



New Phytologist

SylvanSeeds, a database and web app to explore seed germination in temperate broadleaf and mixed forests

Journal:	<i>New Phytologist</i>
Manuscript ID	Draft
Manuscript Type:	MET - Methods Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Fernández-Pascual, Eduardo; Universidad de Oviedo, Departamento de Biología de Organismos y Sistemas
Key Words:	germination database, germination temperature, light and dark germination, seed dormancy, seed traits, temperate coniferous forests, temperate deciduous forests, temperate evergreen forests

SCHOLARONE™
Manuscripts

SylvanSeeds, a database and web app to explore seed germination in temperate broadleaf and mixed forests

Running title: SylvanSeeds, a forest germination database

Author: Eduardo Fernández-Pascual **Institutional address:** Departamento de Biología de Organismos y Sistemas, Universidad de Oviedo, C/ Catedrático Rodrigo Uría, 33006 Oviedo/Uviéu, Spain **Email:** efernandezpascual@gmail.com
Telephone: +34985104787

Total word count 2,577 **Introduction** 530 **Materials and Methods** 968 **Results** 848 **Discussion** 226 **Number of figures** 2 in colour **Tables** 0 **Supporting information** 6

Summary

- Seed traits have functional significance in all levels of plant biology, but there is a lack of germination databases of wide geographical scope. This article presents *SylvanSeeds*, a first global dataset of germination records for an ecologically coherent unit: Temperate Broadleaf and Mixed Forests.
- Data were gathered with a systematic literature search. A list of frequent taxa of the study area was created using 17,852 vegetation relevés from the sPlot database. The list was searched in the Web of Science. 6,791 references were screened, finding 614 articles from which data were extracted.
- *SylvanSeeds* includes 4,828 germination records of 364 species from 74 families (gymnosperms and angiosperms), collected in 46 countries between 1920-2017. It provides raw data for meta-analysis: proportions of seeds germinated in laboratory experiments of scarification, stratification, light-darkness, and constant-alternating temperatures. *SylvanSeeds* is freely

distributed as a .csv. The shiny app *SylvanSeeds* is also presented, to make data accessible to the public.

- *SylvanSeeds* advances functional seed ecology and brings two innovations to plant science. First, the data-gathering methodology can be extended to other biomes. Second, database and app can be a standard in further efforts to compile germination data.

Keywords: germination database, germination temperature, light and dark germination, seed dormancy, seed traits, temperate coniferous forests, temperate deciduous forests, temperate evergreen forests

Introduction

Seed traits have functional significance in all levels of plant biology, from population dynamics (Huang *et al.*, 2016) to community assembly (Larson & Funk, 2016) and species distributions (Larson & Funk, 2016). Recent discussions (Jiménez-Alfaro *et al.*, 2016; Saatkamp *et al.*, 2019) highlighted the lack of reliable germination databases of wide geographical scope as a major limitation to produce a global synthesis of the seed ecological spectrum. Seed germination is the first major developmental transition in the life of a plant, and it has profound effects on the plant's ecology (Donohue *et al.*, 2010). As such, the timing of this transition is controlled by a combination of environmental signals. Amongst these are temperature (Fernández-Pascual *et al.*, 2019), diurnal temperature alternation (Thompson *et al.*, 1977), light (Carta *et al.*, 2017) and seed dormancy inductors and relievers (Finch-Savage & Leubner-Metzger, 2006). All of these signals interact to produce a coarse- and fine-scale regulation of germination timing, integrating inputs from both seasonal climatic cycles (Jurado & Flores, 2005) and local environmental gradients (Fernández-Pascual *et al.*, 2017). The practical consequence of this is that the response of seeds to, for example, light will depend on other conditions set by the experimenter. This makes it difficult to summarise

germination “traits” into a single value in a way comparable to seed mass, specific leaf area or plant height (Pérez-Harguindeguy *et al.*, 2013), and may explain the lack of global germination databases. Germination compilations with a biogeographical background, of which the prime example today is the cornerstone book of Baskin and Baskin (2014), only provide summary information, for instance the interpreted optimal germination temperature instead of the proportion of seeds germinated across a temperature gradient.

