Letters

SeedArc, a global archive of primary seed germination data

Introduction

Biology is elucidating global patterns and processes at a rapid pace, as international initiatives mobilize an ever-expanding amount of data on the environment (Karger *et al.*, 2017; Lembrechts *et al.*, 2020), species occurrences (GBIF, 2023), community composition (Sabatini *et al.*, 2021), phylogenomics (Smith *et al.*, 2020) and functional traits (Kattge *et al.*, 2020). Data synthesis and modern comparative methods (Garamszegi, 2014) are shedding light on the origin and distribution of present-day biodiversity, including the underlying processes of trait evolution and species adaptation (Sauquet & Magallón, 2018). In plant science, trait-based macroecology is exploring how species and communities respond to the environment and affect ecosystem functions (Wright *et al.*, 2004; Chave *et al.*, 2009; Mommer & Weemstra, 2012; Bruelheide *et al.*, 2018). Together with plant height and seed mass, leaf–wood–root traits are helping to describe the global spectrum of plant form and function (Díaz *et al.*, 2016; Carmona *et al.*, 2021).

In contrast to the growing understanding of vegetative plant functions, our knowledge of seed-based plant regeneration is less profound, except for the functions related to seed mass (Jiménez-Alfaro *et al.*, 2016; Larson & Funk, 2016). The regeneration phase plays a crucial role in the plant life cycle, with effects on species turnover, community assembly and plant conservation (Grubb, 1977). The regeneration niche reflects the multifactorial interactions between seeds or seedlings and their environment, including the functions of dispersal, seed persistence, germination and seedling establishment (Poschlod *et al.*, 2013). Unfortunately, little progress has been made in synthesizing data on seed traits (Jiménez-Alfaro *et al.*, 2016; Larson & Funk, 2016).

This is especially evident in the case of seed germination. Accurately describing the seed germination function requires integrating an ample range of metrics reflecting different drivers of physiological seed germination – from germination proportions, the most available germination trait in TRY (Kattge *et al.*, 2020), to calculations of speed and synchrony (Lozano-Isla *et al.*, 2019) or more complex models such as hydrothermal germination thresholds (Dürr *et al.*, 2015). As a wider range of seed germination metrics and response patterns are synthesized, it will be possible to explore the full spectrum of seed germination strategies that exist among plant populations and species.

The seed germination spectrum

Seed germination is a physiological process regulated by integrated responses to multiple environmental cues (Donohue *et al.*, 2010). Among these cues, the most important are water (Rosbakh *et al.*, 2019), temperature (Fernández-Pascual *et al.*, 2019), diurnal temperature fluctuations (Thompson *et al.*, 1977), light (Carta *et al.*, 2017) and seed dormancy regulators (Finch-Savage & Leubner-Metzger, 2006; Baskin & Baskin, 2014). The combination of interactions and trade-offs among responses to environmental germination cues across plant species is what we define as the seed germination spectrum (Saatkamp *et al.*, 2019; Fernández-Pascual *et al.*, 2021).

Despite extensive research on germination ecology in the last decades, there has been little exploration of how the seed germination spectrum integrates with other adult and regeneration trait spectra (Saatkamp *et al.*, 2019). Global-scale studies are needed to clarify the role of regeneration by seeds in shaping plant evolution and the distribution of past and future plant diversity. The theory underlying the seed germination spectrum has been laid out by decades of work on seed ecology (Baskin & Baskin, 2014), but empirical studies testing major ecological hypotheses at both global and local scales remain elusive without a standardized seed germination database.

