

SOFTWARE

Open Access



A centralized interactive tool for genomic data on heat shock protein 90 in Fagaceae species

Rafael Postigo-Luque¹, David Guijo-Rubio², M. Ángela Martín¹, Patricia Castro¹ and José V. Die^{1*}

Abstract

Background The Fagaceae family includes key forestry species in temperate and Mediterranean ecosystems, many of which are increasingly impacted by drought stress and high temperatures, exacerbated by climate change. HSP90 proteins play a crucial role in response to that kind of stress, but genomic information related to these proteins in Fagaceae species are scattered in diverse databases, making integrated analysis challenging.

Results We present HSP90-Fagaceae, an interactive web application built with the Shiny package in R that centralizes curated HSP90-related genomic, structural, and phylogenetic data across major Fagaceae species. Accessible online, the application provides users with an intuitive platform to explore and compare HSP90 protein sequences, annotations, and evolutionary relationships within a gene family-specific framework.

Conclusions The HSP90-Fagaceae app integrates previously fragmented genomic data into a single, interactive, open-source resource that supports structural and evolutionary analyses of the HSP90 gene family across different Fagaceae species. Its modular design also offers a flexible framework that can be adapted to develop similar applications for other protein families or species of agricultural and forestry interest.

Keywords R Shiny, Heat shock protein 90, Fagaceae, Drought stress, Bioinformatic pipeline

Background

Forests play a crucial environmental role by supplying essential ecosystem services such as carbon sequestration, water cycle regulation, and the maintenance of biodiversity [1]. The Fagaceae family—comprising prominent genera such as *Quercus* (oaks), *Castanea* (chestnuts), *Fagus*, *Castanopsis*, *Lithocarpus*, and *Notholithocarpus*—includes many species for which increasingly comprehensive genomic resources are becoming

available. This family holds substantial ecological and economic value, particularly for breeding programs aimed at selecting individuals with improved tolerance to environmental stressors. Climate change is profoundly impacting forest ecosystems worldwide with recurrent summer droughts and elevated temperatures intensifying abiotic stresses that threaten the health and resilience of Fagaceae species [2, 3].

Heat shock proteins (HSPs) are an important class of molecular chaperones that maintain cellular balance by protecting essential proteins, enabling plants to withstand adverse conditions [4]. Although HSPs play a vital role in response to different abiotic stresses in plants, their role is especially prominent in relation to heat stress. HSPs have been classified into five categories on the basis of their molecular weight. Among them, HSP90 proteins are particularly relevant due to their role in

*Correspondence:

José V. Die

jose.die@uco.es

¹Departamento de Genética-ETSIAM, Universidad de Córdoba, Campus de Rabanales, Córdoba 14071, Spain

²Departamento de Ciencia de la Computación e Inteligencia Artificial-Escuela Politécnica Superior de Córdoba, Universidad de Córdoba, Campus de Rabanales, Córdoba 14071, Spain



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

regulating cellular processes related to heat and drought stress response [5]. Recently, large-scale phenotypic analyses using HSP knockout mutants revealed that many HSPs are necessary to enhance heat stress tolerance [6]. Additionally, the knockdown of drought-inducible HSP90 in cassava resulted in decreased drought stress tolerance [7]. In woody species, it has also been shown that HSP90 genes are among the differentially expressed genes (DEGs) associated with chaperone activity, indicating a rapid early response to drought and temperature stress [8]. Together these studies highlight the biological relevance of HSP90 proteins while underscoring the need for robust genomic and evolutionary characterization to support downstream functional investigations.

Despite its importance, knowledge and characterization of the HSP90 family in Fagaceae remain limited. Although substantial genomic resources are available for species within this family, they are scattered across multiple repositories, making integration and access challenging. The study of a specific gene family, such as HSP90, often requires additional steps of searching, downloading, and filtering sequences of interest across multiple repositories. Generalist genomic databases such as Phytozome [9] provide broad coverage and are indispensable resources for molecular plant research. However, their scope is not tailored for gene family-specific curation, comparative analysis, or integrated phylogenetic exploration. In addition, not all Fagaceae taxa are represented in these platforms, which further fragments the available information.