To fill this gap, this article presents *SylvanSeeds*, a first global dataset of raw germination data for an ecologically and biogeographically coherent unit: the Temperate Broadleaf and Mixed Forests biome of the northern hemisphere (*TBMFNH*). The building of the dataset was part of a research project conducted within the *sPlot* initiative (Bruehlheide *et al.*, 2019), specifically *sPlot*’s project #12 (https://www.idiv.de/en/sdiv/working_groups/wg_pool/splot/projects.html). The goal of *sPlot* #12 was to study taxonomical and functional plant diversity in deciduous forests, which are dominant in temperate latitudes with mild to cold winters and relatively warm and wet summers. In the northern hemisphere, they also form a coherent evolutionary and biogeographic unit, sharing a common Tertiary history until they became separated between North America, Asia and Europe (Milne & Abbott, 2002). At the same time, a relatively high number of classic seed biology research groups have worked within the *TBMFNH* areas, e.g. in the USA (Baskin, 1995), the UK (Grime *et al.*, 1981), Belgium (Van Assche & Vandeloock, 2006) and Japan (Washitani, 1988). For this reason, *sPlot* project #12 was chosen to incorporate a biome-wide analysis of seed traits, as it was expected that the *TBMFNH* would provide a wide scope of available data. To gather data for the project, a systematic literature search (Koricheva *et al.*, 2013) was applied to a list of frequent species representing the flora of the *TBMFNH*. This data paper makes the database accessible to the research community. Additionally, the *SylvanSeeds* shiny web app is presented to facilitate the exploration of the data.

Materials and Methods

Species list

A list of species for which to search germination data was created using vegetation relevés (i.e. records of plant species co-occurring in sampling plots) from *sPlot* (Bruehlheide *et al.*, 2019). The terrestrial regionalization of the *World Wildlife Fund* (Olson *et al.*, 2001) was used to select plots according to their coordinates. To ensure that plots were a representative sample of *TBMFNNH* biodiversity, they were chosen from 20 ecoregions distributed throughout the northern hemisphere: 17 ecoregions of the Temperate Broadleaf and Mixed Forests biome (*Appalachian mixed mesophytic forests*, *Atlantic mixed forests*, *Cantabrian mixed forests*, *Caspian Hyrcanian mixed forests*, *Central Korean deciduous forests*, *Dinaric Mountains mixed forests*, *Euxine-Colchic broadleaf forests*, *Hokkaido deciduous forests*, *Manchurian mixed forests*, *Nihonkai evergreen forests*, *Nihonkai montane deciduous forests*, *Pindus Mountains mixed forests*, *Southeastern mixed forests*, *Taiheiyō evergreen forests*, *Taiheiyō montane deciduous forests*, *Western European broadleaf forests*, *Western Great Lakes forests*), plus three neighbouring ecoregions of the Temperate Coniferous Forest biome (*Cascade Mountains leeward forests*, *Central and Southern Cascades forests*, *Eastern Cascades forests*). Although temperate broadleaf and mixed forests also occur in a smaller area of the southern hemisphere, these southern forests form a separate evolutionary and biogeographical unit, not being related by a common history like the northern hemisphere ones (Milne & Abbott, 2002). For this reason, they were not considered. All chosen relevés had been classified as forest plots by *sPlot*, i.e. their plant cover was dominated by tree species. The taxa names were standardized to species level with *The Plant List* (2013) using the *Taxonstand* package (Cayuela *et al.*, 2019) in *R* version 3.6.2 (R Core Team, 2019). The final vegetation database contained 17,852 relevés and 7,684 standardized species names, considering only seed plants (gymnosperms and angiosperms), and encompassing all forest layers. To obtain the final list of species, only species that

were present in at least 5% of the relevés of an ecoregion were kept, rendering a list of 1,391 frequent species. There were two reasons to use this frequency threshold: (i) it made the search achievable with the resources available; and (ii) it filtered out rare species that may not be representative of *TBMFNH*, and which could have been recorded for a variety of reasons (wrong taxonomical identifications, mistakes in plot coordinates, border areas, unnatural areas such as parks, etc.). The result is a list of plant species that occur with relatively high frequency in forest plots of the *TBMFNH*; this includes trees, shrubs and shade-tolerant forbs and grasses, as well as common semi-open species from forest margins, secondary forests, ruderalized forest plots, etc.

