

REPORT

MedGermDB: A seed germination database for characteristic species of Mediterranean habitats

Diana María Cruz-Tejada¹  | Eduardo Fernández-Pascual²  | Alessio Mo¹  |
Efisio Mattana³  | Angelino Carta^{1,4} 

¹Department of Biology, Botany Unit, University of Pisa, Pisa, Italy

²IMIB Biodiversity Research Institute (University of Oviedo—CSIC—Principality of Asturias), University of Oviedo, Mieres, Spain

³Royal Botanic Gardens, Kew, Wakehurst, West Sussex, UK

⁴CIRSEC—Centre for Climate Change Impact, University of Pisa, Pisa, Italy

Correspondence

Diana María Cruz-Tejada, Department of Biology, Botany Unit, University of Pisa, Via Derna 1, I-56126, Pisa, Italy.
Email: diana.cruztejada@phd.unipi.it

Funding information

Ministero dell'Università e della Ricerca

Co-ordinating Editor: Jodi N. Price

Abstract

Seed germination is a crucial phase of plant responses in early life to current and future environmental conditions. However, germination data are still scarce or disaggregated for many plant lineages and regions, including global biodiversity hotspots such as the Mediterranean Basin. We present MedGermDB, the first germination database for characteristic species of Mediterranean habitats, as defined by the EUNIS classification. We also present a systematic approach to build germination databases using automatic and semi-automatic data extraction from the literature. MedGermDB contains germination data for 4680 laboratory tests performed with 236 angiosperm species from 43 families, extracted from 125 literature sources (2837 sources screened). Each test is associated to a seed lot (i.e., a seed collection of a plant species obtained from a specific location at a specific time) and its metadata, recording geographical information and experimental conditions (storage, dormancy-breaking treatments, incubation temperature, and photoperiod). MedGermDB is available as a csv file, and through a web app: <https://dianamariacruztejada.shinyapps.io/medgermdb/>. MedGermDB can be used to explore eco-evolutionary questions and provides a backbone data set for informing effective seed-based conservation and ecological restoration activities targeting EUNIS habitats. Our methodological approach to data extraction can be extended to other study systems, contributing to global efforts to mobilize germination data.

KEYWORDS

Boolean string, climate change, data extraction tools, macroecology, macroevolution, primary data, regeneration, restoration, seed germination, seed traits

1 | INTRODUCTION

The time and place of seed germination determine the success of the transition from seed to seedling, thereby influencing adult plant trait expression (Donohue et al., 2010), the coexistence of species in plant communities (Grubb, 1977; Baskin & Baskin, 2014) and the distribution of species across biogeographical regions (Bykova et al., 2012).

Germination is an eco-physiological process starting with the imbibition of the seeds and terminating with the emergence of the radicle (Bewley et al., 2013). It is controlled by numerous interacting environmental cues, including temperature, light and dormancy-cycling signals (Baskin & Baskin, 2014). The combination of environmental cues under which the germination of a given species occurs constitutes its 'seed germination niche' which is assumed to maximize the

chance of successful plant regeneration (Grubb, 1977). However, describing the diversity of seed germination niches across plant species and environments (e.g., the seed germination spectrum; Saatkamp et al., 2019) requires high-quality data for a wide range of species (Fernández-Pascual et al., 2021) and the application of challenging analytical techniques (Carta et al., 2022).

Several studies have emphasized the scarcity of germination databases with broad geographical coverage as a significant limitation to fully understand seed germination responses at different phylogenetic and geographic scales (Saatkamp et al., 2019; Visscher et al., 2022). The book by Baskin and Baskin (2014) is a fundamental example of data compilation which provides extensive summary results of germination experiments from published studies at global level, with a sound biogeographical background. However, it does not include detailed information on experimental setup and results. At the same time, online databases containing plant traits (e.g., TRY, <https://www.try-db.org>; Kattge et al., 2020) generally have low coverage of regenerative traits (Jiménez-Alfaro et al., 2016), highlighting the need to compile detailed primary information on seed germination data. Despite recent advances and calls for compiling germination data (Fernández-Pascual, 2021; Ordóñez-Parra et al., 2022; Fernández-Pascual et al., 2023), databases still have very limited coverage of many regions and taxa. Such coverage disparities can have profound negative impacts on the predictive power of macroecological research (Visscher et al., 2022).

This is particularly the case for global hotspots of plant diversity, such as the Mediterranean Basin (Myers et al., 2000), where historical and recent changes in climate, land use and fire regimes create challenges for plant conservation and ecological restoration activities (Mattana, Carta, et al., 2022). Addressing these challenges requires knowledge of functional seed traits across this region (Tavşanoğlu & Pausas, 2018). However, while many Mediterranean countries have a long tradition of seed ecological research (e.g., Herranz et al., 1998; Thanos et al., 1991; Thanos et al., 1989), data syntheses of germination traits at large scales are yet to be attempted. Such synthesis is critical to improve wild species use in seed-based conservation and restoration activities (Merritt & Dixon, 2011). A lack of information on seed germination creates a bottleneck that greatly limits species availability for restoration actions (Ladouceur et al., 2018). When such information is available, large-scale ecological restoration activities can be planned at landscape scales (Jiménez-Alfaro et al., 2020).

