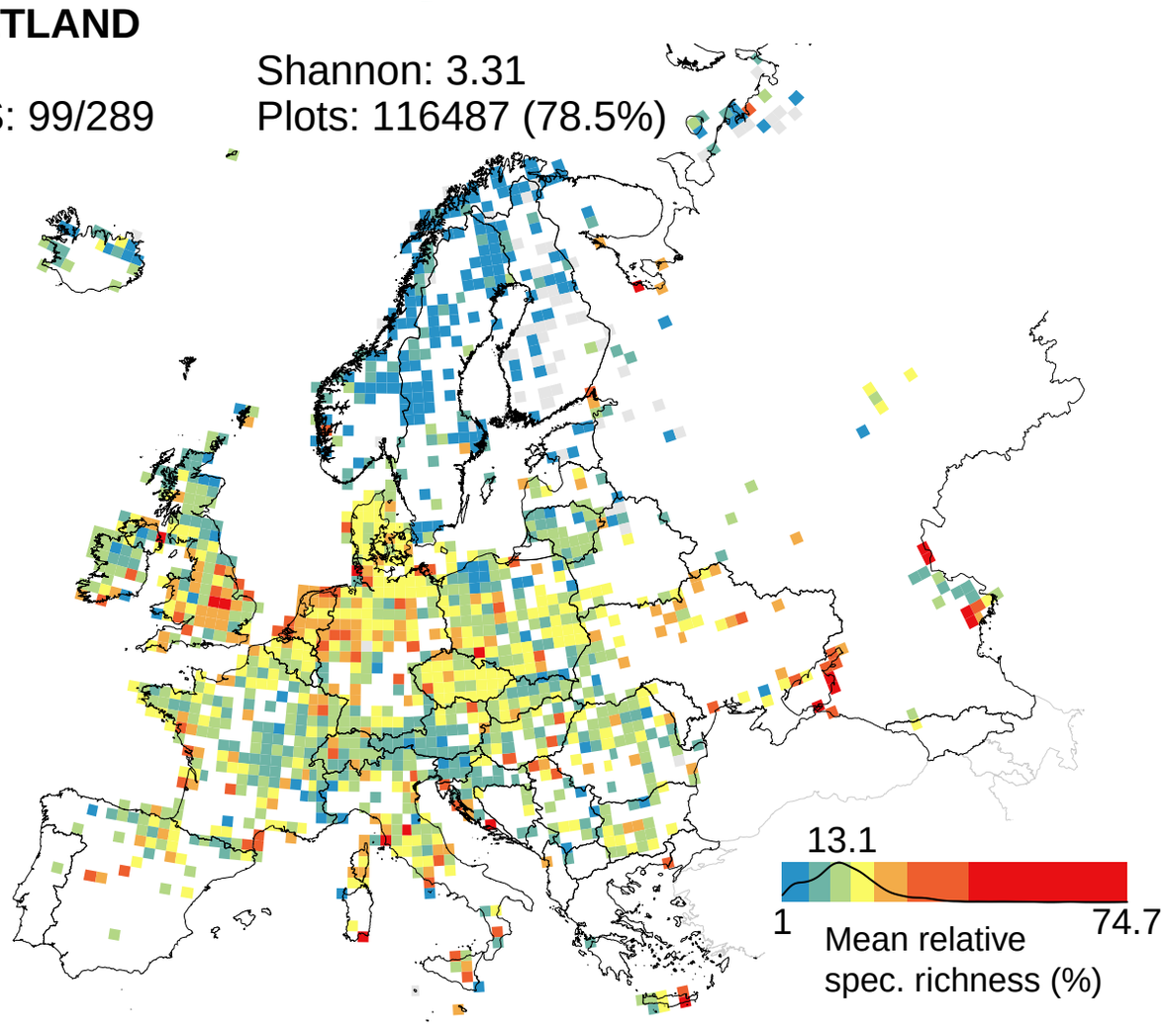
SCRUB

G/S: 138/650
Shannon: 4.38
Plots: 75827 (87.8%)



WETLAND

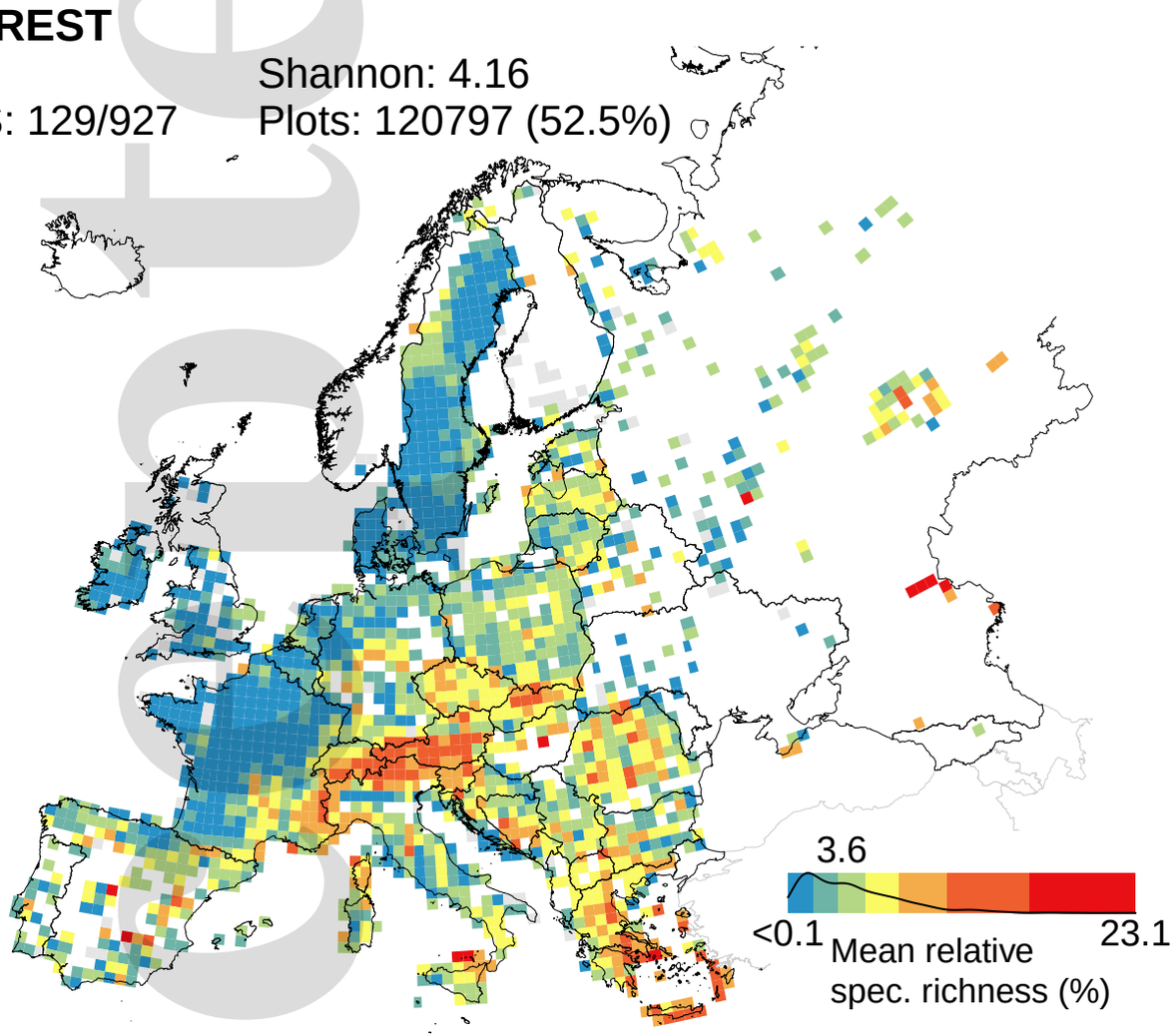
G/S: 99/289
Shannon: 3.31
Plots: 116487 (78.5%)