Gene family-centric repositories have demonstrated their value, such as HSPMdb [10], which compiles molecular information on heat shock proteins across 15 diverse model species, including human, yeast, bacteria, and protozoa. However, comparable resources focusing on woody plant species or forest tree families remain scarce. In our laboratory, we are interested in characterizing the HSP90 family within Fagaceae. Motivated by the need for a curated, family-specific resource that integrates structural, phylogenetic, and motif information in a single interactive environment, we developed HSP90-Fagaceae, an interactive Shiny application that centralizes HSP90 genomic information from ecologically and economically important Fagaceae species. This tool is intended as a structural and evolutionary reference resource that facilitates access to curated data and supports comparative and hypothesis-generating studies in forest genetics and evolutionary biology.

Implementation

HSP90-Fagaceae was developed to address the need to centralize dispersed genomic information and provide accessible tools to analyze HSP90 proteins in species of the Fagaceae family.

Data collection and integration

HSP90-Fagaceae integrates amino acid sequences of HSP90 proteins from public repositories including NCBI [11], NGDC Genome Warehouse [12], Phytozome [9], the *Castanea* genome database [13] and TreeGenes [14]. Detailed information on each species' genome assembly, data source, accession numbers, and taxonomy identifiers is provided in Supplementary Table 1. In a seminal study on stress tolerance in *Quercus* spp., the authors investigated both species-specific and genus-wide genetic responses to water stress, identifying key molecular pathways associated with drought tolerance. Among the candidate genes, members of the HSP90 family featured prominently, along with the drought-induced *ACX1* gene, which is known to interact directly with HSP90 proteins. These findings underscore the crucial role of protein folding and protective mechanisms in the drought stress response [8]. Our research group, which focuses on the transferability of genomic information from *Quercus* spp. to European chestnut (*Castanea sativa*), is developing a genomic atlas of drought adaptation. This atlas provides valuable insights into the genetic basis of drought tolerance and supports the development of molecular markers for more targeted and effective breeding and conservation strategies in chestnuts. Among the candidate genes, we identified three HSP90 homologs with high similarity to sequences from *Quercus suber*: XP_023879721, XP_023890186, and XP_023927413 [15]. In the present study, we used these HSP90 sequences as queries in BLASTp searches against the *Q. suber* genome to further characterize the HSP90 gene family within the species. Next, blast hits were filtered using the NCBI CD-Search tool to verify the presence of both the conserved HSP90 family domain (Pfam: PF00183) and the histidine kinase-like ATPase domain (cd16927, NCBI CDD). Sequences lacking either of these two essential domains were excluded from the database. However, sequences containing both domains in partial or truncated form were retained but explicitly flagged in the database to inform users of their structural integrity. For genes with multiple isoforms, only the longest protein-coding isoform was retained in order to standardize the dataset and avoid redundancy in downstream analyses. Because the genomes used in this study are assembled at the chromosomal level, the retrieved HSP90 sequences correspond to nuclear-encoded genes. This approach yielded 11 HSP90 sequences for *Q. suber*. Using the 11 HSP90 sequences from *Q. suber* as initial queries, BLASTp searches were conducted across various genomes. The selection thresholds were set at a minimum of 30% sequence identity and 50% query coverage to account for interspecific divergence. This approach was independently applied to each species included in the study. Candidate HSP90 sequences identified through