Thus, to foster the applied use of native seeds in conservation activities as defined in the European biodiversity policies (European Commission, 1992), restoration planning must take into account germination requirements to inform species selection, to ensure effective pre-sowing seed management and to maximize seedling establishment efficiency (Kildisheva et al., 2020). To this end, we present MedGermDB, the first database of germination records for characteristic species of Mediterranean habitats based on the European habitat classification system (EUNIS: Chytrý et al., 2020). In doing so, we set up a novel systematic approach which allows to: (1) automatically identify online literature sources with relevant

information; (2) use semi-automatic tools to extract available data; and (3) compile curated data sets. Our approach can be applied to other geographic regions and plant lineages, contributing to global efforts to aggregate germination data (Saatkamp et al., 2019; Fernández-Pascual et al., 2023). MedGermDB can be used to address eco-evolutionary and biogeographical questions and will serve as a backbone data set for informing effective seed-based conservation (e.g., seed banking and plant propagation) and restoration (e.g., natural ecosystems restoration using native seed mixes) activities (Mattana, Ulian, & Pritchard, 2022).

2 | METHODS

We applied a systematic database-building process, composed of six steps to extract germination data from literature sources. The extracted information consisted in detailed germination results uniquely associated to georeferenced populations and experimental conditions. The whole process is based on automatic and semi-automatic tools (Hupe, 2019; R Core Team, 2022) to maximize reproducibility (Figure 1). The workflow is described in a PRISMA flow chart (Appendix S1). Future updates can be automatically uploaded through the GitHub project (<https://github.com/DianaCruzTejada/MedGermDB>) allowing to track new data on seed germination in the Mediterranean.

2.1 | Boolean string construction

We built four alternative Boolean strings to maximize the probability of retrieving results (Appendix S2). All Boolean strings were built based on three types of keywords: one referring to seeds (e.g., seed vs. seed*); one to the function (e.g., germination vs. dormancy); and the last one referring to the geographical distribution (e.g., Mediterranean Basin countries).

2.2 | Literature search

We searched for published scientific articles in the Web of Science (WoS) and Scopus databases using the four Boolean strings described above and available in Appendix S2. The retrieved results were imported into EndNote X9 (Hupe, 2019) to automatically eliminate duplicates, returning 6759 references. We accessed the full text through two main channels: (a) the "Find Full Text" functionality in EndNote and (b) online access to the journals using the institutional subscriptions of the University of Pisa. We obtained full texts for a total of 2837 references (Figure 1a), which were processed in subsequent steps in the workflow (see subsection 2.4. Species names automated scanning in full text and Appendix S1). For the references for which full texts were not available (3958), we examined the titles, year, and journal of publication in order to make sure we were not losing important information. The great majority of them

