



OPEN

DATA DESCRIPTOR

Associations between carabid beetles and fungi in the light of 200 years of published literature

Gábor Pozsgai^{1,2,3,17}✉, Ibtissem Ben Fekih^{1,4,17}, Markus V. Kohnen⁵, Said Amrani⁶, Sándor Bérces^{7,8}, Dávid Fülöp⁹, Mohammed Y. M. Jaber¹⁰, Nicolai Vitt Meyling¹¹, Malgorzata Ruskiewicz-Michalska¹², Walter P. Pfliegler¹³, Francisco Javier Sánchez-García^{14,15}, Jie Zhang¹, Christopher Rensing⁴✉, Gábor L. Lövei^{2,16} & Minsheng You^{1,2}✉

Describing and conserving ecological interactions woven into ecosystems is one of the great challenges of the 21st century. Here, we present a unique dataset compiling the biotic interactions between two ecologically and economically important taxa: ground beetles (Coleoptera: Carabidae) and fungi. The resulting dataset contains the carabid-fungus associations collected from 392 scientific publications, 129 countries, mostly from the Palearctic region, published over a period of 200 years. With an updated taxonomy to match the currently accepted nomenclature, 3,378 unique associations among 5,564 records were identified between 1,776 carabid and 676 fungal taxa. Ectoparasitic Laboulbeniales were the most frequent fungal group associated with carabids, especially with Trechinae. The proportion of entomopathogens was low. Three different formats of the data have been provided along with an interactive data digest platform for analytical purposes. Our database summarizes the current knowledge on biotic interactions between insects and fungi, while offering a valuable resource to test large-scale hypotheses on those interactions.

¹State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, Institute of Applied Ecology, Fujian Agriculture and Forestry University, Fuzhou, 350002, China. ²Joint International Research Laboratory of Ecological Pest Control, Ministry of Education, Fuzhou, 350002, China. ³CE3C – Centre for Ecology, Evolution and Environmental Changes, Azorean Biodiversity Group and Universidade dos Açores, Angra do Heroísmo, 9700-042, Azores, Portugal. ⁴Institute of Environmental Microbiology, College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou, 350002, China. ⁵Basic Forestry and Proteomics Research Center, College of Life Science, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Fujian Agriculture and Forestry University, Fuzhou, 350002, China. ⁶Laboratoire de Biologie et de Physiologie des Organismes, Faculté des Sciences Biologiques, Université des Sciences et de la Technologie Houari Boumediène, BP 32 El Alia, Alger, 16111, Algeria. ⁷Duna-Ipoly National Park Directorate, Költő u. 21, H-1121, Budapest, Hungary. ⁸Juhász-Nagy Pál Doctoral School, University of Debrecen, Egyetem tér 1, H-4032, Debrecen, Hungary. ⁹Department of Zoology, Plant Protection Institute, Centre for Agricultural Research, Nagykovácsi út 26-30, H-1029, Budapest, Hungary. ¹⁰Fujian University Key Laboratory for Plant-Microbe Interaction, College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou, 350002, China. ¹¹Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871, Frederiksberg C, Denmark. ¹²Department of Algology and Mycology Faculty of Biology and Environmental Protection, University of Łódź, Banacha 12/16, PL-90-237, Łódź, Poland. ¹³Department of Molecular Biotechnology and Microbiology, University of Debrecen, Egyetem tér 1, Debrecen, H-4032, Hungary. ¹⁴Fujian Provincial Key Laboratory of Insect Ecology, College of Plant Protection, Fujian Agriculture and Forestry University, Fuzhou, 350002, China. ¹⁵Área de Biología Animal, Departamento de Zoología y Antropología Física, Facultad de Veterinaria, Universidad de Murcia, Murcia, 30100, Spain. ¹⁶Department of Agroecology, Aarhus University, Flakkebjerg Research Centre, Forsøgsvej 1, DK-4200, Slagelse, Denmark. ¹⁷These authors contributed equally: Gábor Pozsgai, Ibtissem Ben Fekih. ✉e-mail: Pozsgai@coleoptera.hu; rensing@iue.ac.cn; msyou@fafu.edu.cn

Background & Summary

One of the striking features of the Anthropocene is a rapid degradation of natural ecosystems^{1,2}, and an alarming decline of many species, which ultimately may lead to extinctions^{3–5}. Whereas conserving ecosystem functions is increasingly recognised as a vital need for humans^{6–8}, the interspecific interactions underpinning these functions are poorly understood^{9,10}. However, conserving such interactions can be particularly important when taxa providing high-value ecosystem services are involved^{10,11}.