The need for a global archive of primary seed germination data

The seed ecology community has recently recognized the need to synthesize knowledge, setting the research agenda for functional seed ecology (Saatkamp *et al.*, 2019). These efforts have led to the creation of several tools like biome- or family-specific seed germination databases (Fernández-Pascual, 2021; Ordóñez-Parra *et al.*, 2023); an updated global list of species dormancy types (Rosbakh *et al.*, 2020); or R packages to analyse seed germination data (Lozano-Isla *et al.*, 2019; Fernández-Pascual & González-Rodríguez, 2020). In parallel, macroecological studies of seed germination are starting to be more frequent (Sentinella *et al.*, 2020; Fernández-Pascual *et al.*, 2021; Carta *et al.*, 2022). Despite these advances, current efforts are biased to specific floras or habitats of mid-latitude temperate biomes, with the major exception of Brazil (Ordóñez-Parra *et al.*, 2023). Large amounts of the Earth's biodiversity hotspots remain unexplored by global-scale germination analyses, especially in tropical areas and continents such as Africa or parts of Asia (Hughes *et al.*, 2021; Visscher *et al.*, 2022).

Archiving seed germination data across species and biomes will allow us to overcome major issues limiting our understanding of seed functions, such as: interspecific variation of germination responses; the heterogeneity of research approaches to measure germination responses; and the need for appropriate statistical methods to model multiple seed responses and different research contexts. Centralizing germination data will also confront new challenges related to the nature of the data: records often track germination proportions in continuous experimental procedures but may also include responses to single and interacting environmental cues that require the application of statistical analyses to derive germination indices as proxy traits (Carta *et al*., 2022).

SeedArc

To advance global understanding of the seed germination spectrum, we are launching *SeedArc*, a global archive of primary seed germination data (Fig. 1). *SeedArc* compiles primary seed germination data to synthesize the seed germination spectrum at a global scale. We define primary germination data as the raw information generated under controlled laboratory conditions in a germination experiment conducted with a specific seed lot (i.e. a collection of seeds of a plant species obtained from a geo-referenced population at a specific date; Fernández-Pascual *et al*., 2021; Carta *et al*., 2022). Collecting primary germination records as opposed to derived traits, such as germination indices (Dürr *et al*., 2015; Lozano-Isla *et al*., 2019), will allow users to extract more (and more comparable) types of information from each experiment, and conduct broad analyses of the database tailored to specific research questions (Fernández-Pascual *et al*., 2021). Importantly, the association between primary germination records and unique geo-referenced populations and collection events will allow researchers to model and explore intraspecific variability in seed germination, as well as variability produced by different experimental settings.

SeedArc includes both published and unpublished germination data and is a unique repository for collecting and archiving