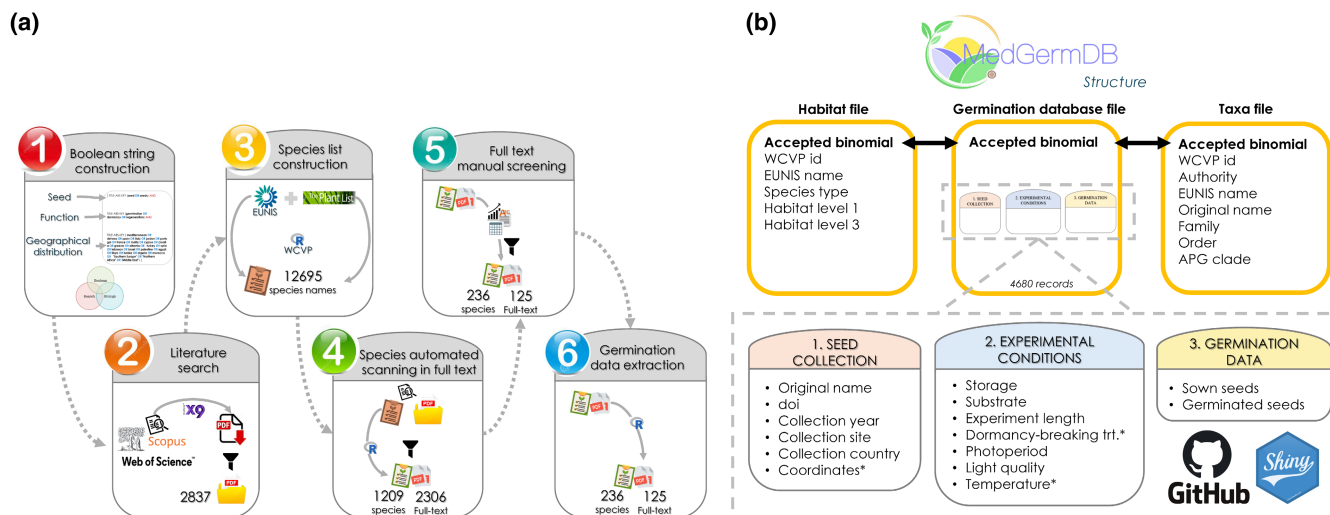


FIGURE 1 Key steps in the systematic process toward building MedGermDB. (a) Workflow for source identification and data extraction: steps 1–2 indicate an automatic process using EndNote and R tools (*Taxonstand*, *pdfsearch*); steps 3–4 indicate a semi-automatic process to screen and extract data using *Tabulizer* and *juicer*. The numbers in black represent the retrieved results at each step (i.e., full-text source and species). (b) Structure of MedGermDB: yellow boxes are the three files that make up the complete database (the *Germination database file*, and the supplementary *Taxa* and *Habitat* files, including the field names of each variable); black thick arrows indicate the link between files using the field “Accepted binomial” as unique identifier. *, “Coordinates” includes four columns: latitude, longitude, altitude, and accuracy; “Dormancy-breaking trt.” includes four columns: scarification, stratification, other treatments, water potential; “Temperature” includes two columns: T_{day}, T_{night}, (see Appendix S4).

(89%) were considered uninformative because they either referred to cultivated crops, to seed collections outside the Mediterranean Basin or did not refer to seed germination (see below definition of seed germination).

2.3 | Species list construction

MedGermDB was generated using a list of characteristic species of Mediterranean habitats based on the EUNIS expert system (Chytrý et al., 2020). This three-level hierarchical classification of European habitats (Moss, 2008) is based on plant species composition and geographic location, which provide potential information for future uses in European nature conservation planning, monitoring, and assessment. As such, the list comprises species representative of the whole Mediterranean region, its habitat variability, and associated ecological processes. This method of species selection allows for maximizing species and ecological diversity to generate a knowledge-species pool for potential seed-based conservation and restoration activities (Ladouceur et al., 2018). We selected all EUNIS level-3 habitats that occur in the Mediterranean Basin (54 habitats) as listed in the EUNIS expert system. We also limited the list to flowering plants (angiosperms) representing the most diverse and dominating monophyletic clade in our system (Sauquet & Magallón, 2018). For the selected habitats, we gathered a list of 1516 species considered by EUNIS to be dominant, constant, or diagnostic of each selected habitat (Appendix S3). To improve matching species names when looking for plant names to be searched in the literature sources (see below), we generated

a list of taxonomic synonyms of the EUNIS original names using the *Taxonstand* R package (Cayuela et al., 2021). This list of synonyms contained 11,179 species names. Then, the original EUNIS names were standardized against The World Checklist of Vascular Plants, WCVP v10 (Govaerts, 2022) using the *U.Taxonstand* package (Zhang & Qian, 2022). From this standardization procedure, we gathered a final list of 12,695 alternative names to be searched in the literature sources.

2.4 | Species names automated scanning in full text

We scanned for the species' alternative names in all downloaded full texts, using the *pdf search* package (LeBeau, 2018) in R. This package allows for keyword searching within a single PDF or a directory of PDF files. We obtained a list of 1209 species and 2306 full texts (Figure 1a).

2.5 | Full-text manual screening

The purpose of this step was to screen the full texts containing species names to determine if they contained quality germination data for inclusion in MedGermDB. To ensure data quality and standardization, we only accepted literature sources reporting laboratory experiments on seed germination. We defined seed germination as radicle protrusion (Baskin & Baskin, 2014) and excluded those sources in which the experiments were related to seedling emergence, a functional trait that substantially differs in ecophysiological

requirements from radicle protrusion. We also discarded those sources where the experimental settings and/or final germination proportions were not fully described or not entirely experimentally controlled. For instance, we discarded greenhouse and field experiments as these, beside not showing results under controlled laboratory conditions, generally provide information on seedling emergence and not on radicle protrusion (Baskin et al., 2006). We also excluded those sources in which seeds of Mediterranean species had been collected in populations outside the Mediterranean Basin (Médail & Quezel, 1999).

At the end of the screening, we identified 125 articles (6% out of 2306) with germination data (Figure 1a; step 5), for which we annotated the exact position in the main text or item (table, figure, references, or supplementary material) in which the data were available. The majority of excluded sources referred to non-standardized laboratory experiments (35%) (e.g., greenhouse or field experiments) or/and to experiments reporting seedling emergence (15%). Also, we excluded experiments conducted with seeds from outside the Mediterranean Basin (13%). On average, the screening of one source took 2 min.

2.6 | Germination data extraction

From the 125 articles previously detected, 4680 primary germination records were extracted by applying semi-automatic tools. For each extracted germination record, we kept the information of which original study supplied the record, by providing a doi reference in the database (see database structure, Figure 1b and Appendix S4). We used the *tabulizer* package (Leeper, 2018) to extract data from tables and, in the case of the data shown in figures, we used the *juicer* package (Lajeunesse, 2021) in R. All data were stored by entering information from the original source into spreadsheets, in which each row represented a unique combination of germination test conditions and each column a single informative variable (Kissling et al., 2019). On average, data extraction for one article took 15 min.