Ground beetles (Coleoptera: Carabidae) have been long known for their benefits in agroecosystems^{12,13}. They play an important role in suppressing pests¹⁴, but several carabid species also consume seeds of herbaceous plants, making them a valuable asset for weed control as well¹⁵.

Fungi are also of vital significance in most of the world's terrestrial ecosystems¹⁶. Mycorrhizal fungi improve nutrient uptake by a large range of plant species through intimate and specialised associations¹⁷, other fungi play a crucial role in decomposition¹⁸, and yet others are pathogens of both crops and pests in agroecosystems¹⁹. Fungal parasitism is one of the crucial agents of evolution²⁰.

Fungi and carabids often co-occur, and they can potentially interact in many ways. The soil environment carabids often inhabit is a reservoir of fungal propagules where the beetles can feed on spores, hyphae or fruiting bodies²¹. They may also be responsible for dispersal of spores of certain fungi²². Several parasitic or entomopathogenic fungi are in an obligatory relationship with their beetle hosts²³, therefore, the population decline of a ground beetle species could potentially lead to overlooked extinction cascades²⁴. However, our knowledge of the fungal-carabid interactions is still limited concerning the frequency of these interactions and on how their exact nature affect the parties involved. Indeed, we do not even have a catalogue of the carabid-fungi interactions, and they have not yet been organized into a comprehensive database. Such a database would be of particular importance from an integrated pest management point of view because both fungi and carabids can deliver ecosystem services, but how their interactions, and potential synergies or antagonisms, influence the delivery of these services is poorly understood.

In order to have a detailed overview of the interactions between Carabidae and the fungal kingdom, we collated a database containing previously reported associations between these taxa. Carabid and fungal species involved in the interaction, the type of the interaction (e. g. parasitic, pathogenic, mutualistic, or trophic interactions), the location (country) the interaction was reported from, and the publication source combined with detailed notes to each questionable entry comprised one record. Publications available in printed formats only were either digitized and data were extracted using semi-automatic text-mining processes, or they were manually screened. We aimed at possible completeness, using a wide range of databases and search engines and several languages to cover most of the published literature.

Both ground beetle and fungal names were validated and their higher taxonomical classifications were also extracted. When it was possible, historical localities were converted to their current country names. The full bibliographical details were also stored in the database.

The database covers a time-period from 1793 to 2020, spans over all geographic sub-regions defined by the United Nations (“UNSD — Methodology”, unstats.un.org. Retrieved 2020–10–11) with recorded associations from 129 countries. Our effort yielded 3,378 unique associations in 5,564 records between 1,776 carabid and 676 fungal species. Although rapidly developing molecular methods have largely facilitated the mapping of complex interaction networks in ecological studies^{25–27}, due to the historic nature of our dataset, most of the records rely on traditional taxonomical identification. Yet, 16 records were based purely on metabarcoding studies; comments linked to these associations clearly identify them.

Whilst we found relatively few pathogenic interactions, a great diversity between ectoparasitic Laboulbeniales fungi and carabids was revealed (Fig. 1). Soft bodied, cave-dwelling members of the Trechinae subfamily were particularly prone to these parasitic infections. Little information was available on mutualistic relationships but the presence of *Yarrowia* yeast reported from the gut of several carabid species²⁸ is probably beneficial for both parties. The data show two distinct peaks in publications registering new associations, in the early 19th century and in the late 20th century (Fig. 2a) but the steady increase in the cumulative number of associations (Fig. 2b) suggests that further research is required to fully resolve this association network. Although we believe that most of the data published so far were collected, data submission will remain open to researchers wishing to contribute.