Asteraceae

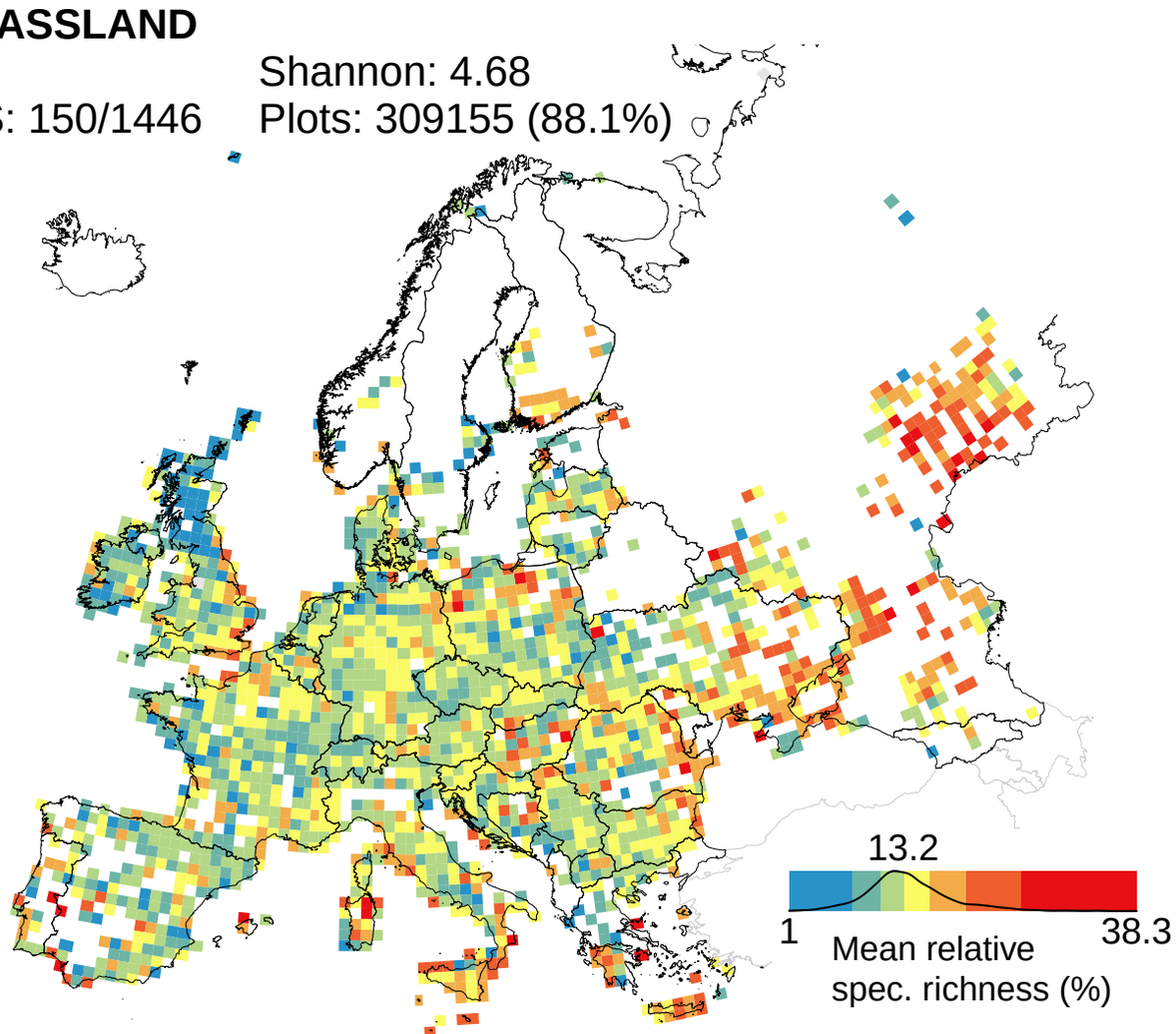
FOREST

G/S: 129/927
Shannon: 4.16
Plots: 120797 (52.5%)



GRASSLAND

G/S: 150/1446
Shannon: 4.68
Plots: 309155 (88.1%)



SCRUB

G/S: 138/1053
Shannon: 5.28
Plots: 44916 (52%)