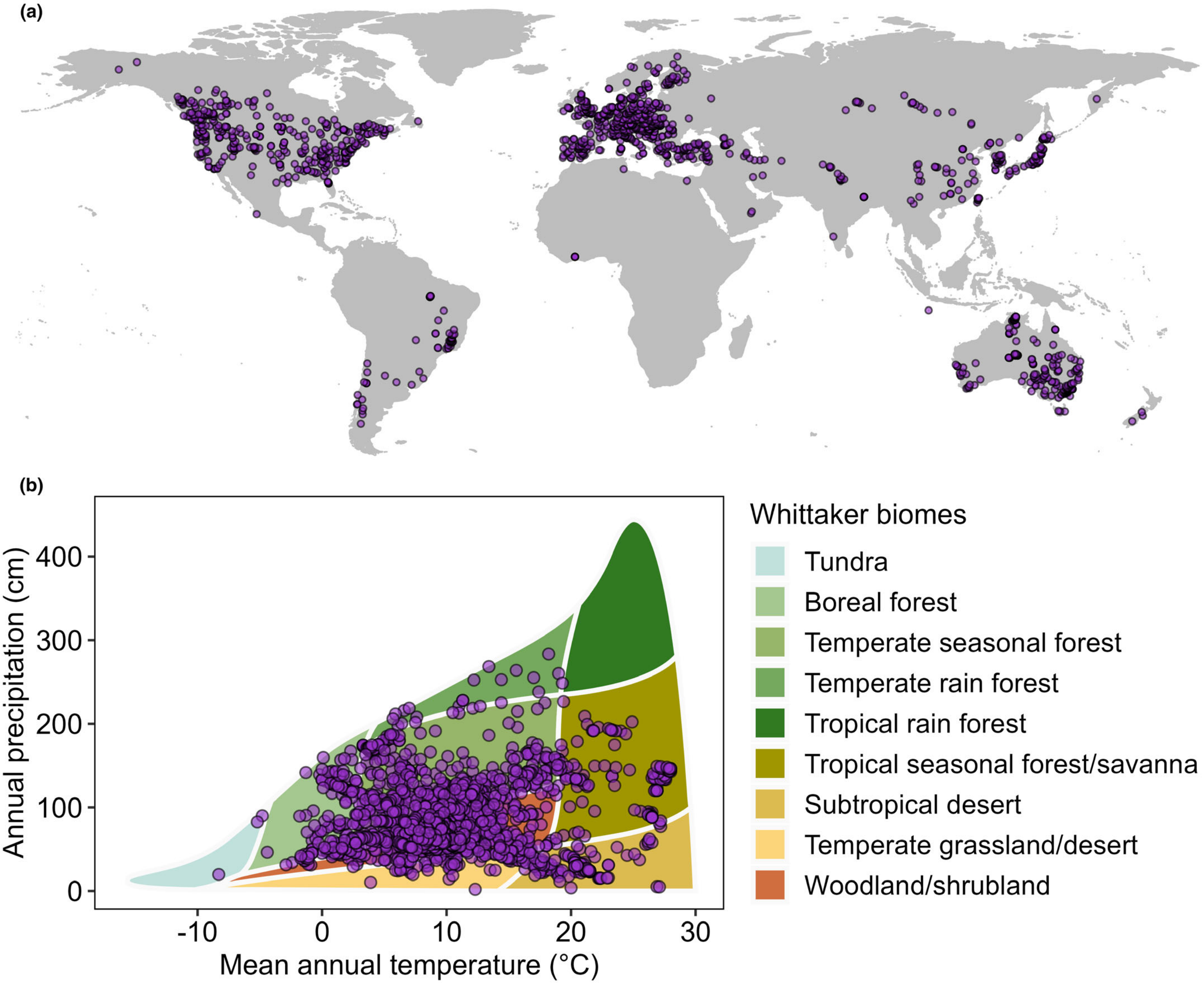

Fig. 1 *SeedArc* v.1.0. The initial version of the *SeedArc* database holds 47 433 primary germination records of 4516 plant species. (a) The map shows the provenance of the germination records, as purple circles. (b) Whittaker biome plot showing the distribution of the records across the Earth's biomes.

experimental data that are currently in notebooks, regional seedbank databases, scattered in scientific articles in different languages or in technical reports. By submitting data to *SeedArc*, contributors will ensure that their efforts will be acknowledged and that seed germination data will be accessible and usable for future studies. In our vision, the ultimate result will be an easily accessible, interactive, updatable database across a wide range of ecosystems, biomes and plant lineages. In the first round of data mobilization, we focus on primary germination data obtained from seeds collected from wild populations of non-cultivated seed plants.

The overarching goal of *SeedArc* is to allow comparative studies of seed germination across various levels of biological organization. *SeedArc* adopts the theoretical background of the research agenda for seed-trait functional ecology (Saatkamp *et al.*, 2019), offering a first step towards a global database of seed functional traits, eventually to be linked with other global plant databases (Kattge *et al.*, 2020; Sabatini *et al.*, 2021). *SeedArc* and its associated metadata will allow users to answer questions related to plant regeneration from seeds, evolution, biogeography and vegetation ecology. Moreover, in the light of the ongoing habitat degradation and climate emergency, *SeedArc* data will support projects in conservation biology and restoration ecology that rely on the successful germination of seeds (Kildisheva *et al.*, 2020).

We envision four critical research questions to address initially from *SeedArc*:

(**1**) How does the seed germination spectrum integrate with other adult and regeneration trait spectra? Answering this question will allow us to map seed traits to germination functions and to assess how regenerative traits are associated with vegetative traits (Hoyle *et al.*, 2015).

