Data Descriptor: Phenotypic trait variation measured on European genetic trials of *Fagus sylvatica* L.

T. Matthew Robson¹*, Marta Benito Garzón²* and BeechCOSTe52 database consortium #

We present BeechCOSTe52, a database of European beech (*Fagus sylvatica*) phenotypic measurements for several traits related to fitness measured in genetic trials planted across Europe. The dataset was compiled and harmonized during the COST-Action E52 (2006–2010), and subsequently cross-validated to ensure consistency of measurement data among trials and provenances. Phenotypic traits (height, diameter at breast height, basal diameter, mortality, phenology of spring bud burst and autumn–leaf discoloration) were recorded in 38 trial sites where 217 provenances covering the entire distribution of European beech were established in two consecutive series (1993/95 and 1996/98). The recorded data refer to 862,095 measurements of the same trees aged from 2 to 15 years old over multiple years. This dataset captures the considerable genetic and phenotypic intra-specific variation present in European beech and should be of interest to researchers from several disciplines including quantitative genetics, ecology, biogeography, macroecology, adaptive management of forests and bioeconomy.

| Design Type(s)                   | time series design |
|----------------------------------|--------------------|
| Measurement Type(s)              | phenotype characterization |
| Technology Type(s)               | measurement device |
| Factor Type(s)                   | geographic location, temporal_interval, age |
| Sample Characteristic(s)         | *Fagus sylvatica* ∙ Germany ∙ Kingdom of Spain ∙ Slovak Republic ∙ Kingdom of Denmark ∙ French Republic ∙ Italy ∙ Czech Republic ∙ Poland ∙ Romania ∙ Sweden ∙ Luxembourg ∙ Republic of Ireland ∙ United Kingdom ∙ Belgium ∙ Kingdom of the Netherlands ∙ Slovenia ∙ Croatia |

¹Viikki Plant Science Centre (ViPS), Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, P.O. Box 65, 00014, Finland. ²UMR 1202 Biodiversité Gènes & Communautés, INRA, Université Bordeaux, Bâtiment B21ée G St Hilaire, CS 50023, 33615 Pessac Cedex, France. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to T.M.R. (email: matthew.robson@helsinki.fi) or to G.M.B. (email: marta.benito-garzon@inra.fr).

#A full list of members appears in the Author Contributions.
**Background & Summary**

Currently climate change threatens to outpace forest migration, making trees' survival dependent on their intrinsic potential to adapt and acclimatise to new climatic conditions. Judging the extent of trees' capacity to cope successfully with new conditions requires knowledge of their tolerance limits; something that is hard to study and to-date cannot be estimated with confidence for many tree species. The long lifespan of forest trees means that we might expect to see a lag in their local adaptation to climate; it also makes experiments requiring comparisons across multiple generations difficult. Genetic trials planted for commercial purposes over recent decades probably hold the greatest potential to provide those data that will allow us to understand adaptive processes and the acclimation of trees to new climates. While analyses of individual trials comparing several provenances can inform us of those differences among populations that are shaped by genetics, only information gleaned from multiple trials covering an entire species range allow us to assess the relative importance of genetics and phenotypic plasticity, and how the two interact, for species survival under a changing climate.

European beech (*Fagus sylvatica* L.) is an important tree species for forestry and culturally. It covers roughly 14-million ha of forested land, ranging from Mediterranean to temperate ecosystems (http://www.euforgen.org). This wide ecological amplitude makes beech a good reference species for large-scale studies of plastic and adaptive responses in its fitness-related traits to climate change over its full distribution range.

Here, we present a consolidated set of phenotypic data from genetic trials of *Fagus sylvatica* across Europe compiled by the COST Action E52 (2006–2010). There have been several efforts to make use of phenotypic data from forestry trials of various species at the European level e.g. to compile metadata and to produce standardised protocols (http://www.trees4future.eu/ and http://www.trees4future.eu/transnational-accesses/treebreedex.html). However, raw data is seldom published, and initiatives to compile old genetic trials tend only to be promoted at the national level; e.g. GENFORED in Spain (www.genfored.es) and PLANTACOMP in France. The BeechCOSTe52 dataset (Data Citation 1) provides a high density map of phenotypic data from 38 genetic trials across Europe, even including trials established outside the range of the species. The wide coverage provided by this dataset makes it a singular resource that can help us to understand the entire gradient of climatic tolerance of European beech (Fig. 1). Likewise, the origin of the 217 provenances planted covers the entire range of the species distribution, which is important for understanding the overall adaptive capacity and the tolerance limits of the species (Fig. 1).