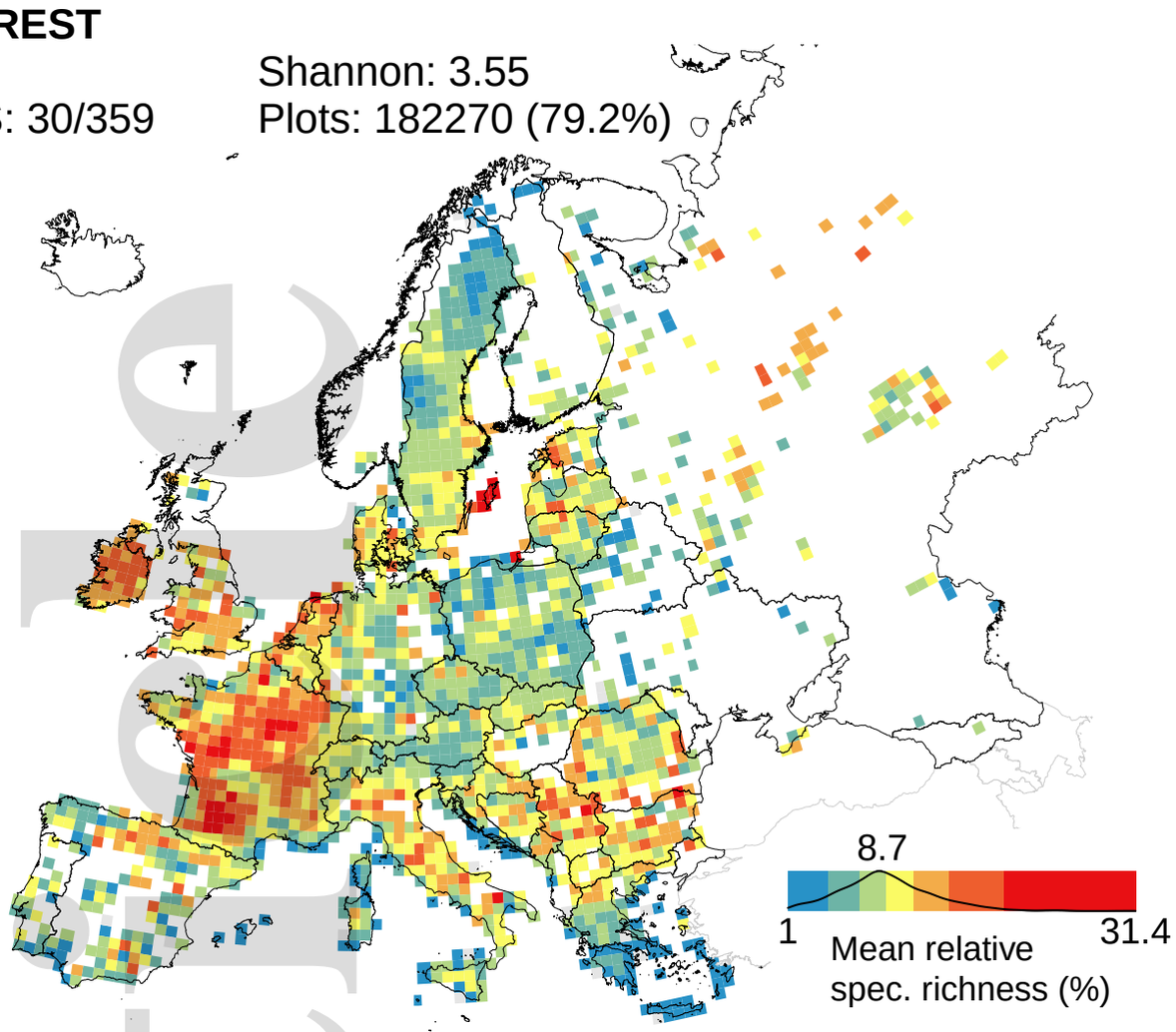
WETLAND

G/S: 91/348
Shannon: 3.03
Plots: 52991 (35.7%)

Rosaceae

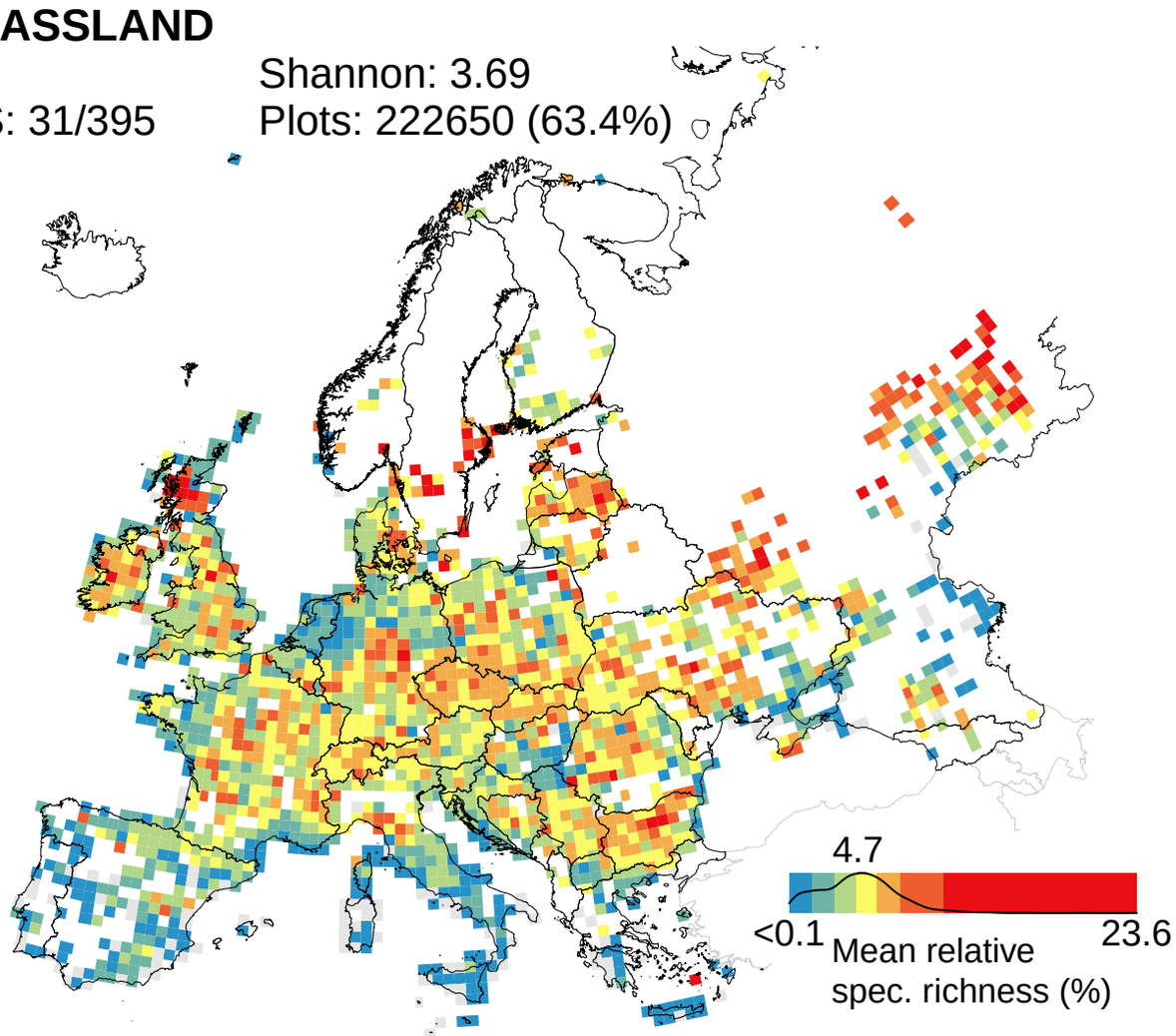
FOREST

G/S: 30/359
Shannon: 3.55
Plots: 182270 (79.2%)



