

Mapping plant-pollinator interactions across Europe

Abstract

Pollinators play a crucial role in maintaining Earth's terrestrial biodiversity and human food production by mediating sexual reproduction for most flowering plants. However, their diversity and role as pollinators are increasingly compromised by rapid human-induced environmental changes. One of the major challenges for pollinator conservation, is the lack of robust generalisable data across space and time to comprehend the conservation status and population trends among different pollinator species. Here, we present a dataset of plant-pollinator interactions at European level that consists of 51 studies distributed across 17 countries that comprise over a million of interactions between plants and pollinators. The dataset includes a total of 2065 pollinator and 1353 plant species that accounts approximately for 30% of each of the main pollinator groups (i.e., bees, syrphids and butterflies) and 5% of flowering plants that inhabit the European continent.

Introduction

1st paragraph

General introduction of how global change impacts plant-pollinator interactions

Maybe expand on some drivers? Climate change, habitat fragmentation, agricultural intensification, urbanization, pollution, pesticides and species' invasions

Highlight the relevance of large scale datasets to understand biological phenomena

2nd paragraph

3rd paragraph

4paragraph

Introduce research questions

LIST THEM HERE (Main ideas so far)

Questions that we would like to answer:

- 1) What are the most common plant a pollinator species? Are those shared across networks? Most common interactions across Europe? Interaction fidelity
- 2) Is generalization the rule? Or specialization? How this impacts indirect interactions? Go in the direction of pollinator importance?

Methods

Dataset description

This European metaweb consist of datasets of plant-pollinator interactions compiled initially by a wide number of researchers and institutions within the European continent. This dataset consist of 51 independent published and unpublished studies conducted during the time period 2004 - 2021 on 17 different countries (**Figure 1a**). The dataset contains a total of 1,150,097 distinct interactions, considering interaction as the contact of a given pollinator to the reproductive structure of a particular plant, from 2,065 pollinator and 1,353 plant species. The majority of plant and pollinator species tend to be regionally specific, with only a minor portion of them being shared across a broad range of studies (**Figure 1b**). These different studies differ in sampling effort and methodology, although most studies took place within a single flowering season (68.63%), sampled a given location an average of 7.54 days, and documented interactions mostly by using transects as sampling method (62.75%).

Data structure?

TABLE? Describe cols?

Taxonomic harmonization

All plant and pollinator species names were checked and standardized according to large scale taxonomic databases. To ensure reproducibility of the workflow, we have conducted this harmonization in R with **rgbif** (Chamberlain, Oldoni, and Waller 2022) as pivotal package to check for species names and retrieve further taxonomic information (i.e., phylum, order, family and genus) from the Global Biodiversity Information Facility (GBIF). The protocol for plants and pollinators is similar but slightly different given the availability of the different taxonomic resources. For transparency, we have included in the database the old species name, the new assigned name, and, if the species name is uncertain (e.g., species complex or species alike).

For plants: (i) we initially verify exact matches with the GBIF species checklist; (ii) select unmatched cases and fix orthographic errors; (iii) retrieve again taxonomic information for those unmatched cases, evaluate accuracy of fuzzing matching and programmatically fix records that are still not found; (iv) finally, we used the World Flora Taxonomic Backbone (Govaerts et al. 2021; WFO, July 7, 2022) as the ultimate filter for taxonomic information as we used it to calculate the plant taxonomic coverage of our database.

For pollinators: (i) we first created a checklist of species names for the most representative pollinator groups at European level by combining the recently published checklists of bees (Ghisbain et al. 2023), syrphids (Kočić et al. 2023) and butterflies (Wiemers et al. 2018); (ii) then, we compared pollinator species names against the checklist and recovered some