BLASTp were subsequently validated via Hidden Markov Model profile confirmation and domain analysis using NCBI's Conserved Domain Database (CDD) tool, ensuring accurate identification of HSP90 family members. The HSP90 characterization is reported for the following species: *Castanea crenata*, *Castanea dentata*, *Castanea mollissima* 'Vanuxem', *Castanea mollissima* 'N-11', *Castanea mollissima* 'H7', *Castanopsis carlesii*, *Castanopsis eyrei*, *Castanopsis hystrix*, *Fagus crenata*, *Fagus grandifolia*, *Fagus sylvatica*, *Lithocarpus litseifolius*, *Lithocarpus polystachus*, *Notholithocarpus densiflorus*, *Quercus acutissima*, *Quercus dentata*, *Quercus gilva*, *Quercus lobata*, *Quercus mongolica*, *Quercus robur*, *Quercus rubra* and *Quercus variabilis*. Although *Quercus ilex* lacks annotated protein files, it was specifically included in this study due to its high ecological and economic relevance in the Mediterranean basin [16]. As one of the most representative evergreen oaks in the region, *Q. ilex* plays a key role in forest resilience under climate change. For this reason, an alternative strategy was employed to obtain the HSP90 protein sequences. The coding sequences (CDS) of previously identified HSP90 genes from *Q. suber* were used as queries in a BLASTn search against the *Q. ilex* genome. The resulting nucleotide sequences were then translated into amino acid sequences. Once all HSP90 proteins were obtained, functional annotation was performed using the Blast2GO software [17]. Structural information, including chromosomal location and strand orientation, was extracted from the genome annotation files (GFF) corresponding to each species. Additionally, subcellular localization was predicted using the online tool DeepLoc 2.0 [18].

Taxonomic and habitat information for each species was obtained from scientific literature and electronic resources such as EUFORGEN [19]. Species images were sourced from the Pl@ntNet and Natusfera image collections. Species description and protein information were stored in a data frame for visualization in the application. Habitat information was used to generate vector files using QGIS, which were processed to create detailed maps of the global distribution of each species.

To have a better understanding of the evolutionary relationships of the HSP90 proteins of the species under study, a phylogeny tree was generated with Mega X software (v.10.2.6), using the Maximum Likelihood method with 1,000 replicates. The sequences were aligned using the MUSCLE algorithm integrated in MEGA X. To make the tree more robust, a threshold of 50% was set for bootstrap support values, thus eliminating nodes of lower reliability. The complete workflow used for database construction and integration is summarized in Fig. 1. In addition, Supplementary Table 2 provides a detailed quantitative overview of the number of sequences

retrieved, filtered, and retained for each species at every step of the pipeline.

Application design

HSP90-Fagaceae was developed using the R *Shiny* package [20]. The *bslib* package [21] was used to organize the application by different modules to improve its navigation. HSP90-Fagaceae relies on several packages for data generation and visualization [*leaflet* [22], *DT* [23], *msa* [24], *msar* [25]]. *Leaflet* allows the visualization of interactive maps, which can be integrated with shapefiles. The *DT* package allows displaying stored data in interactive tables. The packages *msa* and *msar* are respectively used to align sequences and represent them graphically. These packages, along with others, have been implemented in various tools of the application.

This application is hosted on a remote server and can be accessed at: <https://hsp90.ext.uco.es/>

Features

HSP90-Fagaceae integrates several modules with different functions, which are summarized in Fig. 2:

1. *Information on Fagaceae*: This section provides taxonomic descriptions of the main species of the Fagaceae family, accompanied by images and an interactive map showing their natural habitat and areas where they have been introduced.
2. *HSP90 database*: The module includes a table characterizing HSP90 proteins listed in the previous section. Users can customize the view by selecting/deselecting columns, filter by names, and download the table content as a CSV file. In addition, the table provides direct access to protein sequences in FASTA format.
3. *HSP90 amino acid sequences*: A module that provides the protein sequences from Module 2, which can be downloaded in FASTA format.
4. *Phylogenetic tree*: This module provides an interactive implementation of the HSP90 phylogeny through integration with the iTOL platform. Users can zoom, navigate, collapse or expand clades, search for specific identifiers, and visualize bootstrap support values directly within the app. The tree can also be customized and exported either as an image or in Newick format.
5. *BLASTp*: It enables local protein searches by homology using NCBI's BLASTp [26] against a centralized collection of proteomes from all studied species, allowing users to explore candidate homologs within a single integrated environment. Results are displayed in an interactive table, where individual entries can be selected to view alignments.