3 | DATABASE PRESENTATION

3.1 | Database structure

The database is structured in three comma-separated text files (.csv), provided in Appendixes S8, S9, S10 and in GitHub (<https://github.com/DianaCruzTejada/MedGermDB>) (Figure 1b). The actual germination data set is in the “*Germination database File*”, which includes three main blocks of germination information, providing the germination experiments recorded (Figure 2a): the first block corresponds to the seed lot collection (e.g., original taxon name, collection year, collection country, and geographical coordinates; if the latter were not provided in the source, we looked for the closest available toponym and we specified the accuracy of the site of seed collection); the second block reports information about

the experimental conditions (e.g., storage, substrate, experiment length, dormancy-breaking treatments, incubation temperature, photoperiod); and the third block represent the final germination data (number of sown and germinated seeds) (Appendix S4). The supplementary “*Taxa File*” shows the taxonomic information for each species presented in the *Germination database File*, listing the accepted binomial that resulted from the standardization of the original EUNIS taxon names with WCVP v.10 (Govaerts, 2022), the authority, and the original taxon name found in the full-text source (Appendix S5). Finally, the supplementary “*Habitat File*” contains the species list included in the *Germination database File* and the EUNIS habitats where they are present (Appendix S6). These three files are linked by the accepted binomial acting as a unique identifier (Figure 1b); R codes to properly join the three files are provided in the GitHub repository.

3.2 | Database contents

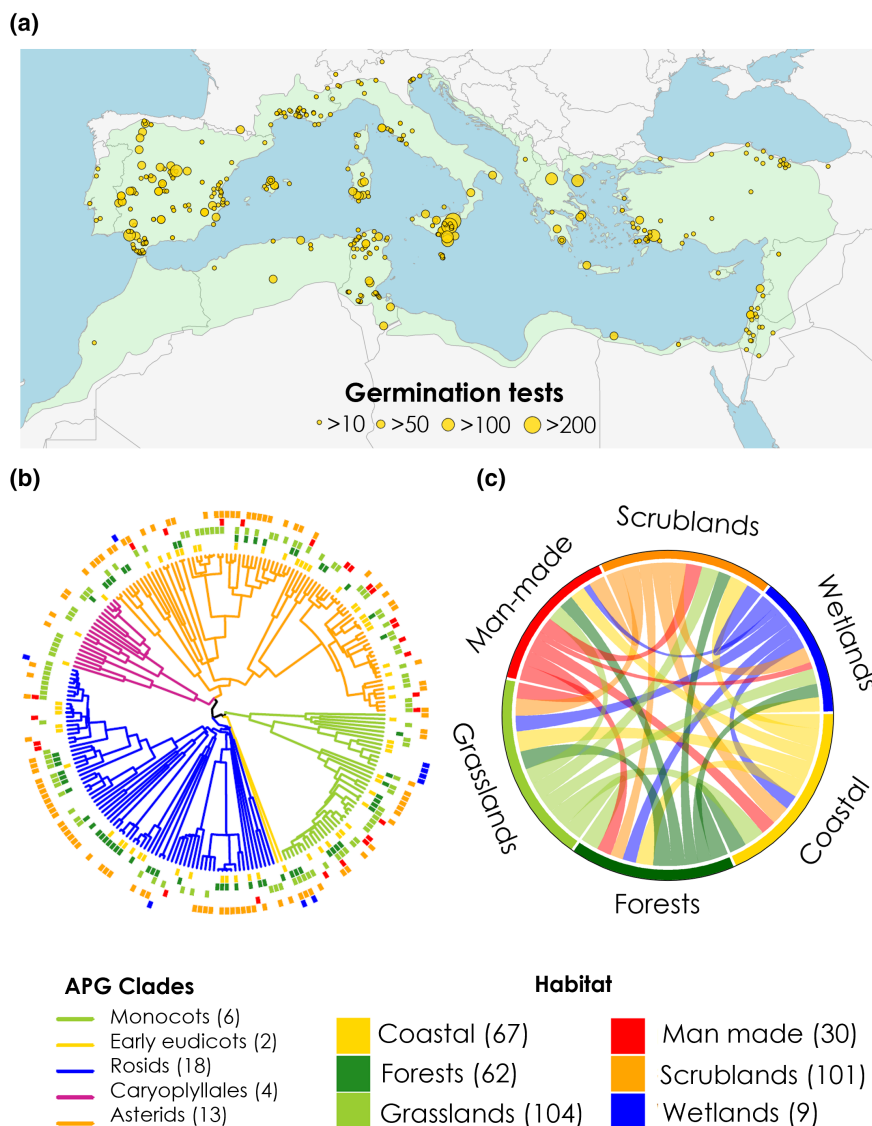
The germination data set (*Germination database File*) contains 4680 records (i.e., germination tests) and 236 angiosperm species (Figure 1a) retrieved from 125 articles (Appendix S1). Seeds were collected across the whole Mediterranean Basin (Figure 2a), with Spain (1951 records), Italy (1276) and Turkey (339) being the top three contributing countries. The species are distributed in all main APG clades (Figure 2b) and all six main habitat groups of the EUNIS system (i.e., Level-1 EUNIS codes, Figure 2c), corresponding to 17% of the total EUNIS Mediterranean species. In terms of quantity of records, these results are fairly comparable with other germination databases that include information from a particular region or ecosystem, reporting 4012 records with 334 species for temperate deciduous forests (Eduardo Fernández-Pascual, 2021) and 10,187 records with 281 taxa for Brazilian rock outcrop vegetation (Ordóñez-Parra et al., 2022).