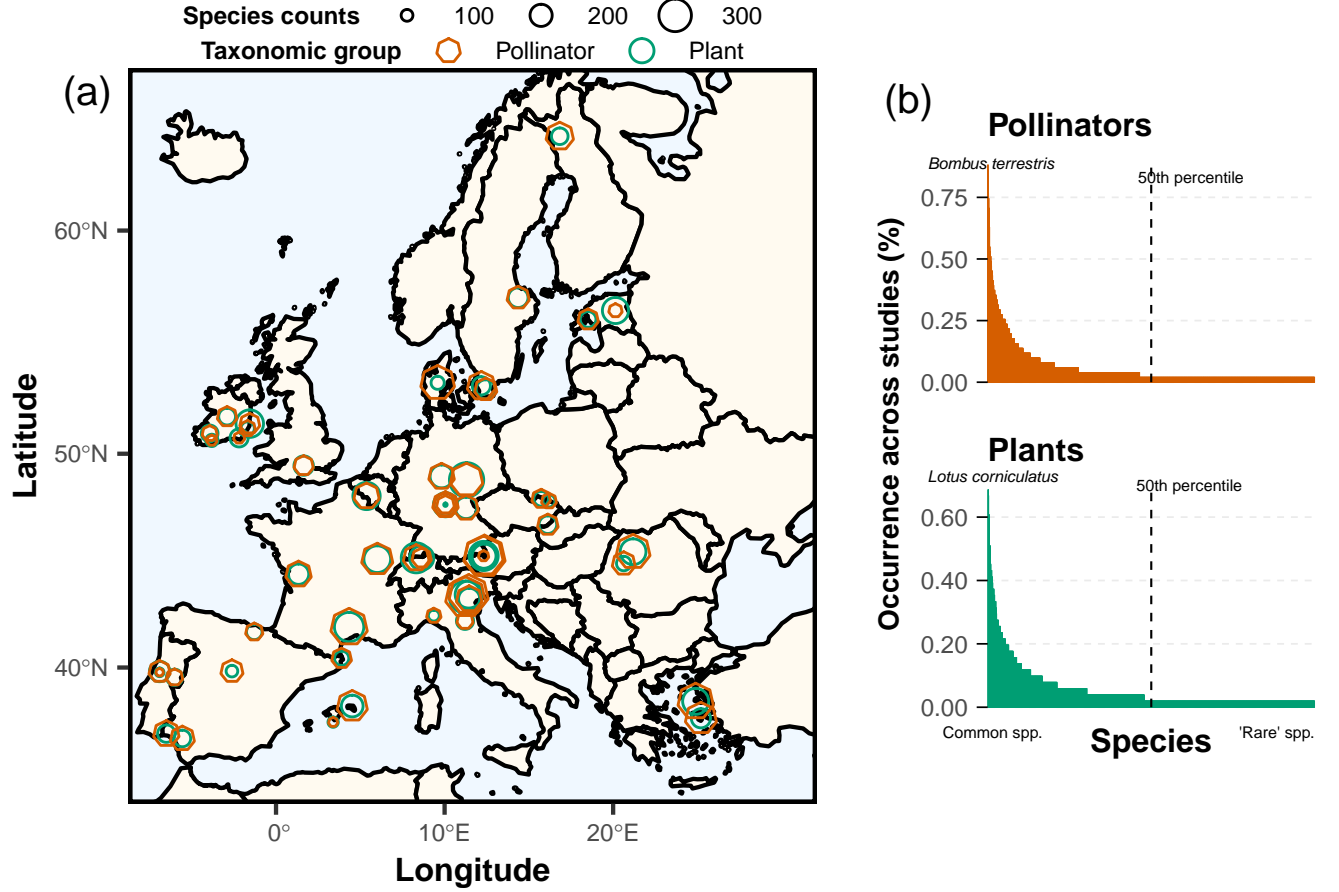


Figure 1. (a) Approximate locations of the studies of the SafeNet database showing the total number of pollinator (i.e., orange heptagon) and plant (i.e., green circles) species per study. The sizes of these shapes are proportional to the respective species counts. For visualization purposes, we have focused only on the European region with studies and selected a single location per study. (b) Percentage of shared occurrences of pollinators and plants across studies. Species on the left are highly shared across studies, while species on the right are found only in a few or a single study. As a reference, we have indicated the 50th percentile with a black dashed vertical line.

unmatched cases with restrictive fuzzy matching by using **stringdist** package (Van der Loo et al. 2014); (iii) we programmatically fixed unmatched records when necessary and retrieved the taxonomic information for all species from GBIF; (iv) we fixed the non-found cases in the GBIF checklist and made sure that all species names from bees, syrphids and butterflies were named according to their respective species checklists.