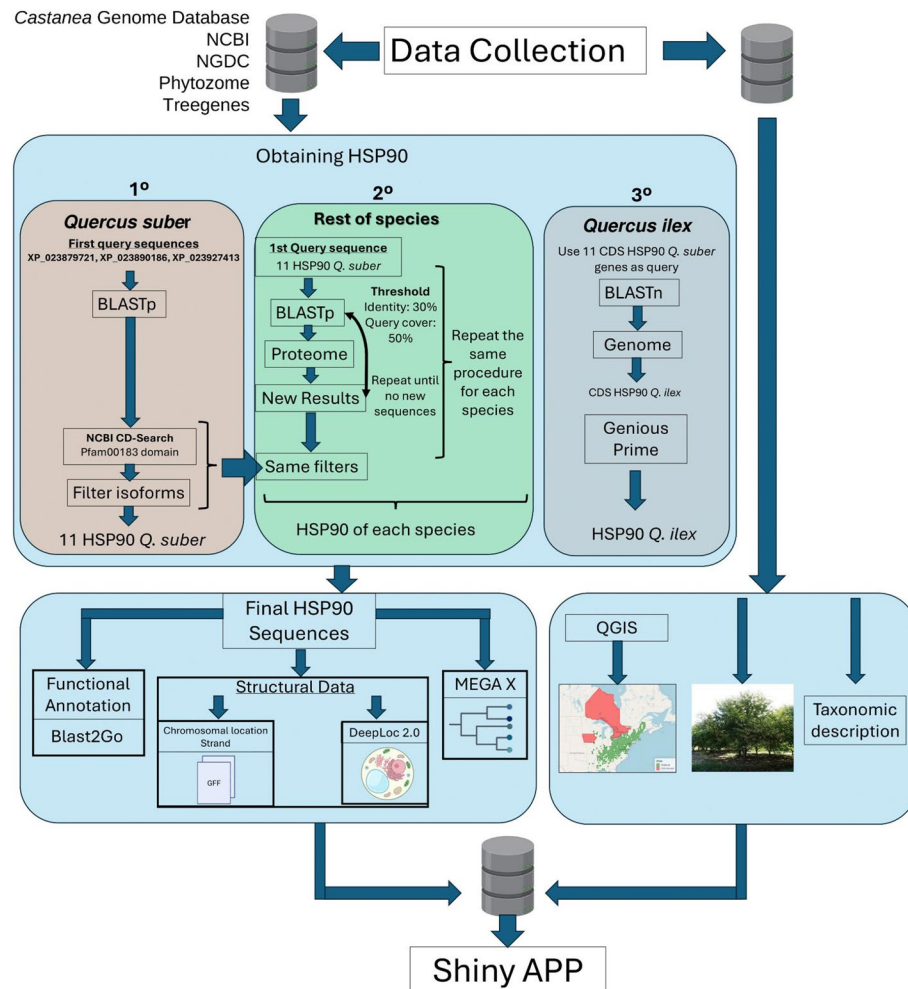


Fig. 1 Workflow followed to generate the HSP90 gene database for the Shiny application Fagaceae-HSP90

6. *Sequence aligner*: The module allows the alignment of amino acid and nucleotide sequences using three algorithms: ClustalW, ClustalOmega, and Muscle. The alignments can be downloaded in FASTA format.

Phylogenetic relationships of HSP90 proteins in Fagaceae

The phylogenetic reconstruction based on full-length HSP90 protein sequences resolved the dataset into seven well-differentiated clades (Groups I–VII) with robust bootstrap support (Fig. 3). All clades included representatives from each of the Fagaceae genera, indicating that the diversification of HSP90 subfamilies predates the radiation of the family.