(**2**) How is seed germination regulated by past, current and future environmental conditions? This will broaden our understanding of links between regeneration and spatio-temporal ecological gradients (Walck *et al.*, 2011; Rubio de Casas *et al.*, 2017), particularly relevant for understanding potential demographic effects of climate change.

(**3**) What processes drive seed germination variability? Clarification is needed to assess the degree of evolvability vs phenotypic plasticity (Fernández-Pascual *et al.*, 2019), which in turn allows us to unveil the macroevolution of seed germination (Carta *et al.*, 2022) and the role of germination syndromes in angiosperm diversification (Linkies *et al.*, 2010; Willis *et al.*, 2014).

(**4**) How do we tackle the methodological and computational challenges faced by macroecological studies of seed germination? Standardized descriptions of the multiple germination functions (probability, velocity, synchrony) can support global meta-analyses (Carta *et al.*, 2022) and summarize the primary germination axes across ecological and evolutionary scales.

Data mobilization and use

The aim of *SeedArc* is to create the largest repository of seed germination data globally. To ensure a fair protocol of data acquisition and use, we have prepared a set of data-management rules accessible on the *SeedArc* website (https://www.unioviedo.es/seedarc). The current version (1.2) was approved by the *SeedArc* coordination team in March 2023. These rules aim at developing a transparent system of data contribution and use describing: the objectives of the database; the coordination team of *SeedArc*, its role and the process to renew it; the nature of the data contributors; the terms of data ownership, data use and co-authorship, highlighting that data contributors retain the ownership of the data and the right to collaborate in *SeedArc* research; and the validity of the rules and the process to update them. In practice, the scientific contribution of *SeedArc* will be generated by research projects. Individual projects using the database will start upon request of one member of the coordination team or the data contributors. The essential workflow of data, projects and publications will be as follows: submitted data (= primary data) remain property of the contributors; primary data are managed by the data coordination team; access and usage of the primary data are done through research projects; each research project must publish, open access, all data that are used by the project (= secondary data). In other words, all data in *SeedArc* will end up being published open access, once they are used and presented in a research publication. The leaders of *SeedArc* projects will need to follow ethical rules of scientific publishing and will be obliged to produce open science research.

Call for data

SeedArc enables exploration of macroevolutionary and ecological drivers of seed germination. It captures both optimal and suboptimal germination conditions of species, and accounts for intraspecific variability as a key aspect of phenotypic variability in an evolutionary context (Violle *et al.*, 2012; Silvestro *et al.*, 2015). Research opportunities within the *SeedArc* initiative will be extensive and organized through research projects to address specific questions; any data contributor will be able to propose and lead a project following the procedure outlined in the data-management rules.

We invite the community of plant scientists to join *SeedArc* by contributing primary data on seed germination from all over the world, regardless of the type of experiment conducted and the number of species tested. We offer data ownership and co-authorship rules that aim to support collaboration and attribution under an open science scheme. By contributing to *SeedArc*, researchers will ensure that their data are preserved and reused. Updated information about *SeedArc* and instructions for contributing data can be found at https://www.unioviedo.es/seedarc.

Acknowledgements

SeedArc is supported by the Jardín Botánico Atlántico de Gijón/Xixón, project SV-20-GIJON-JBA. The EOARC is jointly funded by the USDA-ARS and Oregon State Agricultural Experiment Station. USDA-ARS and Oregon State University are equal-opportunity providers and employers. Mention of a proprietary product does not constitute a guarantee or warranty of the product by USDA-ARS, Oregon State University, or the authors and does not imply approval to the exclusion of other products.

Competing interests

None declared.

Author contributions

EFP and BJ-A conceived the idea. EFP, BJ-A, AC and SR drafted the first version of the data-management rules, research goals and structure of the database. EFP, AC, SR, LG, SSP, FAOS, SCC, JEL and BJ-A contributed data, revised and approved the data-management rules, research goals and structure of the database. EFP led manuscript writing, with contributions from AC, SR, LG, SSP, FAOS, SCC, JEL and BJ-A.