Genetic trials have historically been used in forestry to test which reproductive material to select (http://www.euforgen.org), but the utility of genetic trials in general and of this database in particular go beyond this traditional use. The most immediate use of the BeechCOSTe52 database is in understanding of trade-offs between traits from an ecological perspective. Comparable trials over a larger geographical area covering the range limits of a species can be helpful in assessing local adaptation in populations of trees that have grown nearby during the recent past. They can also be useful for testing the fitness of range-edge populations or those derived from isolated refuges during the last glaciation, which may hold valuable traits that can confer fitness benefits under particular circumstances. Over small geographical areas, the effects of local environmental conditions and past management can be tested and pave the way to guide new innovative management practices involving assisted gene flow and the translocation of populations to compensate lower productivity with higher survival when mitigating climate change and the socio-economic consequences of planting exotic versus local provenances.

A complete network of genetic trials is also of interest for the conservation of genetic resources, in particular those from marginal populations which can hold higher genetic variation than those populations at the core of a distribution. Finally, we hope that this database will encourage researchers and foresters to compile and publish genetic trial data of other species, since databases such as this are an invaluable resource helping us understand the capacity of forests to adapt and acclimatise to climate change.

**Methods**

**Tree plantation in genetic trials**

The international series of genetic trials presented in the BeechCOSTe52 database were established in two different series (years sown/trial planted: 1993/95 and 1996/98) to study genetic variation of European beech across its distribution range. The planning of the international series of beech trials was principally coordinated by researchers from the Institute of Forest Genetics in Grosshansdorf, Germany, including Mirko Liesebach, Georg von Wuehlisch and Hans Muhs, in liaison with the international trial-site holders, and supported by the Concerted Action of the Commission of the European Communities (AAAIR3 Programme, Grant No. CT94-2091). Previous plantations of genetic trials of beech exist but mostly with local provenances, which precludes the study of genetic variation across the species range. The trials were planted in three complete blocks, each block containing all of the provenances planted at each trial site. In total 150 seedlings (two-year old) of each provenance were planted in randomised groups of 50 seedlings within each block, using a spacing of 1 × 0.2 m. Border rows of a local provenance were used to reduce edge effects. More details of the planning, design and planting of the provenance trials are given in a document presented at the IUFRO meetings on beech.
Phenotypic traits measurement
Phenotypic trait measurements included tree height, basal diameter, diameter at breast height, mortality, bud burst phenology in spring and autumn leaf senescence, and in some cases the causes of damage related to mortality. Trait data was collected on an individual-trees basis in consecutive years following plantation, including measurements from 2-to-15 year-old trees, thus providing information through different ontogenic stages. We give an example of how the database can be plotted and further analysed according to trait and ontogenic stage (Fig. 2).

Data compilation
A succession of national and international projects have allowed the collection of data from the trial sites (Additional Information.zip, Data Citation 1 and 1st Meeting9). In 2004, various preliminary databases of individual French, German and Italian provenances were gathered during the DYNABEECH project (Additional Information.zip, Data Citation 1).

Later on, a European Concerted Research Action designated as COST Action E52 "Evaluation of Beech Genetic Resources for Sustainable Forestry" came into force in 2006, running until 2010, in order to consolidate and interpret the data collected from the two series of beech trials set up in 1995 and 1998, and to bring together the participants from countries that were managing trials to pool their expertise. The stated aim of the COST Action was "to make predictions of the future distribution range of beech forest ecosystems under the assumption of certain scenarios of climate change, based on the analysis of the reaction pattern of European beech populations".

An introductory COST Action E52 Working Group and Management Committee meeting took place in Zvolen (Slovakia 2006), with plenary meetings during the subsequent three years as well as various Working Group meetings10. Some partial analyses were done during this period, considering one11–18 or several trials19–23.