These phylogenetic clades broadly correspond to the canonical HSP90 subfamilies described in plants and reflect their cellular specialization. Groups I–IV collectively represent the cytosolic HSP90 subfamily, which in plants comprises various paralogous lineages characterized by the conserved C-terminal MEEVD motif. Within this cytosolic compartment, Groups I and II include

proteins historically annotated as HSP80, while groups III and IV contain proteins annotated as HSP83; both designations reflect molecular weight. Group IV represents the largest and most expanded cytosolic clade in Fagaceae, consistent with lineage-specific diversification patterns. In contrast, groups I–II show a moderate expansion pattern, indicating differential evolutionary trajectories.

Groups V, VI and VII correspond to the organelle targeted HSP90 subfamilies. Group V comprises endoplasmic reticulum localized HSP90.7 in plants, characterized by the presence of C-terminal KDEL or HDEL retention motifs. Group VI corresponds to chloroplast localized HSP90.5, whereas group VII represents mitochondrial HSP90.6, forming a distinct and well supported phylogenetic clade consistent with their mitochondrial location.

Although the main subfamilies were conserved across species, lineage-specific patterns were observed. For example, *Quercus* and *Castanea* proteins tended to form tighter clusters within Groups III and IV, suggesting lineage-specific expansion patterns that may warrant further functional investigation. In contrast, *Fagus* and

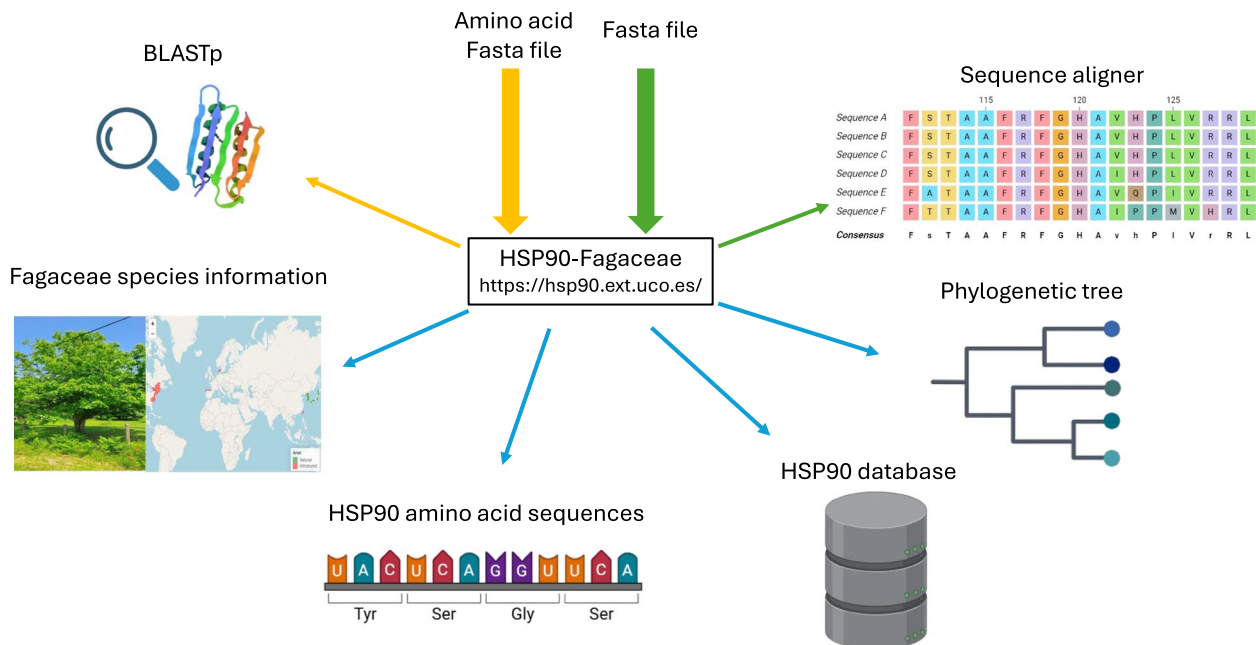


Fig. 2 Overview of the modules and the main features of the HSP90-Fagaceae app. The thick colored arrows indicate the required input data for the corresponding modules: amino acid Fasta file (yellow) and either an amino acid or nucleotide FASTA file (green). The modules marked with blue arrows do not require additional input data for execution