Web of Science literature search

The list of frequent species (plus the synonyms recorded in the relevés) was incorporated into a Boolean search string, together with the words “(seed OR seeds) AND (dormancy OR germination)” (**Supporting Information 1**). This string was searched in the *Thompson Reuters Web of Science* on 5 Mar 2019, returning 6,791 results. The processing of these results is described in a PRISMA flow diagram (**Supporting Information 2**). A first filter of the results by the relevance of the title retained 1,489 references. An effort was made to access the full text of all these references, using in order the following channels: (i) “Find Full Text” functionality in *EndNote X9*, (ii) online access to the journals using the institutional subscriptions of the University of Oviedo; (iii) request to the authors through *ResearchGate*; (iv) general search of *pdf* files in *Google* using the article details. Unfortunately, 236 references could not be accessed. Most of these belonged to smaller publishers not included in the subscriptions of the University of Oviedo, and to older journal volumes which had not been digitalised. The resources of this project did not make it possible to further pursue these references, but they are included as a *BibTeX* file (**Supporting Information 3**) in case they can be processed in the future. The found texts were accessed to retrieve data for the database. At this stage, references were considered to be non-relevant if they (i)

did not pertain to species in the original frequent species list; (ii) did not study germination in species monocultures (e.g. competition studies); (iii) did not report the minimal information needed (number of seeds sown, germination percentage and germination temperature). Records were also excluded when they pertained to special germination treatments that do not allow for general comparisons (i.e. use of germination promoters such as the plant hormone gibberellic acid; drought treatments, including polyethylene glycol solutions; application of heavy metals or other pollutants; smoke-related germination promoters such as karrikins). Of the accessed references, 639 were non-relevant (*BiBTeX* file in **Supporting Information 4**) and 614 provided relevant data (*BiBTeX* file in **Supporting Information 5**).

Recording of the database

For each relevant reference, three blocks of information were recorded. The first block described the plant material, including the species, the populations that had been sampled, the year of sampling (or the year of publication if that information was missing), the country, the geographical coordinates (if not provided in the reference, the closest available toponyms were searched in *Google Maps*; in some cases the only geographical information was the country, in these cases the coordinates of the capital were recorded). The second block described the experimental conditions: experimental setting (greenhouse, incubator or room temperature), germination substrate and characteristics of the containers, seed storage, application of seed sterilization or nutrients, length of the germination incubations, use of stratification (none, cold [$< 15^{\circ}\text{C}$], warm [$\geq 15^{\circ}\text{C}$] or combinations of cold and warm), use of scarification, photoperiod, maximum germination temperature, minimum germination temperature, and weighted average germination temperature. The third block contained the final germination proportion: the reported final germination percentages (retrieved from the text, tables or figures) and the reported number of replicates and seeds per replicate were used to calculate a count of seeds sown and a count of seeds germinated.

Results

Summary of contents

The final database contains 4,828 records (germination proportions for a given seed lot of a species, recorded in a set of experimental conditions) from 614 references. The plant materials had been collected across the Temperate Broadleaf and Mixed Forest biome and surrounding areas (**Fig. 1**) in both hemispheres. The oldest record was from 1920 and the top three contributing countries were the USA (1,357), the UK (591) and Japan (531). There were 364 species represented, from 74 seed plant families. The total estimate of seeds used in the experiments was 953,807. The range of germination temperatures (weighted average of the daily thermoperiod) went from -4 to 43 °C, with 2,107 records of constant temperatures and 2,721 of alternating temperatures. Light was used in 2,849 records, darkness in 1,227 and 752 did not provide information on this parameter. The experiments were performed with unstratified seeds in 3,227 records, and of the rest, the majority (1,421) went through cold stratification. Scarification was applied to 252 records.