Taxonomic coverage

To assess the completeness of plant and pollinator species in the SafeNet database at European level, we used the aforementioned checklists for plants and pollinators. Specifically for plants, we refined the checklist to include only European flowering plants and excluded taxonomic groups not associated with biotic pollination. We did this by first excluding the families considered to have exclusively a wind pollination mode (see Culley, Weller, and Sakai 2002), and then by filtering out the genera with wind or non-biotic pollination from families that exhibit both biotic and non-biotic pollination modes. Additionally, we manually included some exotic species and added unresolved species names that were not present in the accepted names of the checklist at the current version of usage. For pollinators, we compared only the taxonomic coverage of bees, syrphids and butterflies by using their species checklists at European level.

Habitat type and bioclimatic region

The different sites per study were classified with a habitat type by the authors. As these habitats are not standardized across studies, they were standardized with the additional help of land cover information and visual checks on current satellite imagery. For each georeferenced site, the land cover information was extracted from Corine Land Cover (CLC, version 2018) with the help of the Terra package (Hijmans et al. 2022). Based on the habitat classification from the authors and the CLC classification, we created habitat categories that intend to summarize the diversity of habitats in the SafeNet database (see habitat type definition in supplementary material). These categories allow a quick comparison and understanding of the habitat types from the database. However, we advise authors to revise this classification if they intend to rely on this field for their analyses as this is a non-fully objective process. Moreover, Europe is characterized by a great variety of environmental conditions that harbor different biota. Thus, to allow authors to explore set of studies that share similar environmental conditions and species, we assigned to each site a biogeographical region. The biogeographical regions were downloaded from the European Environment Agency (version 2016) and were matched to the different sites with the help of a spatial join from the **sf** package (Pebesma et al. 2018).