3.3 | Germination experimental conditions

The total estimate of seeds used in the experiments was 490,958. The storage time varies from 1 to 3650 days: 1380 records were conducted using *fresh* seeds (storage less than one month from collecting); 3017 records with *stored* seeds (storage more than one month from collecting or records for which no storage information was provided but for which we could assume they experienced some degree of storage before testing); 313 records with *banked* seeds (storage more than one month from collecting at temperatures below 0°C). The experiments were performed with untreated seeds in 1284 records; 1348 records used *scarified* seeds; 701 records *stratified* seeds (526 and 175 records for cold and warm stratification, respectively) and 2048 records used seeds with *other treatments*. Concerning the photoperiod, *light* (diurnal alternating light) was used in 3264 records; *darkness* in 1292; and 124 did not provide information on this parameter. The range of germination temperatures (average of the daily

FIGURE 2 Database contents

(a) Geographical distribution of the germination collections included in the database, with size of circles associated with number of records (i.e., germination tests). Data representativeness from the (b) phylogenetic and (c) habitat perspectives. The phylogenetic tree in (b) represents a wide phylogenetic dispersion, including the most representative clades of the angiosperms and it was generated based on the reference megatree (Smith & Brown, 2018) and plotted with the *ggtree* package (Yu et al., 2017). The squares at the tips of the tree represent the different habitats in which each species of the tree is found and the numbers in parentheses indicate the families in each clade. The chord diagram in (c) shows that the species with germination data are well-distributed across the Level-1 EUNIS Mediterranean habitats. The different lines that join the habitats show the bidirectional relationship between the habitats, representing the shared species and indicating the presence of the same species in more than one habitat. The numbers in parentheses indicate the number of species per habitat and this is reflected in the external portion of the circle.



thermoperiod) went from 3°C to 40°C, with 3142 records of *constant* temperature regimes; and 1538 *alternating* temperature regimes.

3.4 | Database representativeness

To evaluate the representativeness of the database, we considered the phylogeny (family and APG clades; APG IV, 2016) and the habitat types.

In total, 43 families are represented in the database, of which Fabaceae, Cistaceae, Asteraceae, and Poaceae are the top four most studied ones with 1059, 700, 439, and 347 records, respectively. All these families are the most diverse in the original EUNIS species list (Asteraceae, 207 species; Poaceae, 181; Fabaceae, 154), except the Cistaceae, well represented in our data set because of their seed germination ecology (hard seeds and fire-related responses; Luna et al., 2019; Pausas & Lamont, 2022; Thanos et al., 1992). We compiled the phylogenetic tree for all species included in the database using the R package *V.PhylMaker* (Jin & Qian, 2019) and plotted this information with the *ggtree*

package (Yu et al., 2017). The binary phylogenetic signal ($D=0.88$, Orme et al., 2023) of the species included in our database vs all Mediterranean EUNIS species, indicates a wide phylogenetic dispersion (Figure 2b). This result suggests that the species for which germination data are available represent a random subsample of the original list, with no tendencies for a specific group to be under- or overrepresented.

We built a Euclidean distance matrix including the Level-1 EUNIS habitats with the total of the EUNIS Mediterranean species and another distance matrix with only the species included in the database and their habitats. Mantel test suggests a high correlation between both matrixes ($p<0.05$, $r=0.82$), indicating that MedGermDB is a representative sample of Mediterranean species at the habitat level (Figure 2c).

4 | CONCLUSIONS

MedGermDB, the first database of germination data for the Mediterranean Basin, will contribute to improve our knowledge of



the seed germination niche in this biodiversity hotspot. Its wide phylogenetic and ecological representativeness constitute a solid platform for the development of plant ecological, evolutionary and conservation studies. By following the EUNIS classification of species and habitats, it provides a species pool with seed quality data with the potential to inform European policies and ecological restoration activities (De Vitis et al., 2017; Pedrini et al., 2020) in the region. The new systematic approach presented here can be transferred to other geographic regions and taxonomic lineages, contributing to global efforts to aggregate germination data (Saatkamp et al., 2019; Fernández-Pascual et al., 2023), and allowing testing long-standing ecological, evolutionary and biogeographical hypotheses (e.g. Grubb, 1977; Baskin & Baskin, 2014). Particularly, we expect to quantitatively test the main cold-cued germination syndrome of the Mediterranean flora (Thanos et al., 1989; Picciau et al., 2019; Carta et al., 2022) and identify germination optima across species and habitats. We also envisage the use of this kind of data to forecast plant responses to current and future climatic conditions (Sentinella et al., 2020; Mattana et al., 2023).

AUTHOR CONTRIBUTIONS

Angelino Carta and Diana María Cruz-Tejada conceived the study. Eduardo Fernández-Pascual and Efigio Mattana contributed to the study design. Diana María Cruz-Tejada, Alessio Mo, and Angelino Carta contributed to the acquisition of data. Diana María Cruz-Tejada and Angelino Carta wrote the manuscript. All co-authors edited the manuscript.