Database file

The database is provided as a csv file, comma separated, named **Supporting Information 6** (see online Supporting Information). The first row of the file contains the header data, with the following variables: *Taxon* (the original species names, as used in the references), *TPLName* (*The Plant List* species names), *Family* (the botanical family of the species), *Reference* (bibliographic source of the record), *Year* (year the seed lot was collected), *Country* (country where the seed lot was collected), *Population* (geographical information of the seed lot), *Latitude* (approximated latitude where the seed lot was collected, in decimal degrees), *Longitude* (approximated longitude where the seed lot was collected, in decimal degrees), *Dry_storage* (binary variable indicating whether the seed lot was kept at dry storage for more than a month between seed collection and start of the

experiments), *Setting* (environs used to control germination temperature: either a greenhouse, a climatically-controlled incubator, or room temperature), *Sterilization* (note on surface sterilization treatments applied to the seed lot before the experiments), *Nutrients* (note on nutrient application to the germination substrate), *Substrate* (substrate used to hold the seeds during the germination incubation), *Container_type* (type of container used to hold the substrate and the seeds during the experiments), *Container_size* (dimensions of the germination containers), *Container_number* (number of germination containers used for the treatment), *Sown_by_container* (number of seeds sown in each container), *Incubation_days* (number of days between the start of the experiment, not including stratification, and the day when germinated seeds were counted), *Scarification* (binary variable indicating whether the seed lot was scarified before the test), *Stratification_days* (number of days the seed lot was exposed to any type of stratification, before the test), *Stratification_temperature* (temperature or temperatures used during the stratification treatment, in degrees Celsius), *Stratification_type* (type of stratification, which can be none, cold, warm or combinations of cold and warm), *Stratification* (binary variable indicating whether the seed lot was stratified or not before the test), *Light* (binary variable indicating whether the seed lot was germinated in light or in darkness), *Photoperiod* (number of hours of exposure to light in the daily photoperiod), *Alternating* (binary variable indicating whether the germination test was conducted under constant or alternating temperatures), *Tdif* (difference in degrees between the hottest and the coldest temperatures of the daily thermoperiod), *Tmax* (hottest temperature in the daily thermoperiod), *Tmin* (coldest temperature in the daily thermoperiod), *Tmean* (mean germination temperature, weighted by the length of each phase of the daily thermoperiod), *Temperature* (aggregation of the mean germination temperature in 5 °C intervals), *Germinated* (count of seeds that germinated during the experiment), *Germinable* (count of seeds used in the experiment). Each row below the header represents a record for a seed lot germinated in a given set of experimental conditions. Presenting the response variable as a proportion (i.e. number of successes

Germinated, number of trials *Germinable*) has important benefits for meta-analysis of the database, as proportions can be analysed directly using a GLM, and a variance term does not need to be specified since it is incorporated as part of the analysis (Mengersen & Gurevitch, 2013).

***SylvanSeeds* app**

To facilitate the visualization of the database, the *SylvanSeeds* app was written using the *shiny* package (Chang *et al.*, 2020). It is publicly accessible at <http://sylvanseeds.shinyapps.io/sylvanseeds/>. The app uses the *tidyverse* package (Wickham *et al.*, 2019) to aggregate and show results for species and experimental treatments (i.e. aggregating all seed lots of the same species germinated in the same experimental conditions). To facilitate comparisons, germination temperatures are aggregated to 5 °C intervals. When there is only one seed lot per species and combination of experimental conditions, the binomial 95% confidence interval is calculated using the Wilson method in the *binom* package (Dorai-Raj, 2014). When there is more than one seed lot per species and combination of experimental conditions, the aggregate proportion and binomial confidence intervals are calculated using binomial-normal meta-analysis models (Stijnen *et al.*, 2010) as implemented in the package *metafor* (Viechtbauer, 2010). By visiting the app, users can consult the available germination information for a species (**Fig. 2**), the origin of its seed lots, and the bibliographical references for the species.