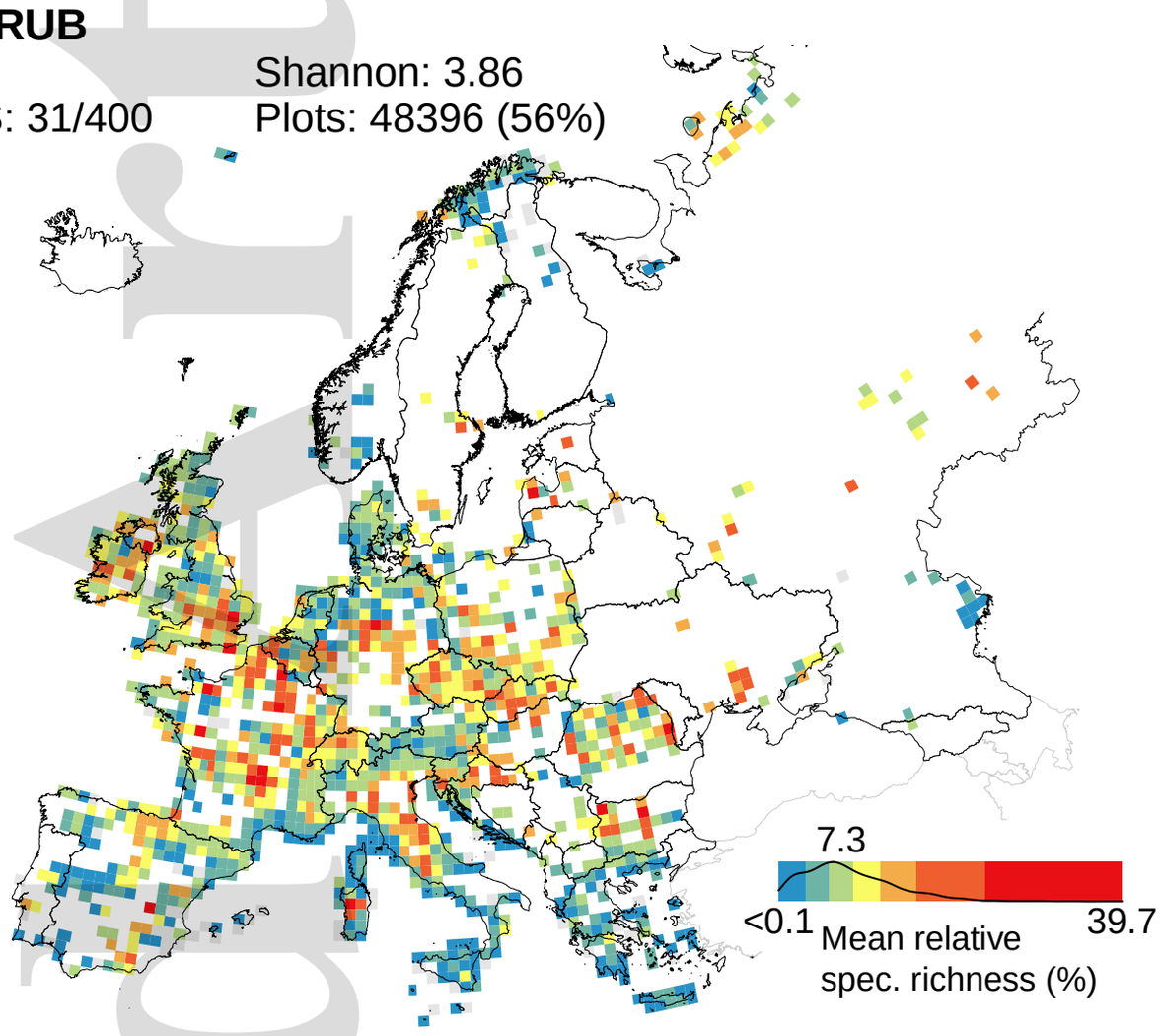
GRASSLAND

G/S: 31/395
Shannon: 3.69
Plots: 222650 (63.4%)



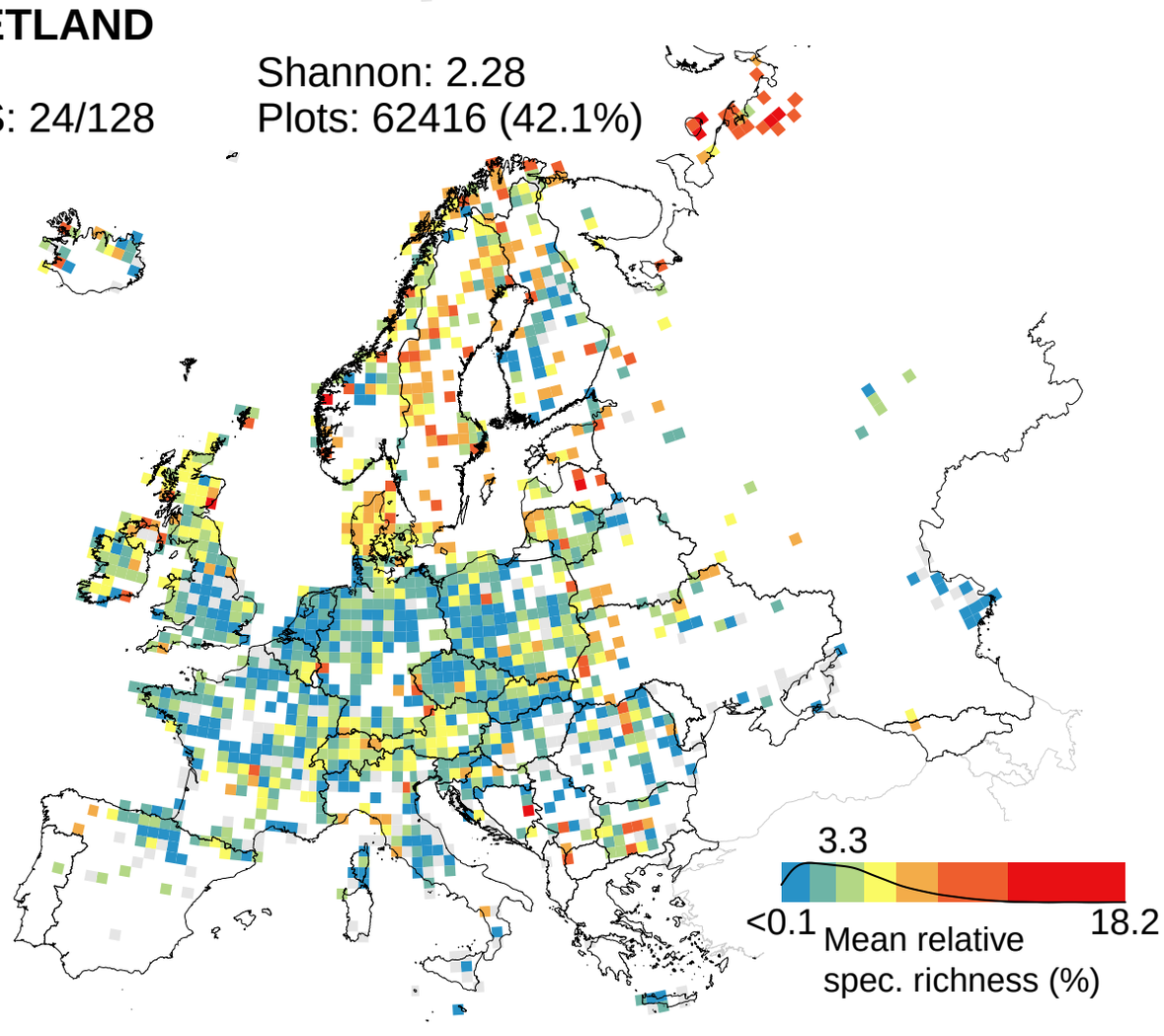
SCRUB

G/S: 31/400
Shannon: 3.86
Plots: 48396 (56%)