ACKNOWLEDGEMENTS

We thank Isabella Pucci, Suada Pastacaldi, Livia Gragnani, Federico Pasquini, and Stefania Bottega for their help in data screening. We also thank Gabriele Midolo for his helpful comments.

FUNDING INFORMATION

Ministry of University and Research (MUR) as part of the PON 2014–2020 “Research and Innovation” resources—Green/Innovation Action—DM MUR 1061/2022, Number: DOT13GFICX-2.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

All data are available as supplementary materials. The data and codes to join the database files are stored at <https://github.com/DianaCruzTejada/MedGermDB> and visualized with the shiny app at <https://diana.mariacruztejada.shinyapps.io/medgermdb/>. A version of record of the repository can be found at <https://doi.org/10.5281/zenodo.10915154>. All people interested in contributing to the growth of this germination database are encouraged to contact the corresponding author.

ORCID

Diana María Cruz-Tejada <https://orcid.org/0000-0003-3220-1619>

Eduardo Fernández-Pascual <https://orcid.org/0000-0002-4743-9577>

Alessio Mo <https://orcid.org/0000-0003-3129-1747>

Efigio Mattana <https://orcid.org/0000-0001-6235-4603>

Angelino Carta <https://orcid.org/0000-0001-8437-818X>

REFERENCES

- APG IV. (2016) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, 181(1), 1–20. Available from: <https://doi.org/10.1111/boj.12385>
- Baskin, C.C. & Baskin, J.M. (2014) *Seeds: Ecology, biogeography, and, evolution of dormancy and germination*. San Diego, CA: Elsevier Inc. Available from: <https://doi.org/10.1016/C2013-0-00597-X>
- Baskin, C.C., Thompson, K. & Baskin, J.M. (2006) Mistakes in germination ecology and how to avoid them. *Seed Science Research*, 16(3), 165–168. Available from: <https://doi.org/10.1079/ssr2006247>
- Bewley, J.D., Bradford, K.J., Hilhorst, H.W.M. & Nonogaki, H. (2013) *Seeds: physiology of development, germination and dormancy*, 3rd edition. (Vol. 9781461446934). New York: Springer. Available from: <https://doi.org/10.1007/978-1-4614-4693-4/COVER>
- Bykova, O., Chuine, I., Morin, X. & Higgins, S.I. (2012) Temperature dependence of the reproduction niche and its relevance for plant species distributions. *Journal of Biogeography*, 39(12), 2191–2200. Available from: <https://doi.org/10.1111/j.1365-2699.2012.02764.x>
- Carta, A., Fernández-Pascual, E., Gioria, M., Müller, J.V., Riviére, S., Rosbakh, S. et al. (2022) Climate shapes the seed germination niche of temperate flowering plants: a meta-analysis of European seed conservation data. *Annals of Botany*, 129, 775–786. Available from: <https://doi.org/10.1093/aob/mcac037>
- Cayuela, A.L., Macarro, I., Stein, A. & Oksanen, J. (2021) Taxonstand: Taxonomic Standardization of Plant Species Names. *R Package Version 2.4*, 1–11. <https://cran.r-project.org/package=Taxonstand>
- Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M., Rodwell, J.S. et al. (2020) EUNIS habitat classification: expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, 23(4), 648–675. Available from: <https://doi.org/10.1111/avsc.12519>
- De Vitis, M., Abbandonato, H., Dixon, K.W., Laverack, G., Bonomi, C. & Pedrini, S. (2017) The European native seed industry: characterization and perspectives in grassland restoration. *Sustainability (Switzerland)*, 9(10), 1–14. Available from: <https://doi.org/10.3390/su9101682>
- Donohue, K., Rubio de Casas, R., Burghardt, L., Kovach, K. & Willis, C.G. (2010) Germination, Postgermination adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 293–319. Available from: <https://doi.org/10.1146/annurev-ecolsys-102.209-144.715>
- European Commission. (1992) Council directive 92/43/EEC of 21 of may 1992 on the conservation of natural habitats.
- Fernández-Pascual, E., Carta, A., Rosbakh, S., Guja, L., Phartyal, S., Silveira, F. et al. (2023) SeedArc, a global archive of primary seed germination data. *New Phytologist*, 240, 466–470. Available from: <https://doi.org/10.1111/nph.19143>
- Fernández-Pascual, E. (2021) SylvanSeeds, a seed germination database for temperate deciduous forests. *Journal of Vegetation Science*, 32(1), e12960. Available from: <https://doi.org/10.1111/jvs.12960>
- Fernández-Pascual, E., Carta, A., Mondoni, A., Cavieres, L.A., Rosbakh, S., Venn, S. et al. (2021) The seed germination spectrum of alpine plants: a global meta-analysis. *New Phytologist*, 229(6), 3573–3586. Available from: <https://doi.org/10.1111/nph.17086>
- Govaerts, R. (2022) The world checklist of vascular plants (WCVP) (10.0). <http://sftp.kew.org/pub/data-repositories/WCVP/>