Lithocarpus sequences were more dispersed, indicating higher divergence within these lineages. Such patterns reflect both the ancient conservation and subsequent diversification of HSP90 subfamilies in Fagaceae.

Discussion and potential applications

HSP90-Fagaceae is a bioinformatics tool designed to address a common challenge in genetic studies: the fragmentation of data across multiple web servers. The dispersion hinders both access and downstream analysis, as exemplified by the case of HSP90 proteins in the Fagaceae family. This application resolves this issue by centralizing relevant information, making key data more accessible and streamlining research workflows. In addition, it provides download options to support interoperability with third-party software, facilitating more advanced and flexible analysis.

By combining structural annotations, phylogenetic relationships, and species distribution data, HSP90-Fagaceae supports comparative and evolutionary analyses within a gene family-centric framework. Rather than functioning as a comprehensive functional genomics platform, the application is intentionally designed as a structural and evolutionary reference resource. Functional dimensions such as transcriptomic profiling, stress-responsive expression analyses, co-expression networks, or selection pressure inference are not included in the current implementation. This focused scope reflects the heterogeneous availability of functional datasets across forest tree species and ensures consistency and reliability of the curated

information across all taxa included. The methodology used to develop this application can be systematically applied to other gene families with scattered information, as well as to agronomic or forestry species of importance for breeding programs or economic value.

Several general-purpose bioinformatics tools are available for gene family analyses, including widely used platforms such as TBtools and SPDEV3.0. These tools offer extensive analytical capabilities and flexibility, particularly for users working with custom datasets. However, these typically require users to independently collect, curate, and format gene family-specific data prior to analysis. In contrast, HSP90-Fagaceae does not aim to replace such tools or introduce new analytical algorithms, but rather to complement them by providing a pre-curated, gene family-specific and species-focused resource. Its contribution lies in the integration, curation, and biological organization of dispersed genomic information into a single interactive framework. By providing pre-curated datasets, standardized annotations, and an embedded phylogenetic context, the platform simplifies analytical workflows, lowers the entry barrier for non-specialist users, and enhances reproducibility within a narrowly defined biological scope.

The phylogenetic component incorporated into HSP90-Fagaceae provides essential biological context for the curated dataset. The resolution of seven distinct clades corresponding to the major cytosolic and organelle-targeted HSP90 subfamilies described in plants highlights both the deep conservation of the family