WETLAND

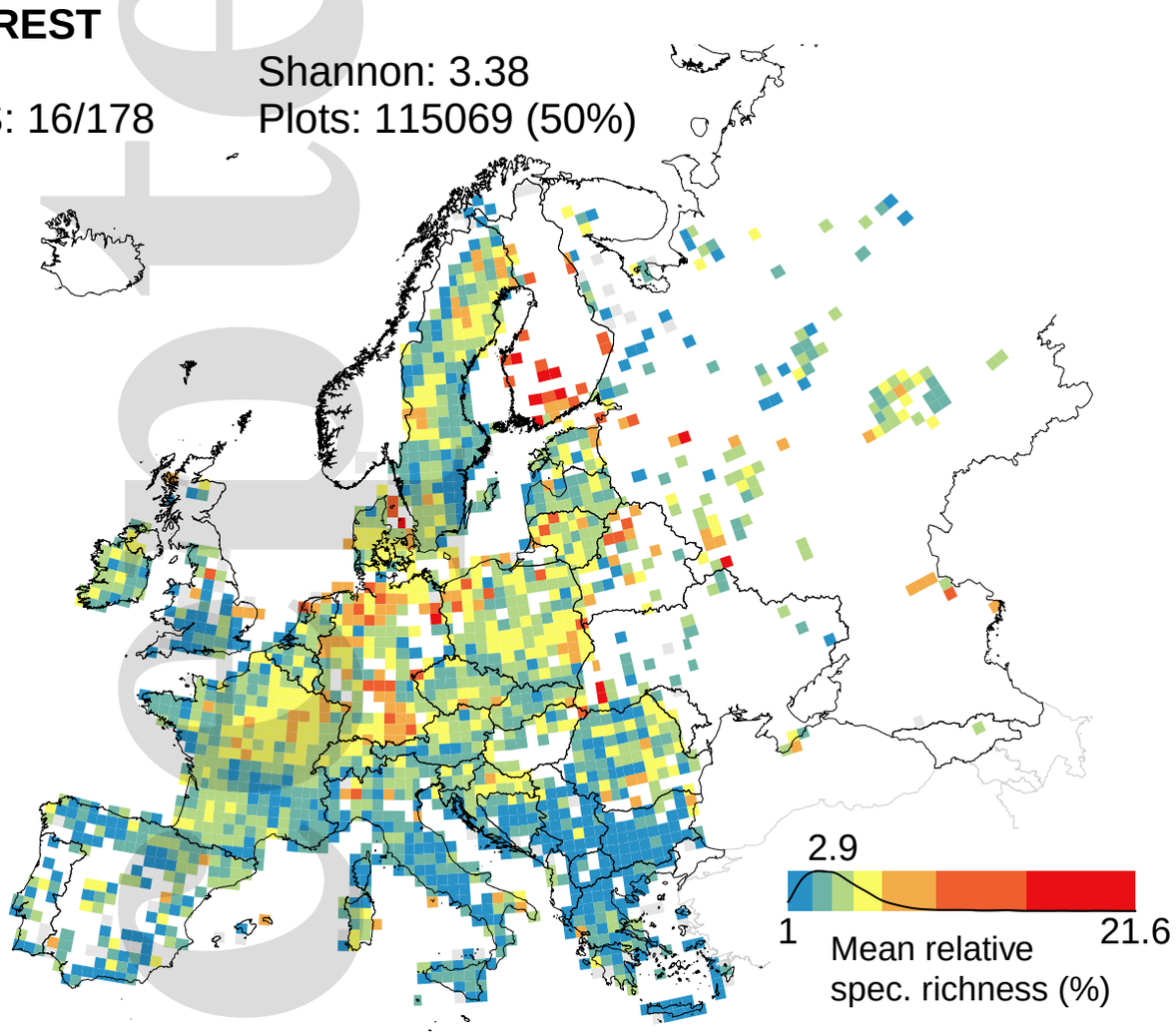
G/S: 24/128
Shannon: 2.28
Plots: 62416 (42.1%)



Cyperaceae

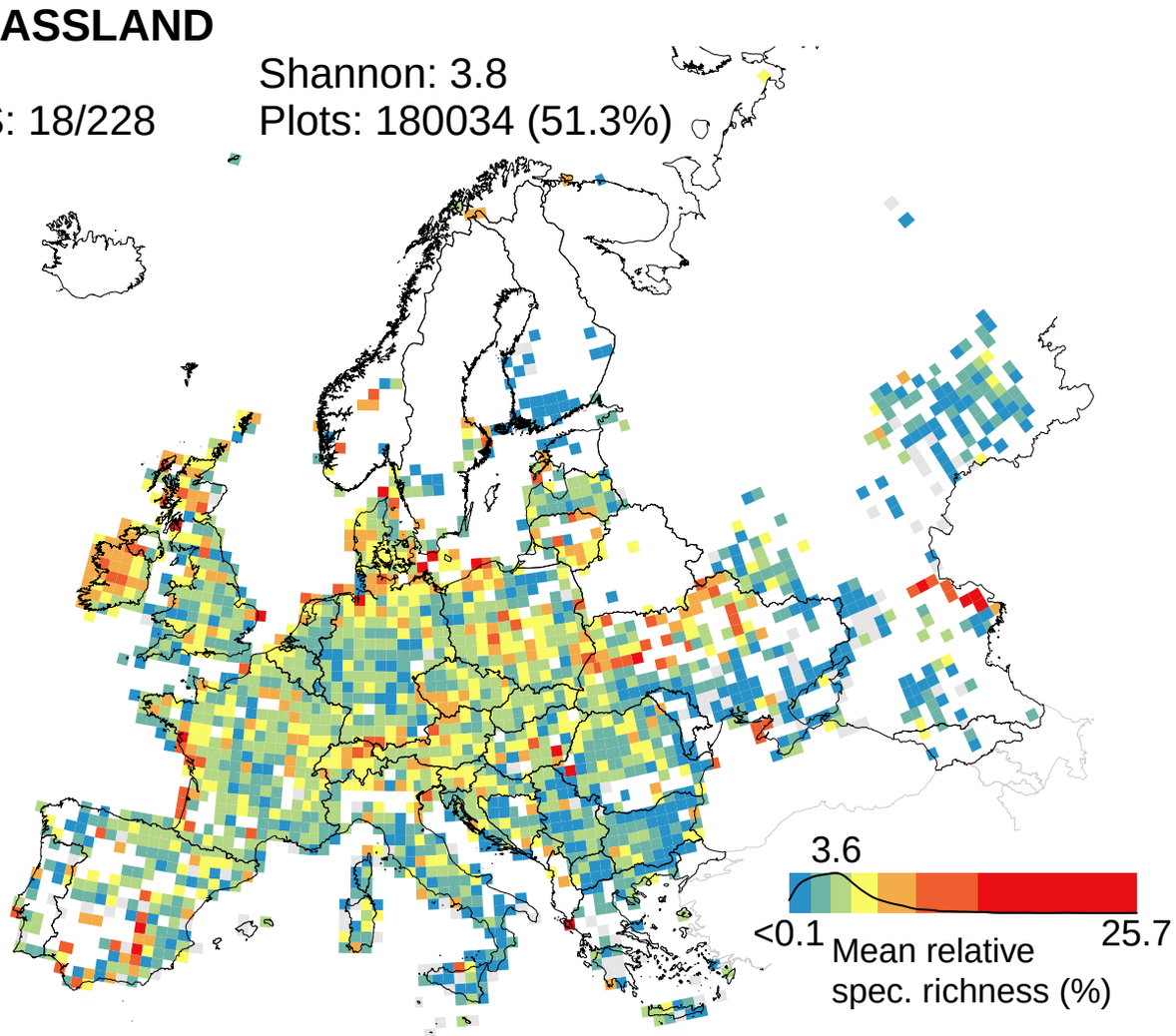
FOREST

G/S: 16/178
Shannon: 3.38
Plots: 115069 (50%)