Methods

Data compilation. *Database scraping and literature search.* Global Biotic Interactions is an open access database for finding species interactions²⁹. As a starting point, we queried GloBI's application programming interface (API) using the *rglobi* R package, with the target taxa set to “Carabidae” and “Fungi”. Since tiger beetles (Carabidae: Cicindelinae) are sometimes considered as a separate family, to avoid taxonomic discrepancies, we also searched “Cicindelidae” and “Fungi”. For this query, we used the application programming interface (API) provided by GloBI. We ran a series of further web searches targeting each registered interaction to find the first published sources and collect additional location data. We also mined academic databases and search platforms, including the Web of Science, Scopus, Google Scholar, and Baidu Research with the keyword sets of “Carabidae AND Fungi” or “Carabid* AND Fung*”. Additional searches were conducted using the words for ground beetles and fungi in Arabic, Chinese, Czech, Danish, Dutch, English, French, German, Hungarian, Japanese, Norwegian, Polish, Russian, Spanish, and Swedish languages (Fig. 3).

Data extraction. Each article in our literature database was individually scanned for associations and interactions and the information was mostly manually extracted, in most cases by collaborators who were native speakers of the language in which the publication was written. Articles reporting experiments of beetles being infected with fungi *in vitro* were excluded. However, reports of naturally occurring carabid-fungus associations, even if they were laboratory experiments, were included as well as taxonomic publications matching this