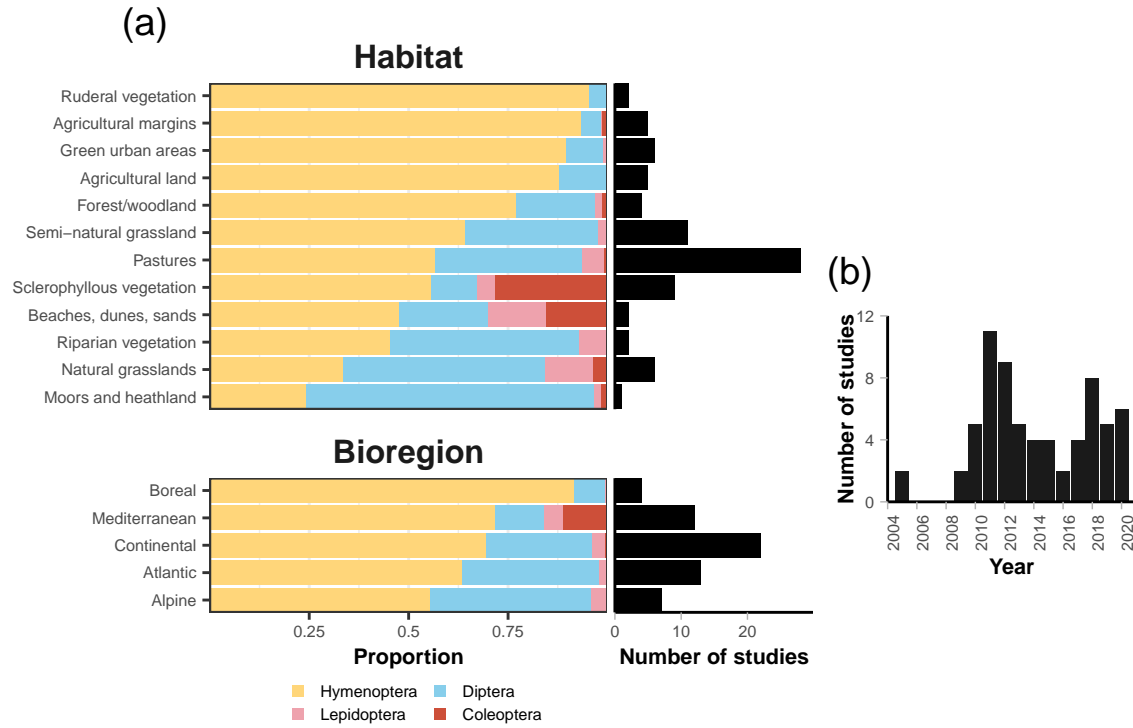


Figure 2. Proportion of the major pollinator orders by habitat types and bioclimatic regions in the SafeNet database. The orders, from left to right, include Hymenoptera, Diptera, Lepidoptera and Coleoptera. The horizontal barplot on the right indicates the number of studies that were conducted on each habitat type or bioclimatic region. Note that a single study can contribute to more than one habitat or bioclimatic region. Areas with a greater number of studies are more likely to depict accurate proportions of pollinators in those systems.

Results

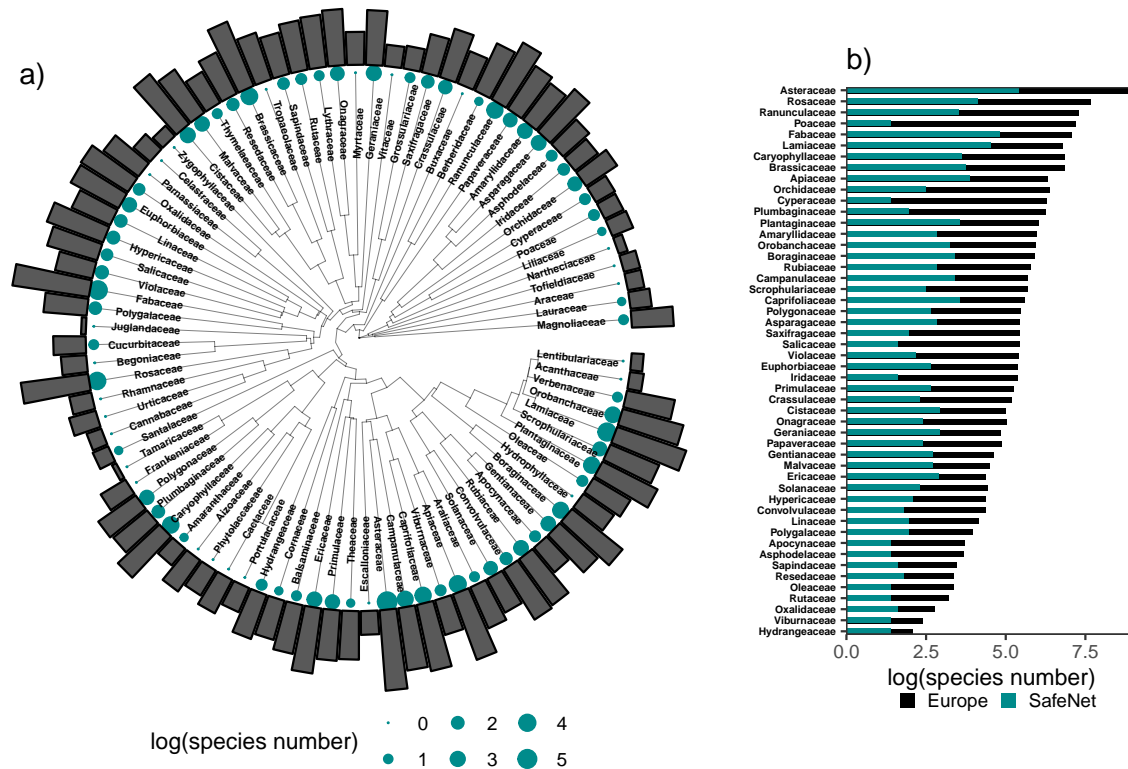


Figure 3. a) Phylogenetic tree of plant families in the SafeNet database, showing the number of species per family in logarithmic scale (blue circles) and the total number of interactions recorded per family (gray bars). b) Comparison of the top 50 families with the highest number of species in the SafeNet database against the total number of species recorded for those families at the European level. Species counts are displayed in logarithmic scale for better visualization.