GRASSLAND

G/S: 18/228
Shannon: 3.8
Plots: 180034 (51.3%)



SCRUB

G/S: 15/189
Shannon: 3.49
Plots: 42128 (48.8%)

WETLAND

G/S: 17/221
Shannon: 3.69
Plots: 120583 (81.3%)

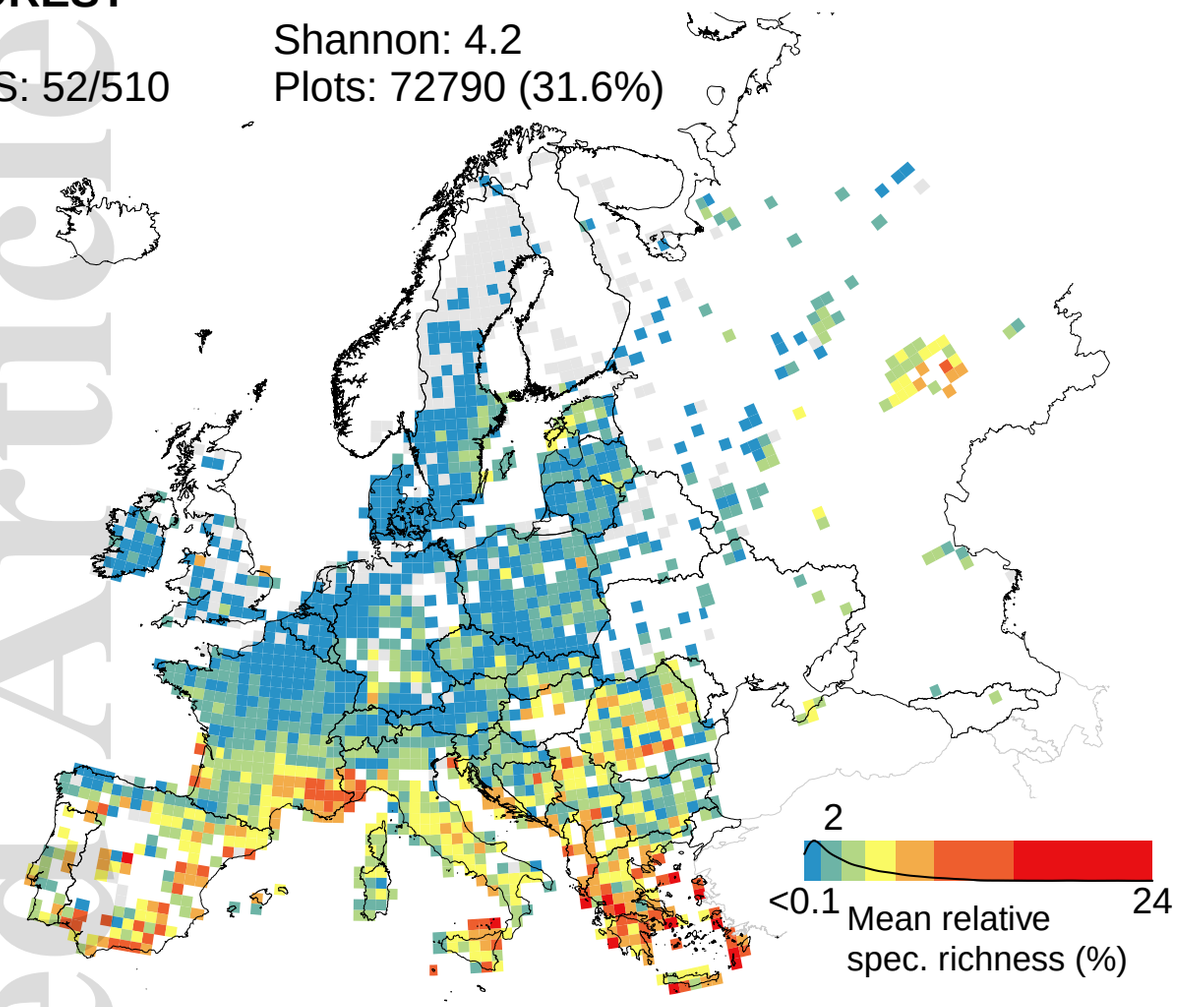
Fabaceae

FOREST

G/S: 52/510

Shannon: 4.2

Plots: 72790 (31.6%)