Discussion

The database provided in this article, and the web app to visualize it, can have wide applicability in science and beyond. It can extend current trends in global analyses of plant traits at species (Díaz *et al.*, 2016) and community levels (Bruehlheide *et al.*, 2018) to seed germination. The ecological determinants of seed germination are also valuable information for species distribution models (Bykova *et al.*, 2012). The visualization of the database with the *SylvanSeeds* app can help

plant ecologists select experimental treatments that are adequate for their experiments (Baskin *et al.*, 2006). Outside of academia, germination data are useful for seed industries (De Vitis *et al.*, 2017), restoration practitioners (Ladouceur *et al.*, 2018) and the implementation of regional schemes for seed-based landscape intervention (Jiménez-Alfaro *et al.*, 2020). The app is accessible to citizens in general who are interested in germinating wild plants. Finally, apart from the dataset itself, this article can bring two innovations to the seed ecology community, helping to advance the agenda of functional seed ecology (Saatkamp *et al.*, 2019). First, the meta-analysis-inspired methodology used to compile the dataset can be extended to other biomes and lists of species, contributing to the creation of a global database for ecologically and biogeographically coherent floras. Second, the database and the app can serve as a standard in further efforts to compile and standardize seed germination data.

Acknowledgements

E.F.P. received financial support from the Government of Asturias and the FP7 – Marie Curie - COFUND programme of the European Commission (Grant ‘Clarín’ ACB17-19). The list of species was built thanks to sPlot, a project of the Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig.

Author Contribution

This is a single-author article.

References

Baskin SEB Carol C.; Meyer. 1995. Two types of morphophysiological dormancy in seeds of two genera (*Osmorhiza* and *Erythronium*) with an Arcto-Tertiary distribution pattern. *American Journal of Botany* **82**: 293–298.

- 277 **Baskin CC, Baskin JM. 2014.** *Seeds. Ecology, Biogeography and Evolution of*
 278 *Dormancy and Germination. Second Edition.* San Diego: Academic Press.
- 279 **Baskin CC, Thompson K, Baskin JM. 2006.** Mistakes in germination ecology and
 280 how to avoid them. *Seed Science Research* **16**: 165–168.
- 281 **Bruehlheide H, Dengler J, Jiménez-Alfaro B, Purschke O, Hennekens SM,**
 282 **Chytrý M, Pillar VD, Jansen F, Kattge J, Sandel B. 2019.** sPlot—A new tool for
 283 global vegetation analyses. *Journal of Vegetation Science* **30**: 161–186.
- 284 **Bruehlheide H, Dengler J, Purschke O, Lenoir J, Jiménez-Alfaro B, Hennekens**
 285 **SM, Botta-Dukát Z, Chytrý M, Field R, Jansen F. 2018.** Global trait–
 286 environment relationships of plant communities. *Nature Ecology & Evolution* **2**:
 287 1906–1917.
- 288 **Bykova O, Chuine I, Morin X, Higgins SI. 2012.** Temperature dependence of the
 289 reproduction niche and its relevance for plant species distributions. *Journal of*
 290 *Biogeography* **39**: 2191–2200.
- 291 **Carta A, Skourti E, Mattana E, Vandeloos F, Thanos CA. 2017.** Photoinhibition
 292 of seed germination: occurrence, ecology and phylogeny. *Seed Science*
 293 *Research* **27**: 131–153.
- 294 **Cayuela L, Stein A, Oksanen J. 2019.** Taxonstand: taxonomic standardization of
 295 plant species names. R package version 2.2.
- 296 **Chang W, Cheng J, Allaire J, Xie Y, Mcpherson J. 2020.** shiny: web application
 297 framework for R. R Package Version 1.4.0.2.
- 298 **De Vitis M, Abbandonato H, Dixon KW, Laverack G, Bonomi C, Pedrini S.**
 299 **2017.** The European native seed industry: characterization and perspectives in
 300 grassland restoration. *Sustainability* **9**: 1682.