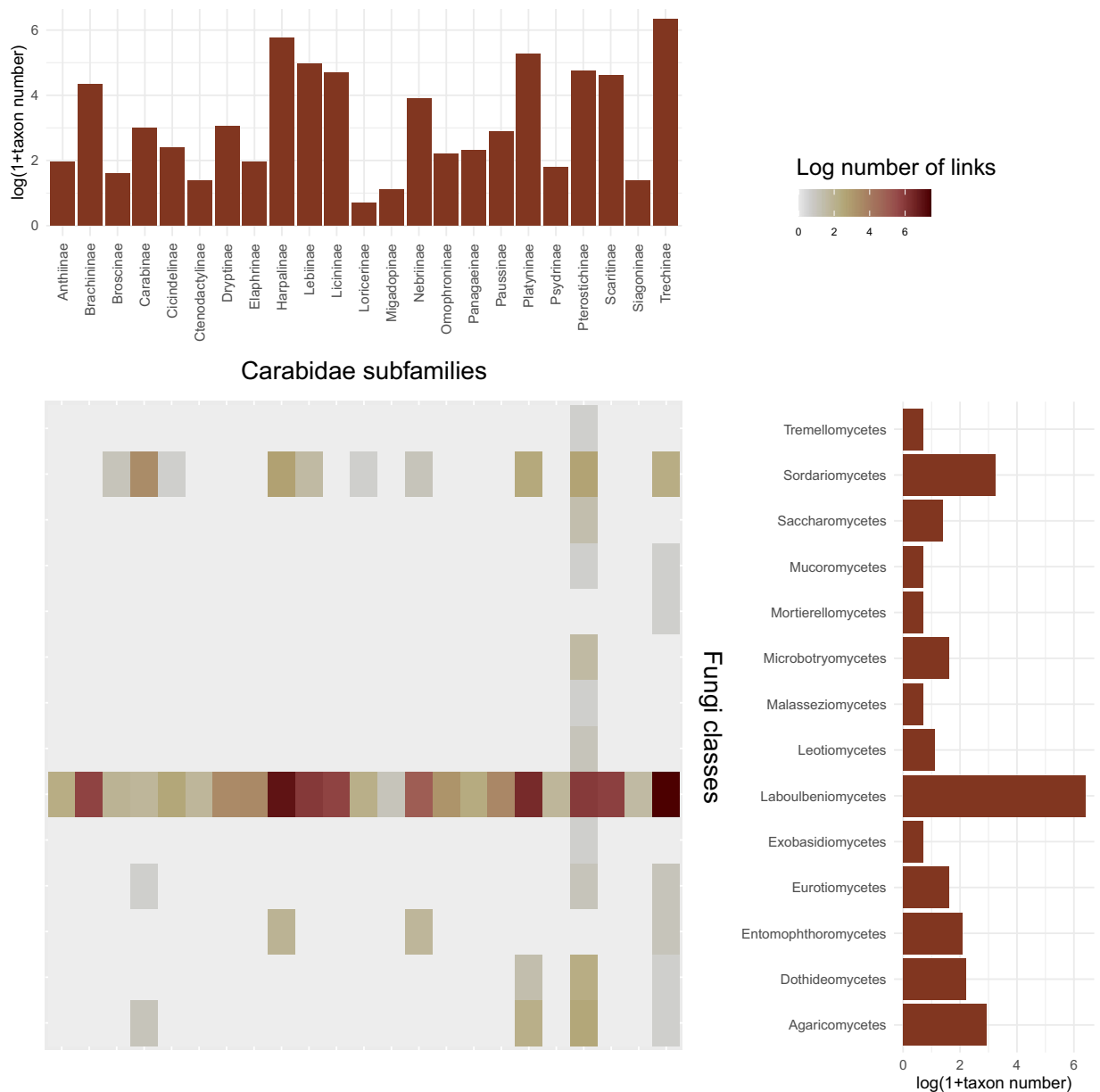


Fig. 1 The number of unique associations between Carabidae subfamilies and fungal classes. Side bar plots show the number of species in each subfamily/class recorded in our dataset.

criterion. Unique associations between carabid and fungal species were individually recorded from each country. Incomplete data with missing species names or country were also extracted and included in the primary dataset. These, however, were filtered out if more precise data were found from the same location or if an earlier, unreported location of a carabid-fungus associations was identified. Alongside each record, the literature source and the contributor who extracted the data were also recorded. This latter served quality control purposes only, and was not included in the final dataset.

Data processing. *Data checking, cleaning, and formatting.* Data were first compiled into one large dataset then organized into separate tables (Online-only Table 1), which were then linked through unique identifiers in the MySQL interactions table. The integrity and uniqueness of each table were checked both programmatically and manually. Taxonomic names were cross-checked with the GBIF Taxonomy Backbone where old names were actualized, and the synonyms changed to currently valid names. When species names published in original articles could not be resolved to any valid taxon, only genera were recorded; if neither genus nor species could be validated, the record was omitted. For instance, *Brachinus* “*laevonaci*” was only recorded as genus *Brachinus*, while the record containing the “*Pardomis*” genus name was excluded. These records, despite not being published in the primary dataset, were kept for further investigations and will be added when new information regarding the taxonomic identity is identified. When species were split into sibling species, it was not possible to back-determine