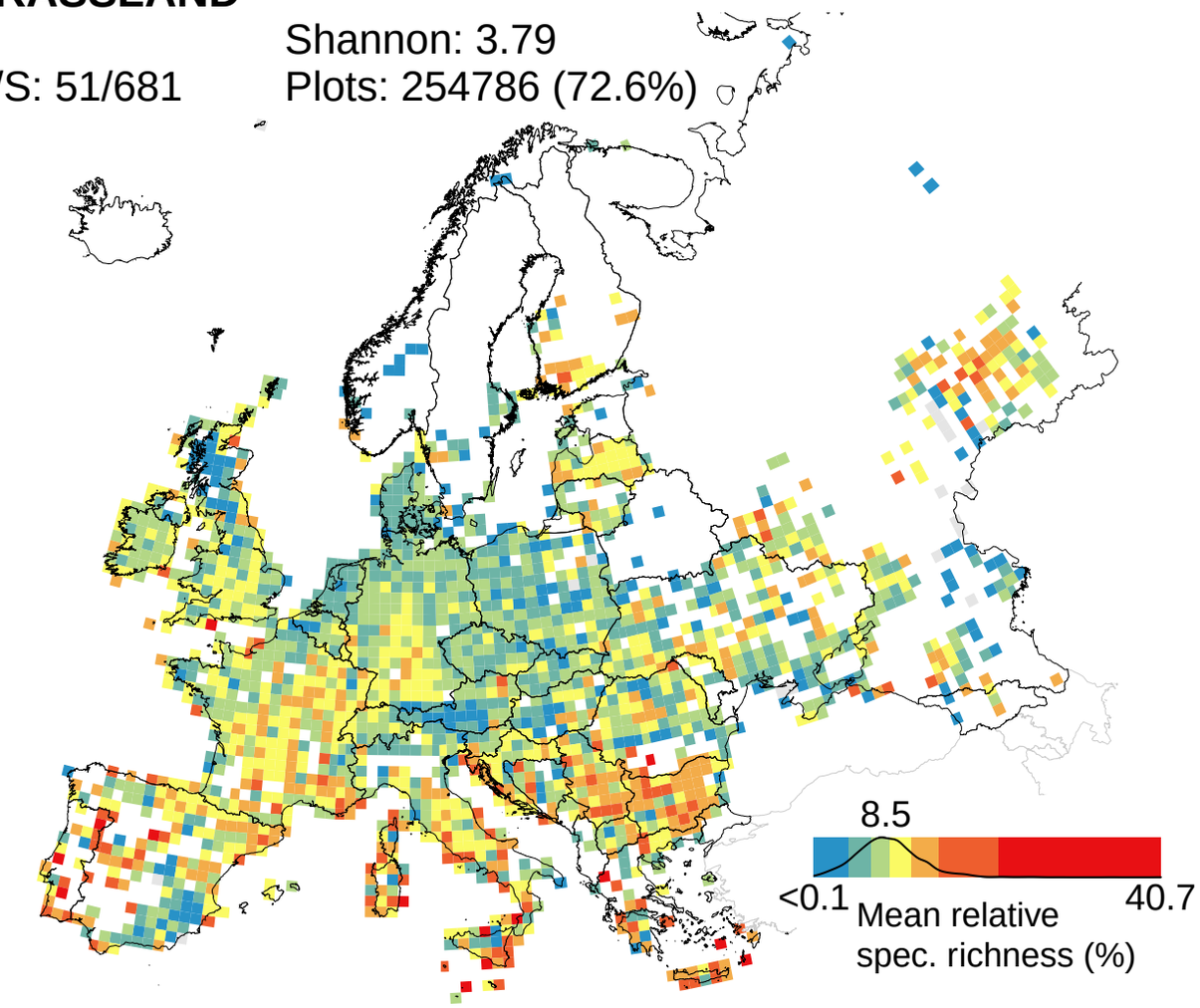


GRASSLAND

G/S: 51/681

Shannon: 3.79

Plots: 254786 (72.6%)

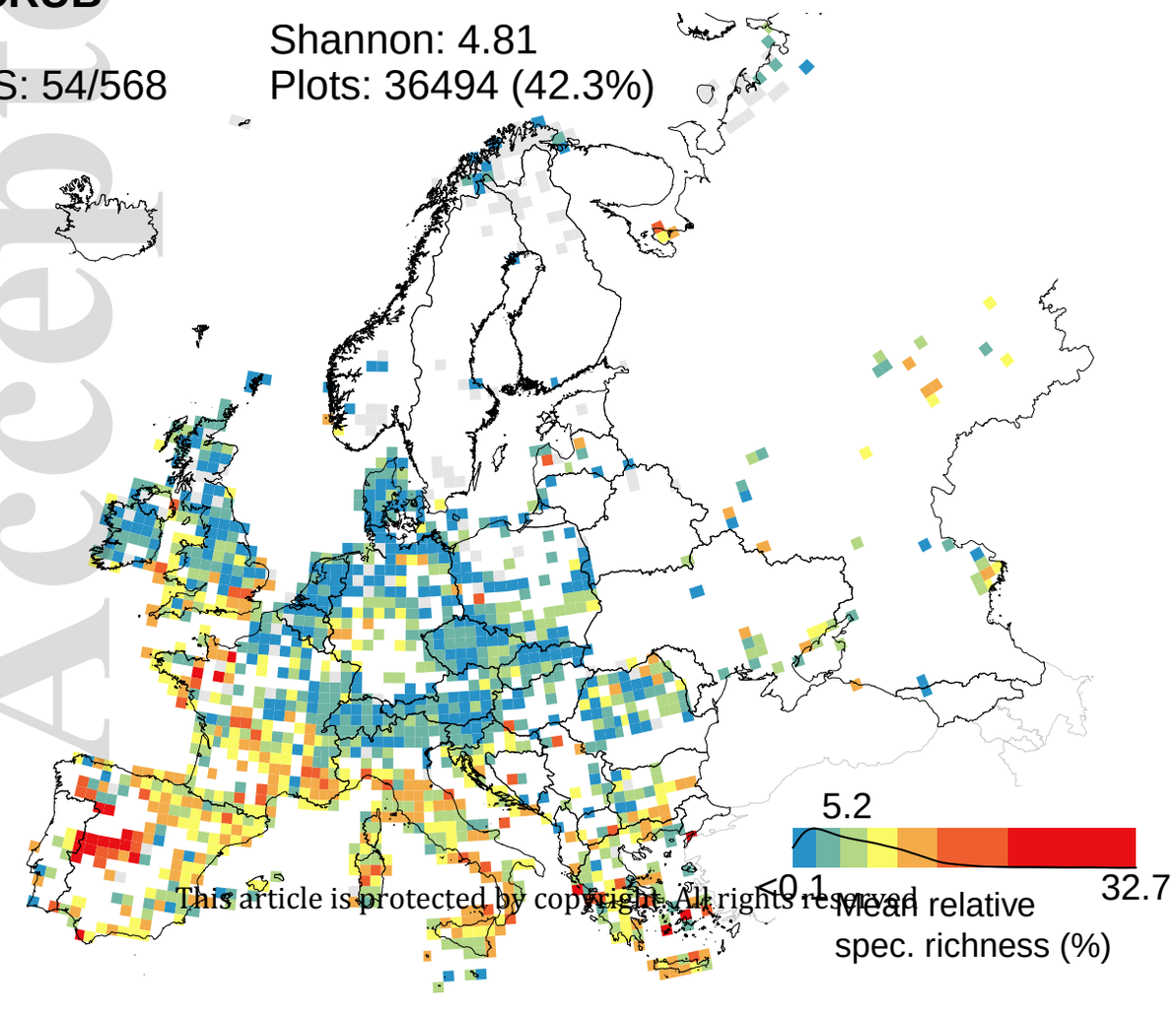


SCRUB

G/S: 54/568

Shannon: 4.81

Plots: 36494 (42.3%)