- 301 **Díaz S, Kattge J, Cornelissen JHC, Wright IJ, Lavorel S, Dray S, Reu B, Kleyer**
 302 **M, Wirth C, Colin Prentice I et al. 2016.** The global spectrum of plant form and
 303 function. *Nature* **529**: 167–171.
- 304 **Donohue K, Rubio de Casas R, Burghardt L, Kovach K, Willis CG. 2010.**
 305 Germination, postgermination adaptation, and species ecological ranges.
 306 *Annual Review of Ecology, Evolution, and Systematics* **41**: 293–319.
- 307 **Dorai-Raj S. 2014.** binom: binomial confidence intervals for several
 308 parameterizations. R package version 1.1-1.
- 309 **Fernández-Pascual E, Mattana E, Pritchard HW. 2019.** Seeds of future past:
 310 climate change and the thermal memory of plant reproductive traits. *Biological*
 311 *Reviews* **94**: 439–456.
- 312 **Fernández-Pascual E, Pérez-Arcoiza A, Prieto JA, Díaz TE. 2017.**
 313 Environmental filtering drives the shape and breadth of the seed germination
 314 niche in coastal plant communities. *Annals of Botany* **119**: 1169–1177.
- 315 **Finch-Savage WE, Leubner-Metzger G. 2006.** Seed dormancy and the control of
 316 germination. *New Phytologist* **171**: 501–523.
- 317 **Grime JP, Mason G, Curtis AV, Rodman J, Band SR, Mowforth MAG, Neal**
 318 **AM, Shaw S. 1981.** A comparative study of germination characteristics in a local
 319 flora. *Journal of Ecology* **69**: 1017–1059.
- 320 **Huang Z, Liu S, Bradford KJ, Huxman TE, Venable DL. 2016.** The contribution
 321 of germination functional traits to population dynamics of a desert plant
 322 community. *Ecology* **97**: 250–261.
- 323 **Jiménez-Alfaro B, Frischie S, Stolz J, Gálvez-Ramírez C. 2020.** Native plants
 324 for greening Mediterranean agroecosystems. *Nature Plants* **6**: 209–214.

- 325 **Jiménez-Alfaro B, Silveira FAO, Fidelis A, Poschlod P, Commander LE. 2016.**
 326 Seed germination traits can contribute better to plant community ecology.
 327 *Journal of Vegetation Science* **27**: 637–645.
- 328 **Jurado E, Flores J. 2005.** Is seed dormancy under environmental control or bound
 329 to plant traits? *Journal of Vegetation Science* **16**: 559–564.
- 330 **Koricheva J, Gurevitch J, Mengersen KL. 2013.** *Handbook of Meta-Analysis in*
 331 *Ecology and Evolution*. Princeton University Press.
- 332 **Ladouceur E, Jiménez-Alfaro B, Marin M, De Vitis M, Abbandonato H,**
 333 **Iannetta PPM, Bonomi C, Pritchard HW. 2018.** Native seed supply and the
 334 restoration species pool. *Conservation Letters* **11**: e12381–e12381.
- 335 **Larson JE, Funk JL. 2016.** Regeneration: an overlooked aspect of trait-based
 336 plant community assembly models. *Journal of Ecology* **104**: 1284–1298.
- 337 **Mengersen K, Gurevitch J. 2013.** Using other metrics of effect size in meta-
 338 analysis. *Handbook of meta-analysis in ecology and evolution*: 72–85.
- 339 **Milne RI, Abbott RJ. 2002.** The origin and evolution of tertiary relict floras.
 340 *Advances in Botanical Research* **38**: 281–314.
- 341 **Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV,**
 342 **Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC. 2001.**
 343 Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global
 344 map of terrestrial ecoregions provides an innovative tool for conserving
 345 biodiversity. *BioScience* **51**: 933–938.
- 346 **Pérez-Harguindeguy N, Díaz S, Garnier É, Lavorel S, Poorter H, Jaureguiberry**
 347 **P, Bret-Harte M, Cornwell WK, Craine J, Gurvich D. 2013.** New handbook for
 348 standardised measurement of plant functional traits worldwide. *Australian*
 349 *Journal of Botany* **61**: 167–234.
- 350 **R Core Team. 2019.** R: A language and environment for statistical computing.