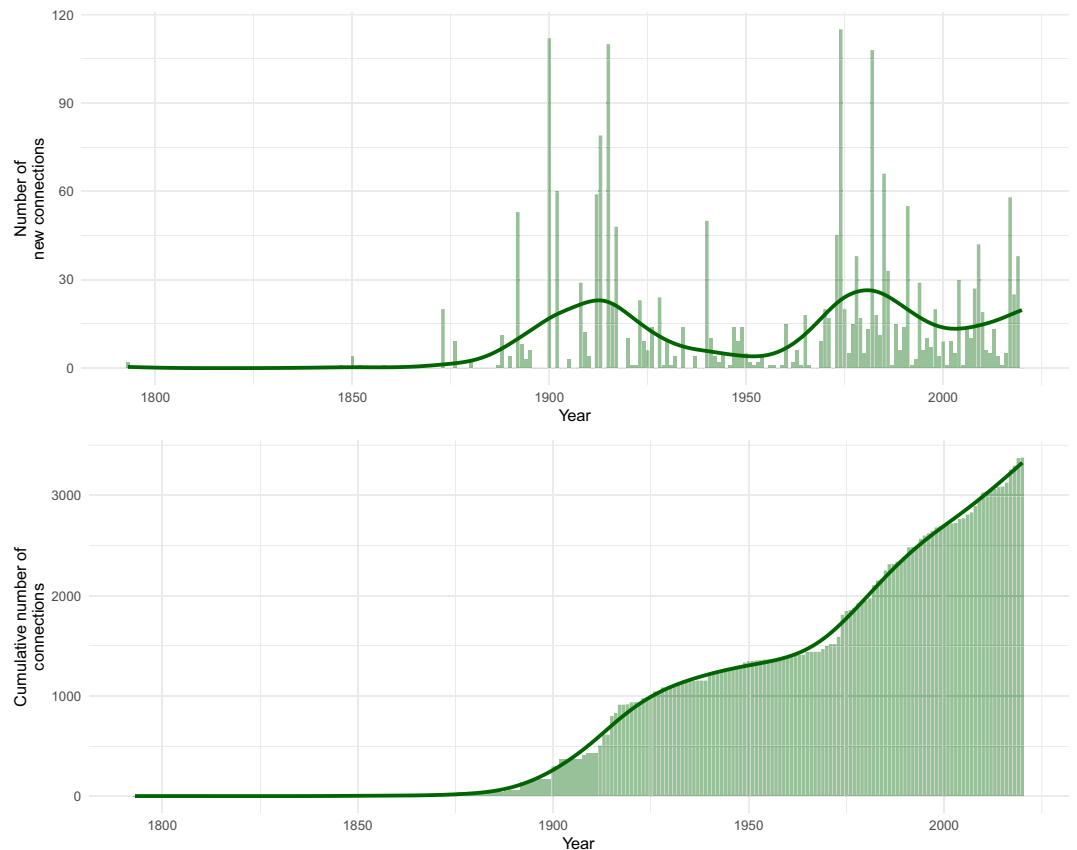


Fig. 2 The number of recorded unique associations over time. Changes in the number of new records (a) and in the cumulative number (b) per year. Dark green lines indicate smoothed trends.

the valid taxon. This was particularly problematic in some recently revised fungal taxa, where the use of novel molecular methods substantially changed the taxonomy. These taxa were recorded as “sensu lato”.

Country names, and their geographic categorization were extracted from an altered version of the United Nations Geoscheme (“UNSD — Methodology”, unstats.un.org. Accessed 2020–10–11). We separated some politically joint but geographically distant regions, mostly islands, to better suit our biodiversity mapping purposes. For instance, Hawaii was handled separately from the U.S. and the Canary Islands were from Spain. A unique combination of one valid carabid, one valid fungal species, and the country where this association was recorded from was used as one data unit. A reference for this unique connection was joined with the data line. When different researchers reported the same association from the same country, the earliest published report was consistently used as a reference, while the later reports were omitted.

Data Records

Available formats and database structure. For our database to be easily accessible to those who are not familiar with MySQL databases, but at the same time keeping the dataset to be expandable, convertible, and queryable, we have provided the data in three different formats: a multi-table relational database in tables of comma separated values, one single MySQL file for the collection of those tables, and a single-file merged data table. All data files are available on *figshare*³⁰: <https://doi.org/10.6084/m9.figshare.14602479.v5>.

Collection of data tables in comma separated values. Both data table collection and the importable SQL database consist of five tables: 1) collecting the unique associations/interactions between Carabidae and Fungi from a particular country, 2) listing carabid and fungal taxonomy, 3) listing the references used in the data collection, 4) geographic data for the countries from where the data were recorded, and 5) a metadata table for annotation (Online-only Table 1). The metadata table contains information on edits and uncertainties within the database. Unlike in the MySQL database, polygonal geometries for countries’ geographic information was not provided in these data tables.

Single data table. Data were also merged into a single.csv file containing 5,564 rows and 42 columns. In this table, column names are the same as in the other data storing methods, but columns were ordered according to their likely importance for the user. Thus, species names and GBIF IDs are presented first, followed by the