ORCID

Angelino Carta https://orcid.org/0000-0001-8437-818X
Si-Chong Chen https://orcid.org/0000-0002-6855-2595
Eduardo Fernández-Pascual https://orcid.org/0000-0002-4743-9577
Lydia Guja https://orcid.org/0000-0001-5945-438X
Borja Jiménez-Alfaro https://orcid.org/0000-0001-6601-9597
Julie E. Larson https://orcid.org/0000-0001-7968-916X
Shyam S. Phartyal https://orcid.org/0000-0003-3266-6619
Sergey Rosbakh https://orcid.org/0000-0002-4599-6943
Fernando A. O. Silveira https://orcid.org/0000-0001-9700-7521

Data availability

The data and code used to produce this article are available at https://github.com/efernandezpascual/seedarcms. A version of record of the repository is deposited in Zenodo at doi:10.5281/zenodo.8114443.

Eduardo Fernández-Pascual[1]*, **Angelino Carta**[2,3], **Sergey Rosbakh**[4], **Lydia Guja**[5,6], **Shyam S. Phartyal**[7], **Fernando A. O. Silveira**[8], **Si-Chong Chen**[9,10], **Julie E. Larson**[11] **and Borja Jiménez-Alfaro**[1]

[1]IMIB Biodiversity Research Institute (University of Oviedo – CSIC – Principality of Asturias), University of Oviedo, E-33600, Mieres, Spain;
[2]Department of Biology, Botany Unit, University of Pisa, 56122, Pisa, Italy;
[3]CIRSEC – Centre for Climate Change Impact, University of Pisa, 56122, Pisa, Italy;
[4]Department of Plant and Environmental Sciences, University of Copenhagen, DK-1871, Frederiksberg C, Denmark;
[5]National Seed Bank, Australian National Botanic Gardens, Parks Australia, 2601, Acton, ACT, Australia;
[6]Centre for Australian National Biodiversity Research (A Joint Venture Between Parks Australia and CSIRO), CSIRO, 2601, Acton, ACT, Australia;
[7]School of Ecology and Environment Studies, Nalanda University, 803116, Rajgir, India;
[8]Department of Genetics, Ecology and Evolution, Federal University of Minas Gerais, 31320290, Belo Horizonte, Brazil;
[9]Wuhan Botanical Garden, Chinese Academy of Sciences, 430074, Wuhan, China;
[10]Millennium Seed Bank, Royal Botanic Gardens Kew, RH176TN, Wakehurst, UK;
[11]USDA Agricultural Research Service, Eastern Oregon Agricultural Research Center, Burns, OR 97720, USA
(*Author for correspondence:
email fernandezpeduardo@uniovi.es)

References

Baskin CC, Baskin JM. 2014. *Seeds. Ecology, biogeography and evolution of dormancy and germination, 2nd edn.* San Diego, CA, USA: Academic Press.

Bruelheide H, Dengler J, Purschke O, Lenoir J, Jiménez-Alfaro B, Hennekens SM, Botta-Dukát Z, Chytrý M, Field R, Jansen F. 2018. Global trait–environment relationships of plant communities. *Nature Ecology & Evolution* 2: 1906–1917.

Carmona CP, Bueno CG, Toussaint A, Träger S, Díaz S, Moora M, Munson AD, Pärtel M, Zobel M, Tamme R. 2021. Fine-root traits in the global spectrum of plant form and function. *Nature* **597**: 683–687.

Carta A, Fernández-Pascual E, Gioria M, Müller JV, Rivière S, Rosbakh S, Saatkamp A, Vandelook F, Mattana E. 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.

Carta A, Skourti E, Mattana E, Vandelook F, Thanos CA. 2017. Photoinhibition of seed germination: occurrence, ecology and phylogeny. *Seed Science Research* **27**: 131–153.