- Grubb, P. (1977) The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews*, 52(1), 107–145. Available from: <https://doi.org/10.1111/j.1469-185x.1977.tb01347.x>
- Herranz, J.M., Ferrandis, P. & Martínez-Sánchez, J.J. (1998) Influence of heat on seed germination of seven Mediterranean Leguminosae species. *Plant Ecology*, 136(1), 95–103. Available from: <https://doi.org/10.1023/A:1009702318641>
- Hupe, M. (2019) EndNote X9. *Journal of Electronic Resources In Medical Libraries*, 16(3–4), 117–119. Available from: <https://doi.org/10.1080/15424065.2019.1691963>
- Jiménez-Alfaro, B., Frischie, S., Stolz, J. & Gálvez-Ramírez, C. (2020) Native plants for greening Mediterranean agroecosystems. *Nature Plants*, 6(3), 209–214. Available from: <https://doi.org/10.1038/s41477-020-0617-3>
- Jiménez-Alfaro, B., Silveira, F.A.O., Fidelis, A., Poschlod, P. & Commander, L.E. (2016) Seed germination traits can contribute better to plant community ecology. *Journal of Vegetation Science*, 27(3), 637–645. Available from: <https://doi.org/10.1111/jvs.12375>
- Jin, Y. & Qian, H. (2019) V. PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography*, 42(8), 1353–1359.
- Kattge, J., Bönsch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P. et al. (2020) TRY plant trait database—enhanced coverage and open access. *Global Change Biology*, 26(1), 119–188. Available from: <https://doi.org/10.1111/gcb.14904>
- Kildisheva, O.A., Dixon, K.W., Silveira, F.A.O., Chapman, T., Di Sacco, A., Mondoni, A. et al. (2020) Dormancy and germination: making every seed count in restoration. *Restoration Ecology*, 28(S3), S256–S265. Available from: <https://doi.org/10.1111/rec.13140>
- Kissling, W.D., Balslev, H., Baker, W.J., Dransfield, J., Gödel, B., Lim, J.Y. et al. (2019) PalmTraits 1.0, a species-level functional trait database of palms worldwide. *Scientific Data*, 6(1), 1–13. Available from: <https://doi.org/10.1038/s41597-019-0189-0>
- Ladouceur, E., Jiménez-Alfaro, B., Marin, M., De Vitis, M., Abbandonato, H., Iannetta, P.P.M. et al. (2018) Native seed supply and the restoration species Pool. *Conservation Letters*, 11(2), 1–9. Available from: <https://doi.org/10.1111/conl.12381>
- Lajeunesse, M.J. (2021) Juicer: automated and manual extraction of numerical data from scientific images. *R Package Version 0.1*, 1–8. <https://cran.r-project.org/package=juicer>
- LeBeau, B. (2018) Pdfsearch: search tools for PDF files. *Journal of Open Source Software*, 3, 668. Available from: <https://doi.org/10.21105/joss.00668>
- Leeper, T.J. (2018) Tabulizer: bindings for tabula PDF table extractor library. *R Package Version 0.2.2*.
- Luna, B., Chamorro, D. & Pérez, B. (2019) Effect of heat on seed germination and viability in species of Cistaceae. *Plant Ecology and Diversity*, 12(2), 151–158. Available from: <https://doi.org/10.1080/17550874.2019.161091>
- Mattana, E., Carta, A., Fernández-Pascual, E., Keeley, J.E. & Pritchard, H.W. (2022) Climate change and plant regeneration from seeds in Mediterranean regions of the northern hemisphere. In: *Plant regeneration from seeds*. Cambridge, Massachusetts: Academic Press, pp. 101–114. Available from: <https://doi.org/10.1016/B978-0-12-823731-1.00015-9>
- Mattana, E., Chapman, T., Miles, S., Ulian, T. & Carta, A. (2023) Regeneration from seeds in a temperate native flora: A climate-smart and natural-capital-driven germination risk modelling approach. <https://doi.org/10.1002/ppp3.10378>
- Mattana, E., Ulian, T. & Pritchard, H.W. (2022) Seeds as natural capital. *Trends in Plant Science*, 27(2), 139–146. Available from: <https://doi.org/10.1016/j.tplants.2021.08.008>
- Médail, F. & Quezel, P. (1999) Biodiversity hotspots in the Mediterranean. *Conservation Biology*, 13(6), 1510–1513.
- Merritt, D.J. & Dixon, K.W. (2011) Restoration seed banks—a matter of scale. *Science*, 332, 21–22.
- Moss, D. (2008) *EUNIS habitat classification—A guide for users*. European Topic Centre on Biological Diversity.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858. Available from: <https://doi.org/10.1080/21564574.1998.9650003>
- Ordóñez-Parra, C.A., Dayrell, R.L.C., Negreiros, D., Andrade, A.C.S., Andrade, L.G., Antonini, Y. et al. (2022) Rock n' Seeds: A Database of Seed Functional Traits and Germination Experiments from Brazilian Rock Outcrop Vegetation. *Ecology*, 104, e3852. Available from: <https://doi.org/10.1002/ecy.3852>
- Orme, D. (2023) The caper package: Comparative analysis of phylogenetics and evolution in R. pp. 1–36.
- Pausas, J.G. & Lamont, B.B. (2022) Fire-released seed dormancy – A global synthesis. *Biological Reviews*, 97(4), 1612–1639. Available from: <https://doi.org/10.1111/brv.12855>
- Pedrini, S., Balestrazzi, A., Madsen, M.D., Bhalsing, K., Hardegree, S.P., Dixon, K.W. et al. (2020) Seed enhancement: getting seeds restoration-ready. *Restoration Ecology*, 28(S3), S266–S275. Available from: <https://doi.org/10.1111/rec.13184>
- Picciau, R., Pritchard, H.W., Mattana, E. & Bacchetta, G. (2019) Thermal thresholds for seed germination in Mediterranean species are higher in mountain compared with lowland areas. *Seed Science Research*, 29(1), 44–54. Available from: <https://doi.org/10.1017/S0960258518000399>
- R Core Team. (2022) R: A Language and Environment for Statistical Computing.
- Saatkamp, A., Cochrane, A., Commander, L., Guja, L.K., Jimenez-Alfaro, B., Larson, J. et al. (2019) A research agenda for seed-trait functional ecology. *New Phytologist*, 221(4), 1764–1775. Available from: <https://doi.org/10.1111/nph.15502>
- Sauquet, H. & Magallón, S. (2018) Key questions and challenges in angiosperm macroevolution. *New Phytologist*, 219(4), 1170–1187. Available from: <https://doi.org/10.1111/nph.15104>
- Sentinella, A.T., Warton, D.I., Sherwin, W.B., Offord, C.A. & Moles, A.T. (2020) Tropical plants do not have narrower temperature tolerances, but are more at risk from warming because they are close to their upper thermal limits. *Global Ecology and Biogeography*, 29(8), 1387–1398. Available from: <https://doi.org/10.1111/geb.13117>
- Smith, S.A. & Brown, J.W. (2018) Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, 105(3), 302–314. Available from: <https://doi.org/10.1002/ajb2.1019>
- Tavşanoğlu, Ç. & Pausas, J.G. (2018) Data descriptor: A functional trait database for Mediterranean Basin plants. *Scientific Data*, 5, 1–18. Available from: <https://doi.org/10.1038/sdata.2018.135>
- Thanos, C.A., Georgiou, K., Douma, D.J., Christie, J., Thanos, C.A., Georgiou, K. et al. (1991) Photoinhibition of seed germination in Mediterranean maritime plants. *Annals of Botany*, 68(5), 469–475. <https://www.jstor.org/stable/42761648>
- Thanos, C.A., Georgiou, K., Kadis, C. & Pantazi, C. (1992) CISTACEAE: A plant family with hard seeds. *Israel Journal of Plant Sciences*, 41(4–6), 251–256. Available from: <https://doi.org/10.1080/0021213X.1992.10677232>
- Thanos, C.A., Georgiou, K. & Skarou, F. (1989) *Glaucium flavum* seed germination - an ecophysiological approach. *Annals of Botany*, 63(1), 121–130. Available from: <https://doi.org/10.1093/oxfordjournals.aob.a087714>
- Visscher, A.M., Vandelook, F., Fernández-Pascual, E., Pérez-Martínez, L.V., Ulian, T., Diazgranados, M. et al. (2022) Low availability of functional seed trait data from the tropics could negatively affect global macroecological studies, predictive models and plant conservation. *Annals of Botany*, 130(6), 773–784. Available from: <https://doi.org/10.1093/aob/mcac130>