Saatkamp A, Cochrane A, Commander L, Guja LK, Jimenez-Alfaro B, Larson J, Nicotra A, Poschlod P, Silveira FA, Cross AT. 2019. A research agenda for seed-trait functional ecology. *New Phytologist* **221**: 1764–1775.

Stijnen T, Hamza TH, Özdemir P. 2010. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in medicine* **29**: 3046–3067.

The Plant List. 2013. Version 1.1.

Thompson JP, Mason GK, Grime. 1977. Seed germination in response to diurnal fluctuations of temperature. *Nature* **267**: 147–149.

Van Assche JA, Vandelook FEA. 2006. Germination ecology of eleven species of Geraniaceae and Malvaceae, with special reference to the effects of drying seeds. *Seed Science Research* **16**: 283–290.

Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**: 1–48.

Washitani I. 1988. Effects of high temperatures on the permeability and germinability of the hard seeds of (*Rhus javanica* L. *Annals of Botany* **62**: 13–16.

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J. 2019. Welcome to the Tidyverse. *Journal of Open Source Software* **4**: 1686.

Supporting Information

Supporting Information 1 Boolean search string used in the Web of Science.

Supporting Information 2 PRISMA 2009 flow chart describing the process of reference screening and inclusion. **Supporting Information 3** BibTeX bibliography listing the not found full texts. **Supporting Information 4** BibTeX bibliography listing the non-relevant references. **Supporting Information 5** BibTeX bibliography

376 listing relevant references included in the *SylvanSeeds* database. **Supporting**
377 **Information 6** *SylvanSeeds* database in a single .csv file.

378 **Data accessibility**

379 All persons can use the database providing they cite this paper properly in any
380 publications or in the metadata of any derived products that are produced using the
381 database. The database is provided as Supporting Information. It can be visualized
382 with the *SylvanSeeds* app at <http://sylvanseeds.shinyapps.io/sylvanseeds/>. The
383 code of the app is stored at (note: the GitHub repo is kept private until publication
384 of the manuscript).

385 **Figures**

386

387 **Figure 1** Geographical distribution of the germination records in the database.
388 Each golden circle is a record. The green areas correspond to the extension of the
389 Temperate Broadleaf and Mixed Forests biome according to the WWF.

390

391 **Figure 2** Example of the germination records as shown by the *SylvanSeeds* app.
392 Records for one species, the European pedunculate oak, *Quercus robur*. Each
393 panel shows the results for a combination of experimental conditions, with the
394 germination temperature varying within each panel. Bars represent the mean
395 germination proportion and brackets the 95% binomial confidence interval.

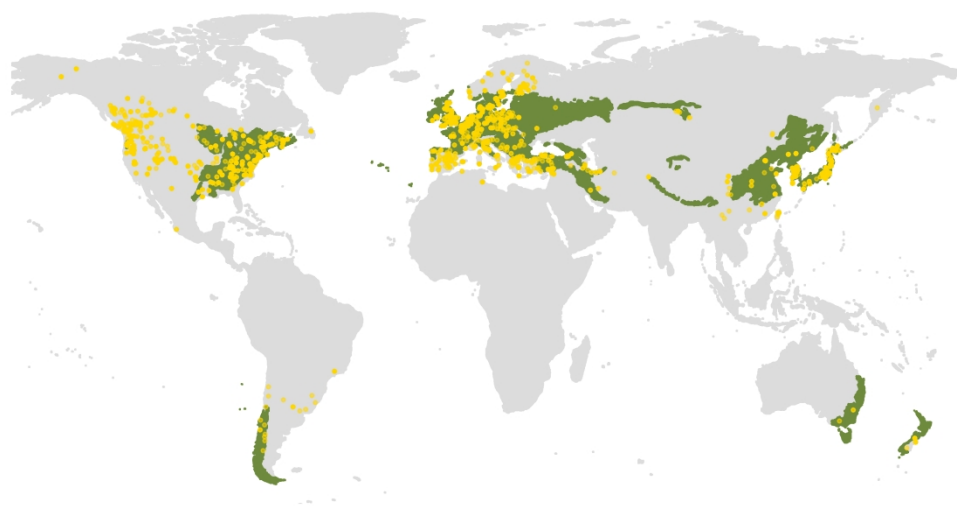


Figure 1 Geographical distribution of the germination records in the database. Each golden circle is a record. The green areas correspond to the extension of the Temperate Broadleaf and Mixed Forests biome according to the WWF.