Chave J, Coomes D, Jansen S, Lewis SL, Swenson NG, Zanne AE. 2009. Towards a worldwide wood economics spectrum. *Ecology Letters* **12**: 351–366.

Díaz S, Kattge J, Cornelissen JHC, Wright IJ, Lavorel S, Dray S, Reu B, Kleyer M, Wirth C, Colin Prentice I *et al.* 2016. The global spectrum of plant form and function. *Nature* **529**: 167–171.

Donohue K, Rubio de Casas R, Burghardt L, Kovach K, Willis CG. 2010. Germination, postgermination adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution, and Systematics* **41**: 293–319.

Dürr C, Dickie JB, Yang XY, Pritchard HW. 2015. Ranges of critical temperature and water potential values for the germination of species worldwide: contribution to a seed trait database. *Agricultural and Forest Meteorology* **200**: 222–232.

Fernández-Pascual E. 2021. *SylvanSeeds*, a seed germination database for temperate deciduous forests. *Journal of Vegetation Science* **21**: e12960.

Fernández-Pascual E, Carta A, Mondoni A, Cavieres LA, Rosbakh S, Venn S, Satyanti A, Guja L, Briceño VF, Vandelook F *et al.* 2021. The seed germination spectrum of alpine plants: a global meta-analysis. *New Phytologist* **229**: 3573–3586.

Fernández-Pascual E, González-Rodríguez G. 2020. *Seedr: hydro and thermal time seed germination models in r.* R package v.0.3.0. [WWW document] URL https://CRAN.r-project.org/package=seedr.

Fernández-Pascual E, Mattana E, Pritchard HW. 2019. Seeds of future past: climate change and the thermal memory of plant reproductive traits. *Biological Reviews* **94**: 439–456.

Finch-Savage WE, Leubner-Metzger G. 2006. Seed dormancy and the control of germination. *New Phytologist* **171**: 501–523.

Garamszegi LZ. 2014. *Modern phylogenetic comparative methods and their application in evolutionary biology.* Berlin, Heidelberg, Germany; New York, NY, USA: Springer.

GBIF. 2023. *What is GBIF?* [WWW document] URL https://www.gbif.org/what-is-gbif [accessed 21 July 2023].

Grubb PJ. 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews* **52**: 107–145.

Hoyle GL, Steadman KJ, Good RB, McIntosh EJ, Galea LME, Nicotra AB. 2015. Seed germination strategies: an evolutionary trajectory independent of vegetative functional traits. *Frontiers in Plant Science* **6**: 1–13.

Hughes AC, Orr MC, Ma K, Costello MJ, Waller J, Provoost P, Yang Q, Zhu C, Qiao H. 2021. Sampling biases shape our view of the natural world. *Ecography* **44**: 1259–1269.

Jiménez-Alfaro B, Silveira FAO, Fidelis A, Poschlod P, Commander LE. 2016. Seed germination traits can contribute better to plant community ecology. *Journal of Vegetation Science* **27**: 637–645.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* **4**: 170122.

Kattge J, Bönisch G, Díaz S, Lavorel S, Prentice IC, Leadley P, Tautenhahn S, Werner GDA, Aakala T, Abedi M *et al.* 2020. TRY plant trait database – enhanced coverage and open access. *Global Change Biology* **26**: 119–188.

Kildisheva OA, Dixon KW, Silveira FAO, Chapman T, Di Sacco A, Mondoni A, Turner SR, Cross AT. 2020. Dormancy and germination: making every seed count in restoration. *Restoration Ecology* **28**: S256–S265.

Larson JE, Funk JL. 2016. Regeneration: an overlooked aspect of trait-based plant community assembly models. *Journal of Ecology* **104**: 1284–1298.

Lembrechts JJ, Aalto J, Ashcroft MB, De Frenne P, Kopecký M, Lenoir J, Luoto M, Maclean IMD, Roupsard O, Fuentes-Lillo E *et al.* 2020. SoilTemp: a global database of near-surface temperature. *Global Change Biology* **26**: 6616–6629.