One striking quality of the BeechCOSTe52 database presented here is that it has been made possible through the coordinated effort of researchers and foresters from many countries across Europe, who have managed and maintained the trial sites over an extended period of time and have altruistically contributed to the efforts of data collection and coordination over a long time period. The compilation of the phenotypic information from all the genetic trials included in the BeechCOSTe52 database only refers to trials and provenances in which traits have been measured repeatedly (Table 1). The database includes 862,095 trait records measured in 38 trial sites and a total of 217 provenances varying from 13 to 101 among trial sites. We kept only traits measured in at least six trial sites to assure a good coverage of the distribution of the species. Measurements are heterogeneously repeated over years and trees, from 1995 to 2008.
Code availability
R code to merge phenotypic trait measurements, geographical coordinates of the trials and geographical coordinates of the provenance locations is included (merge-trial-prov-coords_individualmeasurements.R). Mapping phenotypic traits by provenances and trials as shown in Figs. 1 and 2 can be reproduced using the following scripts: Hplot.R & Plot_Trial-Prov_locations.R. Likewise, individual data by trial and trait can be extracted from the global database using the specific script Subset_by_trial.R and Subset_by_trait.R, respectively.

Data Records
The database is structured in three independent files that can be merged using the code identifier for the provenance and the code combining the trial and provenance codes ("Trial" and "ID_ProvCode", respectively). These data files together with the metadata descriptor document are available at Zenodo data repository (BeechCOSTe52_metadata_descriptor.docx, Data Citation 1). The first file contains individual measurements of phenotypic traits (F.sylvatica.csv, Data Citation 1): height, basal diameter, diameter at breast height, mortality, spring phenology, autumn leaf phenology, and associated damage causing to tree mortality. The second file contains the description of the trials including their geographical locations (Trials_coord.csv, Data Citation 1) and the third file describes the locations of the seed sources (Prov_coord.csv, Data Citation 1).

Technical Validation
The database has been checked for consistency at different stages by various researchers. During the COST Action E52 (2006 to 2010) raw data were submitted by each country and harmonized in electronic format for the first time. Results from the individual trial sites were sent to Georg von Wuehlisch until 2008. In 2008 and 2009 this information was assimilated into electronic form and standardised by Diana Barba and T. Matthew Robson. At that time, it was decided to have one database by genetic trial, harmonised among trials and checked for consistency. The geographical locations, names and identities of trial sites and provenances were rechecked, and the management and biogeographical history of provenances considered. In a workshop (19th-23rd January 2009) the data evaluation working group met in Valsain, Spain, where the database was checked for completeness and consistency of data collected by the members of the COST Action E52. Then, members of the COST Action E52 had one year to present their findings prior to the 2010 final Cost Action E52 meeting in Burgos, where various studies of the provenance trials were presented, and later published in the Proceedings of the Meeting (Additional Information.zip, Data Citation 1).
The studies reported that the data gathered expose high variability in fitness-related traits such as height, both among provenances and among trials. Likewise, there was large variation in the timing of bud burst and leaf flush in spring, when data collected using different scales among sites were harmonised and considered both in terms of day-of-the-year and degree-hours of forcing prior to bud burst.

Finally, in 2017, the originally-compiled trial-site database was checked again for consistency of entries and fields in R by Marta Benito Garzón prior to harmonization of all the trial sites into the single

**Table 1. Summary table showing the information contained in the dataset by trial site.** The summary information refers to trait measurements in one or several years and can include several trait measurements per tree. PROV = number of provenances planted. Height = number of height measurements. DBH = number of diameter at breast height measurements. BasalD = number of basal diameter measurements. Mort = number of dead trees. SP = number of bud burst measurements. AP = number of records of autumn leaf discolouration. D = number of damage measurements. Zero values indicate no data available.
phenotypic database presented here. Outliers and errors in the database were checked by calculating descriptive statistics of quantitative traits for each trial averaged by provenance.