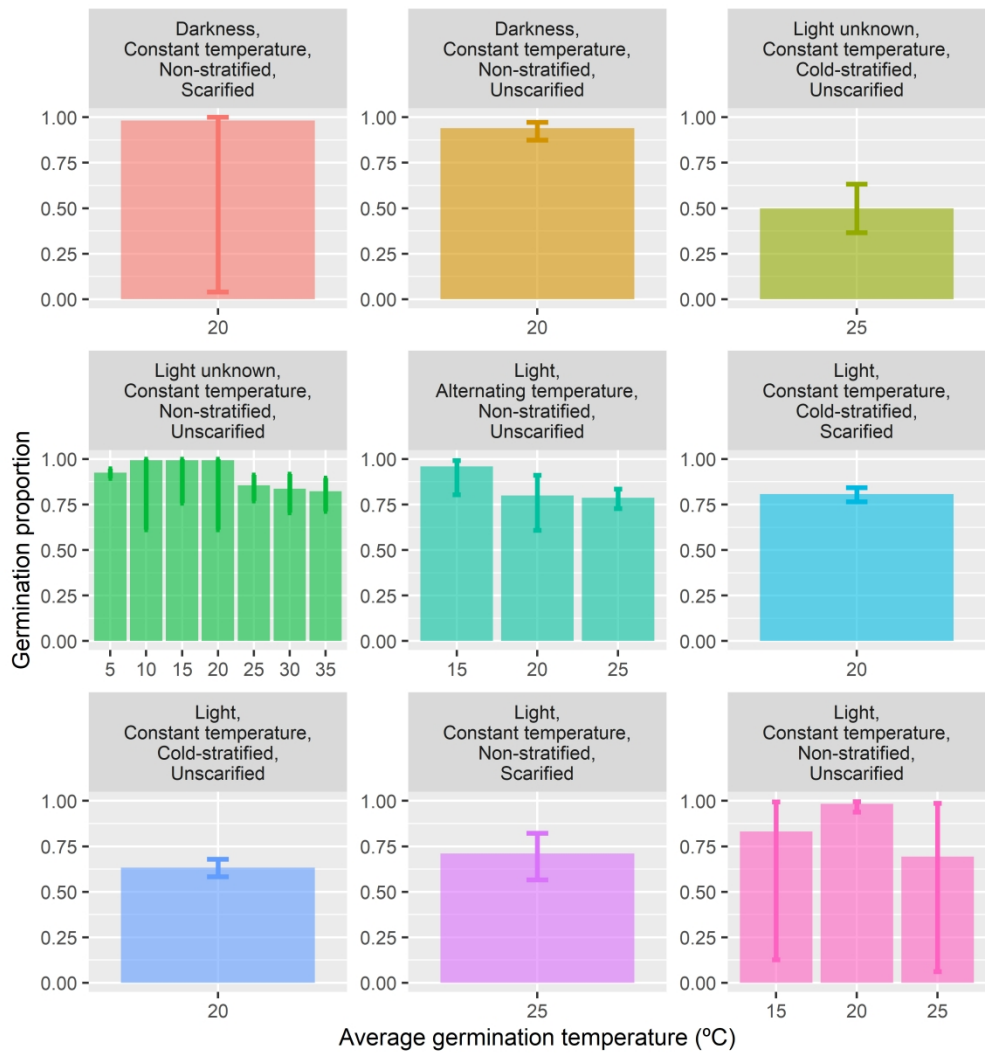


Figure 2 Example of the germination records as shown by the SylvanSeeds app. Records for one species, the European pedunculate oak, *Quercus robur*. Each panel shows the results for a combination of experimental conditions, with the germination temperature varying within each panel. Bars represent the mean germination proportion and brackets the 95% binomial confidence interval.