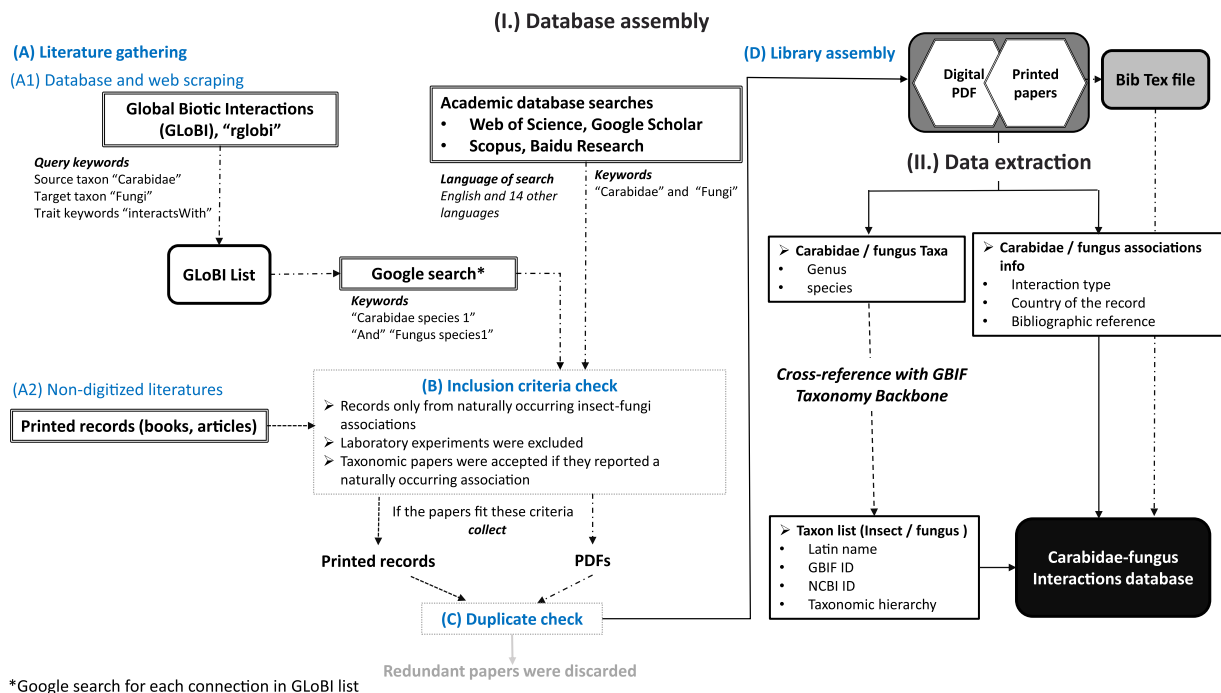


Fig. 3 Flow chart of data collection of carabid-fungus associations from the literature.

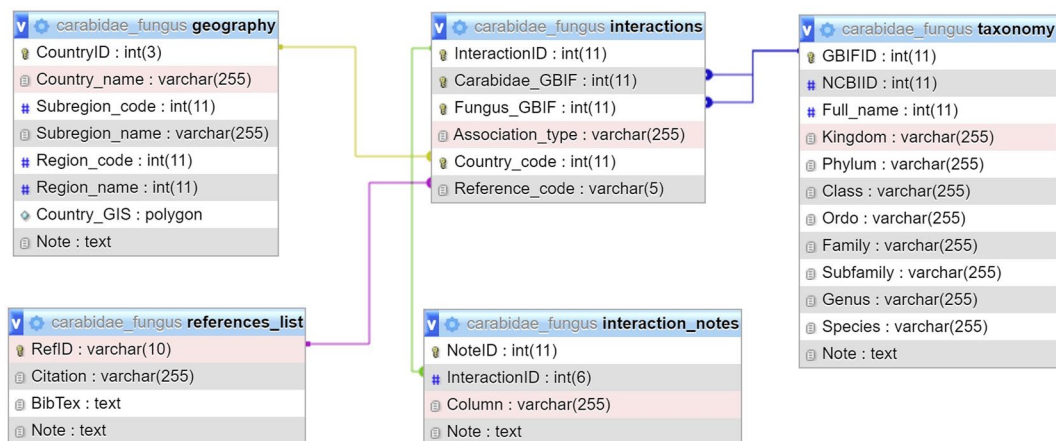


Fig. 4 Structure of the MySQL database. Tables with data recording fields are shown and relationships between tables and linking data fields are marked with coloured lines.

association type, country, and the country code. Next, the unique identifier of the reference source, and the human-readable references follow (Online-only Table 1).

MySQL database. A single.sql file has been provided for importing the entire database with a line of command, or with a few clicks, into an empty MySQL database. After importing, the aforementioned data tables and table links will be made visible (Fig. 4), and the database can be queried directly (e.g. through R).

Interactive web-interface. An interactive Shiny³¹ app has been provided on the <http://54.252.94.194:3838/Carafun/> website to facilitate general data digestion, advanced filtering, and data visualisation.