FIGURE 4 POLLINATOR COVERAGE

FIGURE 5? CLIFFHANGER OF NETWORK METRICS

Taxonomy and taxonomic coverage

Habitat type and bioclimatic region

Discussion

Geographic and taxonomic biases!

References

- Chamberlain, Scott, Damiano Oldoni, and John Waller. 2022. “Rgbif: Interface to the Global Biodiversity Information Facility API.”
- Culley, Theresa M, Stephen G Weller, and Ann K Sakai. 2002. “The Evolution of Wind Pollination in Angiosperms.” *Trends in Ecology & Evolution* 17 (8): 361–69.
- Ghisbain, Guillaume, Paolo Rosa, Petr Bogusch, Simone Flaminio, ROMAIN LE DIVELEC, Achik Dorchin, Max Kasperek, et al. 2023. “The New Annotated Checklist of the Wild Bees of Europe (Hymenoptera: Anthophila).” *Zootaxa* 5327 (1): 1–147.
- Govaerts, Rafaël, Eimear Nic Lughadha, Nicholas Black, Robert Turner, and Alan Paton. 2021. “The World Checklist of Vascular Plants, a Continuously Updated Resource for Exploring Global Plant Diversity.” *Scientific Data* 8 (1): 215.
- Hijmans, Robert J, Roger Bivand, Karl Forner, Jeroen Ooms, Edzer Pebesma, and Michael D Sumner. 2022. “Package ‘Terra’.”
- Kočić, Anja, Ante Vujić, Tamara Tot, Marina Janovikć Milosavljevuć, and Maarten De Groot. 2023. “An Updated Checklist of the Hoverflies (Diptera: Syrphidae) of Slovenia.” *Zootaxa* 5297 (2): 189–227.
- Pebesma, Edzer J et al. 2018. “Simple Features for r: Standardized Support for Spatial Vector Data.” *R J.* 10 (1): 439.
- Van der Loo, Mark PJ et al. 2014. “The Stringdist Package for Approximate String Matching.” *R J.* 6 (1): 111.
- Wiemers, Martin, Emilio Balletto, Vlad Dincă, Zdenek Faltýnek Fric, Gerardo Lamas, Vladimir Lukhtanov, Miguel L Munguira, et al. 2018. “An Updated Checklist of the European Butterflies (Lepidoptera, Papilionoidea).” *ZooKeys*, no. 811: 9.

Supplementary material

Supplementary text 1

Habitat definitions:

- 1) **Ruderal vegetation:** Plants growing on highly disturbed sites such as road sides or mineral extraction sites.
- 2) **Agricultural margins:** Sides of crops that can include any type of vegetation from low growing plants to trees.
- 3) **Green urban areas:** Parks, private gardens or small pastures within an urban setting. Botanical gardens are included in this category.
- 4) **Agricultural land:** Includes any type of crop and any type of vegetation growing within them.

- 5) **Forest/woodland understory:** Any plant community sampled under a wooded group of plants. The forest could be embedded in an agricultural setting or in a fully natural scenario. We have included here agro-forestry areas and open to dense forest. Note that we have excluded from this category forest that contains sclerophyllous vegetation.
- 6) **Semi-natural grassland:** Low growing plant community with relatively low disturbances but under low pressure such as seasonal mowing or extensive grazing.
- 7) **Pastures:** Any type of low growing plant community that is highly influenced by human disturbance. For instance, agriculture, mowing, moderate to high grazing or urban environments. Note that this category also includes old pastures with regrowth of woody vegetation.
- 8) **Sclerophyllous vegetation:** Any type of system with a dominant shrub community adapted to drought. Typical of the Mediterranean region. Note, that we have include in this category also woodlands (open coniferous forest) where the shrub community was the main focus of the study.
- 9) **Beaches, dunes, sands:** Plant communities growing on sandy soil.
- 10) **Riparian vegetation:** Plant communities growing on river margins.
- 11) **Natural grasslands:** Low growing plant communities with little or none human disturbance. Often located in high elevation areas within Europe.
- 12) **Moors and heathland:** Low growing woody vegetation characteristic from low fertile soils near the coast or in alpine areas.