References
1. Pedlar, J. H. & McKenney, D. W. Assessing the anticipated growth response of northern conifer populations to a warming climate. *Sci. Rep.* 7, 703881 (2017).
2. Saenz-Romero, C. et al. Adaptive and plastic responses of *Quercus petraea* populations to climate across Europe. *Glob. Change Biol.* 23, 2831–2847 (2017).
3. Siepieński, A. M. et al. Precipitation drives global variation in natural selection. *Science* 355, 959–962 (2017).
4. Valladares, F. et al. The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecol. Lett.* 17, 1351–1364 (2014).
5. Steinbach, D. et al. GnpIS: an information system to integrate genetic and genomic data from plants and fungi. *Database* 2013 bat058, 1–9 (2013).
6. Fady, B. et al. Forests and global change: what can genetics contribute to the major forest management and policy challenges of the twenty-first century? *Reg. Environ. Change* 16, 927–939 (2016).
7. Bontemps, A., Lefèvre, F., Davi, H. & Oddou-Muratorio, S. In situ marker-based assessment of leaf trait evolutionary potential in a marginal European beech population. *J. Evol. Biol* 29, 514–527 (2016).
8. von Wuelichs, G. Improvement and Silviculture of Beeches. In *Proceedings from the 7th International Beech Symposium IUFRO Research Group 135–144 (IUFRO,2007)*.
9. Turok, J., Kremer, A. & de Vries, S. First EUFORGEN Meeting on Social Broadleaves. 1-177 (IPGRI,1998).
10. Fryd, J., Novotný, P., Fennessey, J. & von Wühlisch, G. (eds) Implementing output of COST Action E 52 Project ‘Evaluation of beech genetic resources for sustainable forestry’ (2006–2010). *Commun. Inst. For. Bohemicae* 25 (2010).
11. Gömöry, D. & Paule, L. Trade-off between height growth and spring flushing in common beech (*Fagus sylvatica* L.). *Ann. For. Sci.* 68, 975–984 (2011).
12. Gömöry, D. et al. Differentiation in phenological and physiological traits in European beech (*Fagus sylvatica* L.). *Eur. J. For. Res.* 134, 1075–1085 (2015).
13. Hofmann, M., Durka, W., Liesebach, M. & Brueleheide, H. Intraspecific variability in frost hardiness of *Fagus sylvatica* L. *Eur. J. For. Res.* 134, 433–441 (2015).
14. Wortemann, R. et al. Genotypic variability and phenotypic plasticity of cavitation resistance in *Fagus sylvatica* L. across Europe. *Tree Physiol.* 31, 1175–1182 (2011).
15. Robson, T. M., Rasztovits, E., Aphalo, P. J., Alia, R. & Aranda, I. Flushing phenology and fitness of European beech (*Fagus sylvatica* L.) provenances from a trial in La Rioja, Spain, segregate according to their climate of origin. *Agric. For. Meteorol* 180, 76–85 (2013).
16. Eilmann, B. et al. Wood structural differences between northern and southern beech provenances growing at a moderate site. *Tree Physiol.* 34, 882–893 (2014).
17. Hajek, P., Kurjak, D., von Wühlisch, G., Delzon, S. & Schuld, B. Intraspecific Variation in Wood Anatomical, Hydraulic, and Foliar Traits in Ten European Beech Provenances Differing in Growth Yield. *Front. Plant Sci.* 7, 791 (2016).
18. Horvath, A. & Mátayás, C. The decline of vitality caused by increasing drought in a beech provenance trial predicted by juvenile growth. *South-East Eur. For.* 7, 21–28 (2016).
19. Mátayás, C., Bozic, G., Gömöry, D., Ivanikovic, M. & Rasztovits, E. Transfer analysis of provenance trials reveals macroclimatic adaptedness of European beech (*Fagus sylvatica* L.). *Acta Silv. Lign. Hung* 5, 47–62 (2009).
20. Mátayás, C., Bozic, G., Gömöry, D., Ivanikovic, M. & Rasztovits, E. Juvenile growth response of European beech (*Fagus sylvatica* L.) to sudden change of climatic environment in SE European trials. *IForest - Biogeosciences For.* 2, 213–220 (2009).
21. von Wuelichs, G. & Alia, R. COST E52 Final Report-Genetic Resources of European Beech (*Fagus sylvatica* L.) for sustainable forestry. *Proceedings of the COST E52 Evaluation of beech genetic resources for sustainable forestry*. 1-148 (Monografías INIA: Serie Forestal 22, 2011).
22. Kramer, K. et al. Chilling and forcing requirements for foliage bud burst of European beech (*Fagus sylvatica* L.) differ between provenances and are phenotypically plastic. *Agric. For. Meteorol* 234, 172–181 (2017).
23. Schueller, S. & Liesebach, M. Latitudinal population transfer reduces temperature sum requirements for bud burst of European beech. *Plant Ecol.* 216, 111–122 (2015).
24. Alia, R. et al. The survival and performance of beech provenances over a Europe-wide gradient of climate. in *COST E52 Final Report-Genetic Resources of European Beech (Fagus sylvatica L.) for sustainable forestry. Proceedings of the COST E52 Evaluation of beech genetic resources for sustainable forestry* 115–126 (Monografías INIA: Serie Forestal 22, 2011).
25. Robson, T. M. et al. The timing of leaf flush in European beech (*Fagus sylvatica* L.) saplings. in *COST E52 Final Report-Genetic Resources of European Beech (Fagus sylvatica L.) for sustainable forestry. Proceedings of the COST E52 Evaluation of beech genetic resources for sustainable forestry* 61–80 (Monografías INIA: Serie Forestal 22, 2011).