Technical Validation

The list of species was both programmatically cross-checked against taxonomic databases, and screened by experts in carabid and fungal taxonomy. Data with uncertain taxonomies (i.e. when resolving a reported taxon to a valid GBIF ID has failed) were excluded from the primary database but stored in a separate dataset. These uncertainties affected < 10% of the collected species names, and < 5% of the data. The authors validated association types, and questionable cases were determined by multi-expert consensus. When country categorizations

were unclear, either due to historical changes in boundaries or because of data recording incongruities, the original text was re-checked and, if provided, the country was re-confirmed based on the exact sampling location. Data with missing country records were still included in the main database and the most precise available location was given in a note. However, if an association with no geographic information was a duplicate of a more precise one, only the latter was retained and the other one deleted.

Each data point was also subjected to manual checks and all the edits in the database were recorded and stored as metadata. As a result of the structured relational database design, our database is easy to update, query and is convertible to other structures. Moreover, since the taxonomy in our database was carefully checked and unique identifiers from other major databases are assigned to each taxonomical entry, it can easily be linked to other international databases. Besides the data table format, the bibliographical information is available both in a human-readable citation format (APA style), and as a BibTex entry that facilitates several ways of usage. The database will also be maintained in the future, and will be made available for registered users to upload new data through a web-based graphical interface. New data will be curated, which will allow the database to be regularly expanded. Moreover, biotic interactions not found on GloBi's server will be uploaded to that database as well.

Code availability

Computer codes used to generate and clean the Carabidae – Fungi database, as well as those to produce plots in the article and the Shiny app are freely available at <https://github.com/pozsgai/CaraFun>.

Received: 14 January 2021; Accepted: 8 October 2021;