- Yu, G., Smith, D.K., Zhu, H., Guan, Y. & Lam, T.T.Y. (2017) Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8(1), 28–36. Available from: <https://doi.org/10.1111/2041-210X.12628>
- Zhang, J. & Qian, H. (2022) U.Taxonstand: an R package for standardizing scientific names of plants and animals. *Plant Diversity*, 45(1), 1–5. Available from: <https://doi.org/10.1016/j.pld.2022.09.001>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. PRISMA flow chart (2020) with the results of the systematic approach retrieved via online literature sources.

Appendix S2. Source searches results from online databases using alternative Boolean strings considering different terms and their combinations in the title, abstract, or keywords.

Appendix S3. Species list generation from the EUNIS habitat classification system.

Appendix S4. Description of columns in the *Germination database File* including three blocks of the recorded germination information.

Appendix S5. Description of columns in the supplementary *Taxa File* listing the original EUNIS names and the corresponding taxonomical information.

Appendix S6. Description of columns in the supplementary *Habitat File* containing the species list included in the database and the EUNIS habitats where they are present.

Appendix S7. Bibliography in BibTeX format listing the relevant references included in MedGermDB.

Appendix S8. MedGermDB *Germination database File* in a single csv file.

Appendix S9. MedGermDB supplementary *Taxa File* in a single csv file.

Appendix S10. MedGermDB supplementary *Habitat File* in a single csv file.

How to cite this article: Cruz-Tejada, D.M., Fernández-Pascual, E., Mo, A., Mattana, E. & Carta, A. (2024) MedGermDB: A seed germination database for characteristic species of Mediterranean habitats. *Applied Vegetation Science*, 27, e12771. Available from: <https://doi.org/10.1111/avsc.12771>