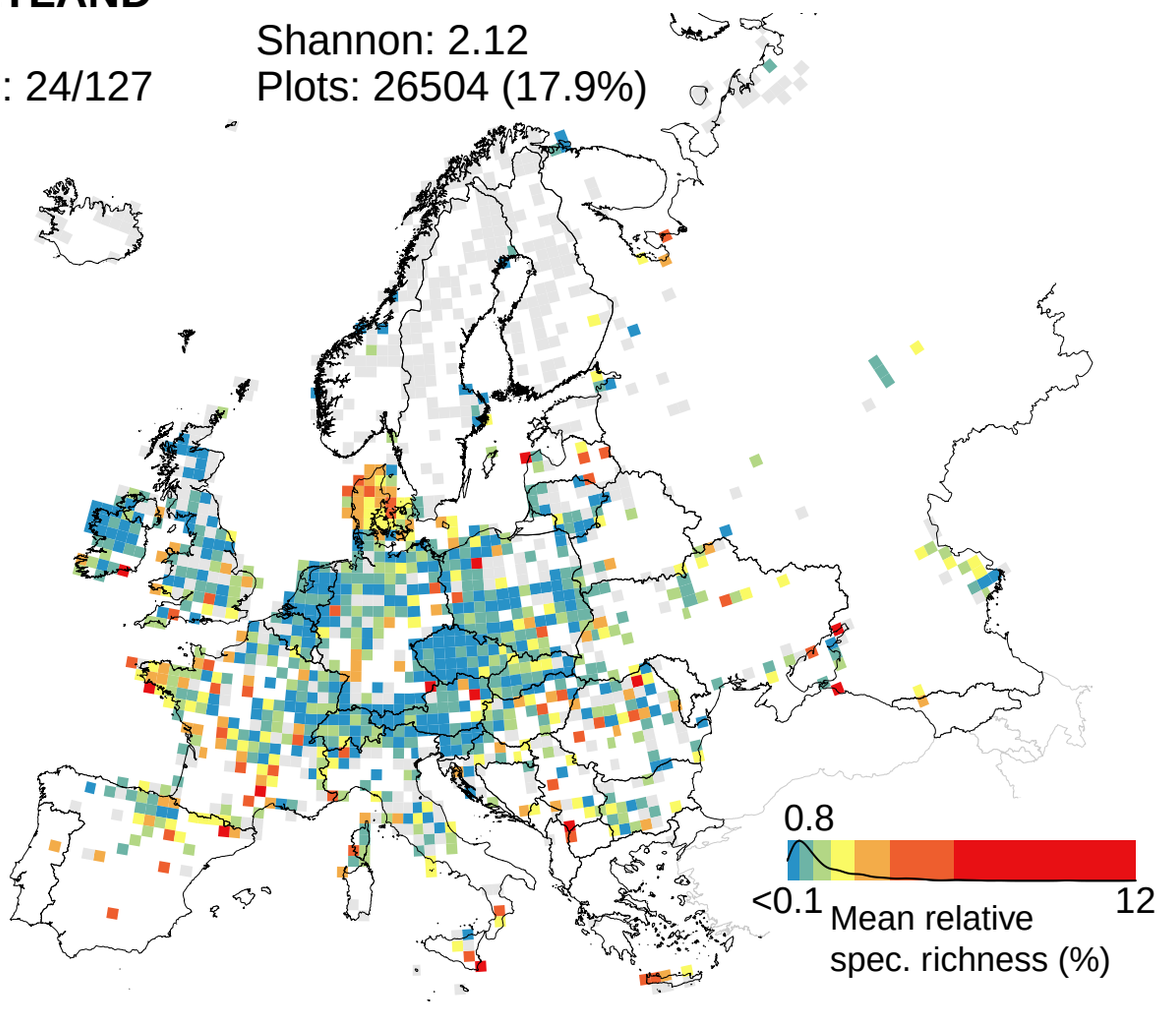


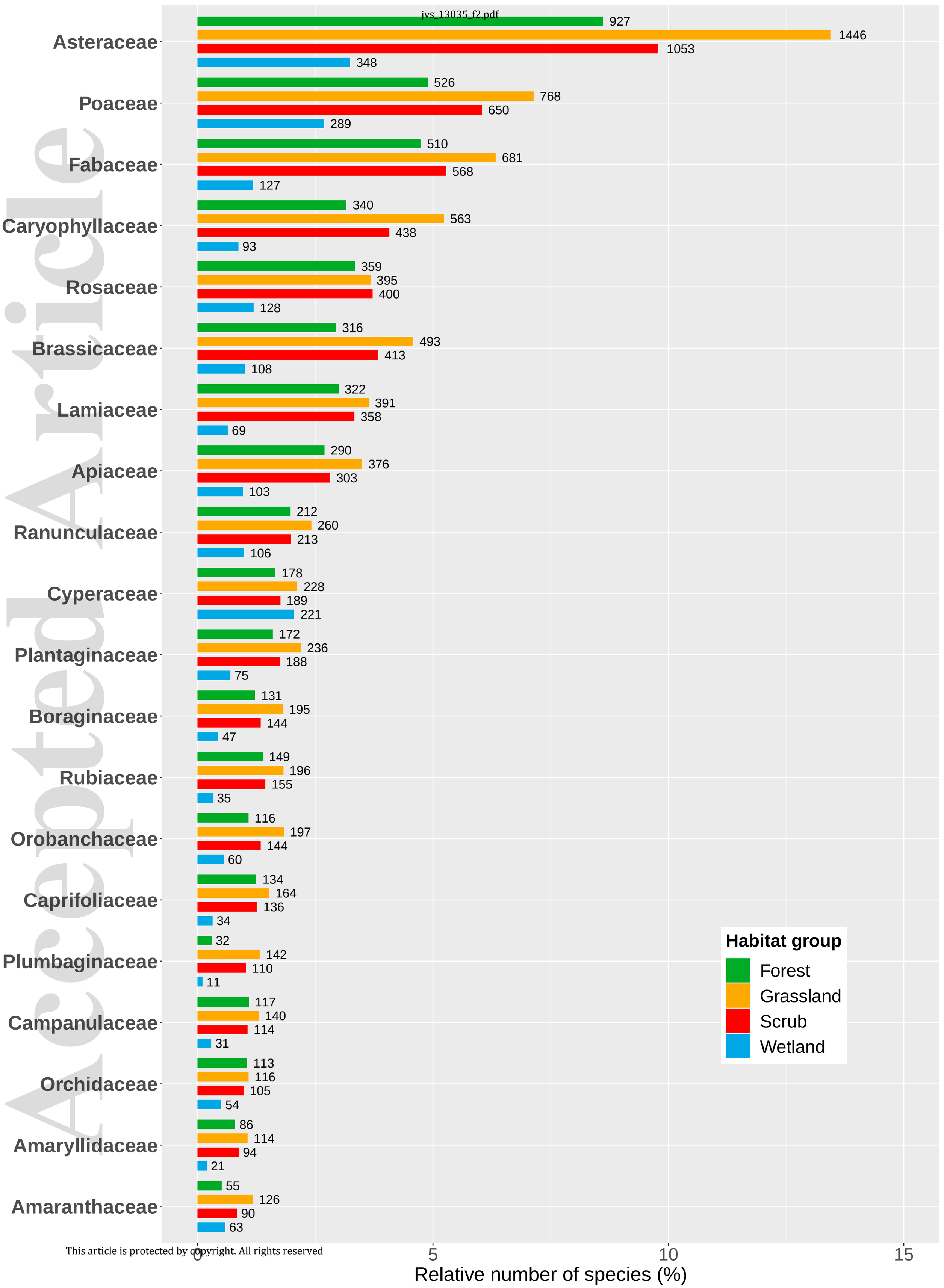
WETLAND

G/S: 24/127

Shannon: 2.12

Plots: 26504 (17.9%)





(a) Mean number of families

(b) Mean species:family ratio

FOREST

15
4
34

1.5
1
2.7

GRASSLAND

12
3
22

1.9
1
3.3

SCRUB

10
1
28

1.6
1
3.2

WETLAND

7
1
23

1.5
1
2.8