Data Citations
1. Robson, T. M. et al. Zenodo https://doi.org/10.5281/zenodo.1240931 (2018).

Acknowledgements
The establishment of the trial series was supported by the Concerted Action of the Commission of the European Communities (AAAIR3 Programme, Grant No. CT94-2091). The compilation and harmonization of the database was supported by the COST Action E52 (2006-2010). We are very grateful to all the technicians and researchers that measured trials over years in the field. The final compilation was funded by the Investments for the Future IDEX Bordeaux (ANR-10-IDEX-03-02).

Author Contributions
T.M.R. compiled and checked the original trial by trial data. T.M.R. and M.B.G. compiled all the individual information from each trial to a single database, checked the data and were largely responsible for writing the manuscript. The other authors (BeechCOSTe52 database consortium) provided the single trial information, contributed to harmonising and formatting the data at an early stage and participated in writing the final manuscript.
BeechCOSTe52 database consortium

Ricardo Alia Miranda¹, Diana Barba Egido², Saša Bogdan³, Attila Borovics⁴, Gregor Božič⁵, Oliver Brendel⁶, Jo Clark⁷, Sven M.G. De Vries⁸, Ivan I Delehan⁹, Alexis Ducousoo¹⁰, Bruno Fady¹¹, John Fennessey¹², Manfred Forstreuter¹³, Josef Fryd³¹, Thomas Geburek¹⁴, Dušan Gömöry¹⁷, Maria Hauke-Kowalska¹⁸, Gerhard Huber¹⁹, Juan-Ignacio Ibañez²⁰, Lucia Ionită²¹, Mladen Ivanković²², Jon Kehlet Hansen²³, Anikó Köczán-Horváth²⁴, Hojka Kraigher²⁵, Steve Lee²⁶, Mirko Liesebach²⁷, Csaba Mátyás²⁸, Patrick Mertens²⁷, Hans-Jakob Muhs²⁹, Petr Novotný¹⁵, Gheorghe Parnuta²¹, László Paule¹⁷, Álvaro Picardo²⁸, Ervin Rasztovics³, Martin Rogge²⁹, Lars-Göran Stener³⁰, Malgorzata Sulikowska³¹, Otmar Urban³², Georg Von Wuehlisch³³, Giovanni V Vendramin³⁴, Cristina Vettori³⁵ & Wojciech Wesoly³⁶

Additional Information

Competing interests: The authors declare no competing interests.

How to cite this article: Robson, T. M. et al. Phenotypic trait variation measured on European genetic trials of Fagus sylvatica L. Sci. Data 5:180149 doi: 10.1038/sdata.2018.149 (2018).

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

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

The Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files made available in this article.

© The Author(s) 2018