Linkies A, Graeber K, Knight C, Leubner-Metzger G. 2010. The evolution of seeds. *New Phytologist* **186**: 817–831.

Lozano-Isla F, Benites-Alfaro OE, Pompelli MF. 2019. GerminaR: an r package for germination analysis with the interactive web application 'GerminaQuant for r'. *Ecological Research* **34**: 339–346.

Mommer L, Weemstra M. 2012. The role of roots in the resource economics spectrum. *New Phytologist* **195**: 725–727.

Ordóñez-Parra CA, Dayrell RLC, Negreiros D, Andrade ACS, Andrade LG, Antonini Y, Barreto LC, Barros F d V, Carvalho V d C, Corredor BAD *et al.* 2023. Rock n' seeds: a database of seed functional traits and germination experiments from Brazilian rock outcrop vegetation. *Ecology* **104**: e3852.

Poschlod P, Abedi M, Bartelheimer M, Drobnik J, Rosbakh S, Saatkamp A. 2013. Seed ecology and assembly rules in plant communities. In: van der Maarel E, Franklin J, eds. *Vegetation Ecology.* Chichester, UK: John Wiley & Sons, 164–202.

Rosbakh S, Baskin CC, Baskin JM. 2020. Nikolaeva *et al.*'s reference book on seed dormancy and germination. *Ecology* **101**: e03049.

Rosbakh S, Phartyal SS, Poschlod P. 2019. Seed germination traits shape community assembly along a hydroperiod gradient. *Annals of Botany* **125**: 67–78.

Rubio de Casas R, Willis CG, Pearse WD, Baskin CC, Baskin JM, Cavender-Bares J. 2017. Global biogeography of seed dormancy is determined by seasonality and seed size: a case study in the legumes. *New Phytologist* **214**: 1527–1536.

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

Sabatini FM, Lenoir J, Hattab T, Arnst EA, Chytrý M, Dengler J, De Ruffray P, Hennekens SM, Jandt U, Jansen F *et al.* 2021. sPlotOpen – an environmentally balanced, open-access, global dataset of vegetation plots. *Global Ecology and Biogeography* **30**: 1740–1764.

Sauquet H, Magallón S. 2018. Key questions and challenges in angiosperm macroevolution. *New Phytologist* **219**: 1170–1187.

Sentinella AT, Warton DI, Sherwin WB, Offord CA, Moles AT. 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**: 1387–1398.

Silvestro D, Kostikova A, Litsios G, Pearman PB, Salamin N. 2015. Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution* **6**: 340–346.

Smith SD, Pennell MW, Dunn CW, Edwards SV. 2020. Phylogenetics is the new genetics (for most of biodiversity). *Trends in Ecology & Evolution* **35**: 415–425.

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

Violle C, Enquist BJ, McGill BJ, Jiang L, Albert CH, Hulshof C, Jung V, Messier J. 2012. The return of the variance: intraspecific variability in community ecology. *Trends in Ecology & Evolution* **27**: 244–252.

Visscher AM, Vandelook F, Fernández-Pascual E, Pérez-Martínez LV, Ulian T, Diazgranados M, Mattana E. 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**: 773–784.

Walck JL, Hidayati SN, Dixon KW, Thompson K, Poschlod P. 2011. Climate change and plant regeneration from seed. *Global Change Biology* **17**: 2145–2161.

Willis CG, Baskin CC, Baskin JM, Auld JR, Venable DL, Cavender-Bares J, Donohue K, Rubio de Casas R. 2014. The evolution of seed dormancy: environmental cues, evolutionary hubs, and diversification of the seed plants. *New Phytologist* **203**: 300–309.

Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JHC, Diemer M *et al.* 2004. The worldwide leaf economics spectrum. *Nature* **428**: 821–827.

Key words: data synthesis, database, germination, open science, plant reproduction, repository, seed, trait.

Received, 9 June 2023; accepted, 27 June 2023.