Published: 4 November 2021

References

- Foley, J. A. *et al.* Global consequences of land use. *Science* **309**, 570–574 (2005).
- Chase, J. M., Blowes, S. A., Knight, T. M., Gerstner, K. & May, F. Ecosystem decay exacerbates biodiversity loss with habitat loss. *Nature* **584**, 238–243 (2020).
- Turvey, S. T. & Crees, J. J. Extinction in the Anthropocene. *Curr. Biol.* **29**, R982–R986 (2019).
- Cardoso, P. *et al.* Scientists' warning to humanity on insect extinctions. *Biol. Conserv.* **242**, 108426 (2020).
- Ceballos, G., Ehrlich, P. R. & Dirzo, R. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc. Natl. Acad. Sci. USA* **114**, E6089–E6096 (2017).
- Hooper, D. U. *et al.* Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol. Monogr.* **75**, 3–35 (2005).
- Chapin, F. S. *et al.* Consequences of changing biodiversity. *Nature* **405**, 234–42 (2000).
- Costanza, R. *et al.* Changes in the global value of ecosystem services. *Glob. Environ. Change* **26**, 152–158 (2014).
- Gray, C. *et al.* Ecological networks: The missing links in biomonitoring science. *J. Appl. Ecol.* **51**, 1444–1449 (2014).
- Valiente-Banuet, A. *et al.* Beyond species loss: The extinction of ecological interactions in a changing world. *Funct. Ecol.* **29**, 299–307 (2015).
- Traill, L. W., Lim, M. L. M., Sodhi, N. S. & Bradshaw, C. J. A. Mechanisms driving change: altered species interactions and ecosystem function through global warming. *J. Anim. Ecol.* **79**, 937–47 (2010).
- Lövei, G. L. & Sunderland, K. D. Ecology and behavior of ground beetles. *Annu. Rev. Entomol.* **41**, 231–256 (1996).
- Luff, M. L. Biology of polyphagous ground beetles in agriculture. *Agric. Zool. Rev.* **2**, 237–278 (1987).
- Symondson, W. O. C., Sunderland, K. D. & Greenstone, M. H. Can generalist predators be effective biocontrol agents? *Annu. Rev. Entomol.* **47**, 561–594 (2002).
- Honek, A., Martinkova, Z. & Jarosik, V. Ground beetles (Carabidae) as seed predators. *Eur. J. Entomol.* **100**, 531–544 (2003).
- Peay, K. G., Kennedy, P. G. & Talbot, J. M. Dimensions of biodiversity in the Earth mycobiome. *Nat. Rev. Microbiol.* **14**, 434–447 (2016).
- Tedersoo, L., Bahram, M. & Zobel, M. How mycorrhizal associations drive plant population and community biology. *Science* **367**, (2020).
- Lustenhauer, N. *et al.* A trait-based understanding of wood decomposition by fungi. *Proc. Natl. Acad. Sci. USA* **117**, 11551–11558 (2020).
- Fisher, M. C. *et al.* Emerging fungal threats to animal, plant and ecosystem health. *Nature* **484**, 186–194 (2012).
- Vilcinskas, A. Evolutionary ecology of parasitic fungi and their host insects. *Fung. Ecol.* **38**, 12–20 (2019).
- Thiele, H. U. *Carabid Beetles in Their Environments: A Study on Habitat Selection by Adaptations in Physiology and Behaviour.* (Springer, 1977).
- Traugott, M., Weissteiner, S. & Strasser, H. Effects of the entomopathogenic fungus *Beauveria brongniartii* on the non-target predator *Poecilus versicolor* (Coleoptera: Carabidae). *Biol. Control* **33**, 107–112 (2005).
- Haelewaters, D., Blackwell, M. & Pfister, D. H. Laboulbeniomyces: Intimate fungal associates of arthropods. *Annu. Rev. Entomol.* **66**, annurev-ento-013020–013553 (2021).
- Brodie, J. F. *et al.* Secondary extinctions of biodiversity. *Trends in Ecol. Evol.* **29**, 664–672 (2014).
- Evans, D. M. & Kitson, J. J. Molecular ecology as a tool for understanding pollination and other plant–insect interactions. *Curr. Opin. Insec. Sci.* **38**, 26–33 (2020).
- Deroches, S. A. P. *et al.* Biomonitoring for the 21st Century: Integrating next-generation sequencing into ecological network analysis. *Adv. Ecol. Res.* **52**, 1–62 (2018).
- Sow, A., Haran, J., Benoit, L., Galan, M. & Brévault, T. DNA Metabarcoding as a tool for disentangling food webs in agroecosystems. *Insects* **11**, 294 (2020).
- Kudo, R., Masuya, H., Endoh, R., Kikuchi, T. & Ikeda, H. Gut bacterial and fungal communities in ground-dwelling beetles are associated with host food habit and habitat. *ISME J* **13**, 676–685 (2019).
- Poelen, J. H., Simons, J. D. & Mungall, C. J. Global Biotic Interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecol. Inform.* **24**, 148–159 (2014).
- Pozsgai, G. *et al.* Carabid – fungal associations. *figshare* <https://doi.org/10.6084/m9.figshare.14602479.v5> (2021).
- Chang, W., Cheng, J., Allaire, J., Xie, Y. & McPherson, J. *shiny: Web application framework for r.* <https://CRAN.R-project.org/package=shiny> (2020).

Acknowledgements

G.P., I.B.F., J.Z. and M.Y. were supported by a grant of the “111 project” in China. G.P. was supported by a postdoctoral fellowship by the State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, and M.R.M. by the Statutory Funds of Institute for Agricultural and Forest Environment of Polish Academy of Sciences.

Author contributions

G.P. and I.B.F. contributed equally in work and share first authorship. The idea was conceived by G.P. and finalized with the help of I.B.F. The data collection procedure, and data representation were settled by G.P. and I.B.F. The data collection flowchart was drawn by I.B.F. and edited by G.L.L. Database structure was designed by G.P. R codes facilitating data collection and data cleaning were written by G.P. and M.V.K. Each of the contributors participated in data collection, S.B. and D.F. provided unpublished original data. Interactive visualisation of data was programmed by M.V.K. and G.P. Carabidae taxonomy was validated by G.P., S.B., and G.L.L., fungal taxonomy was validated by I.B.F., W.P.P., and N.V.M. The first draft of the text was written by G.P., then edited by I.B.F. and G.L. The final version of the manuscript was edited by G.L.L., C.R. and M.Y.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.P., C.R. or M.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2021, corrected publication 2022