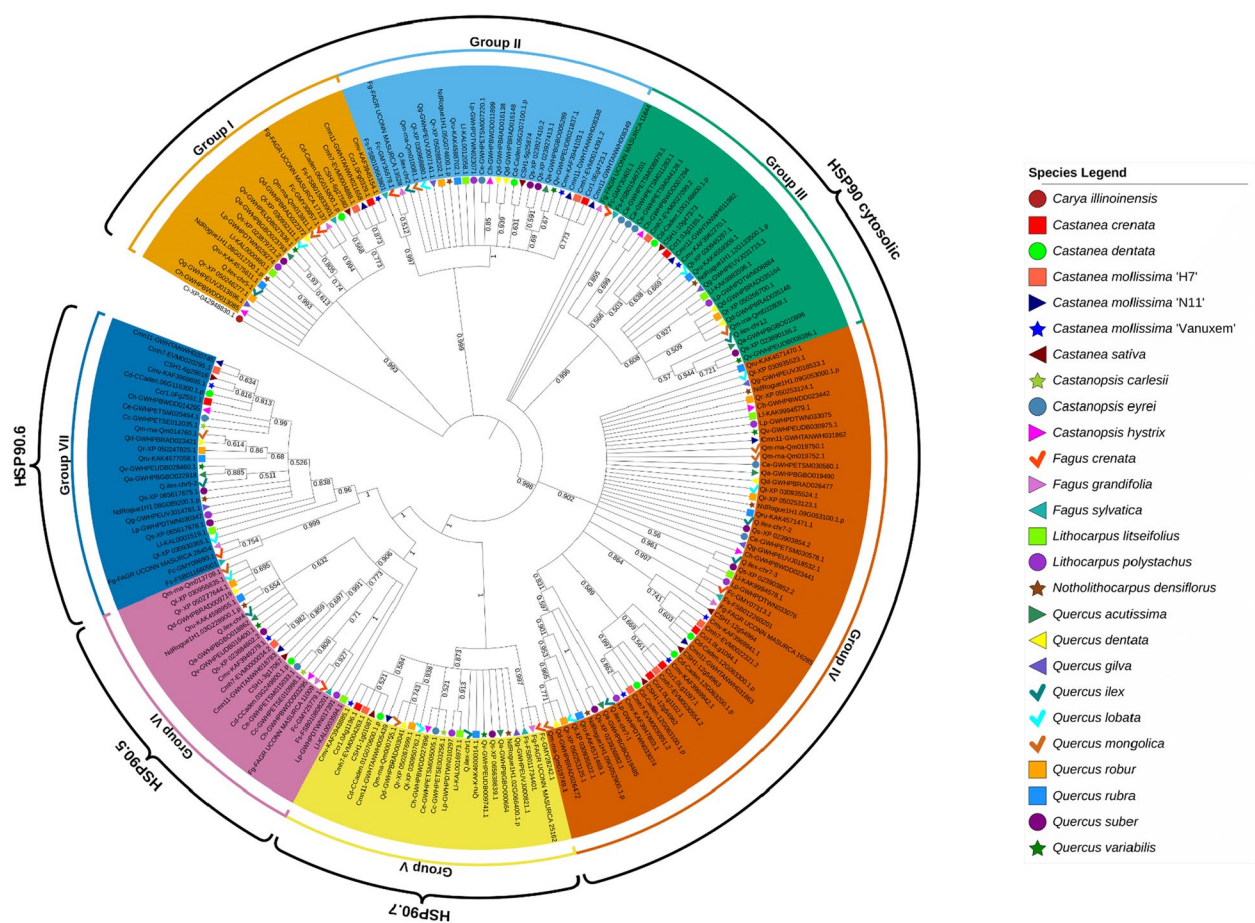


Fig. 3 Phylogenetic tree of HSP90 proteins across Fagaceae species. The tree was constructed using the Maximum Likelihood method (1,000 bootstrap replicates) with a threshold of 50% for bootstrap support values. Colored symbols correspond to species as indicated in the legend. The phylogeny resolves seven major clades (Groups I–VII). Groups I–IV collectively represent the cytosolic HSP90 subfamily. Groups V, VI and VII correspond to the organelle-targeted HSP90 subfamilies HSP90.7 (endoplasmic reticulum), HSP90.5 (chloroplast) and HSP90.6 (mitochondria), respectively

and its lineage-specific diversification within Fagaceae. Observed clustering patterns, including apparent expansions in *Quercus* and *Castanea*, generate testable hypotheses regarding the evolutionary dynamics of HSP90 genes in forest trees exposed to diverse environmental conditions. In this way, HSP90-Fagaceae not only enables efficient data retrieval and comparison but also serves as a foundation for hypothesis driven functional and ecological studies conducted using complementary datasets and analytical frameworks.

Conclusion

HSP90-Fagaceae is a bioinformatics tool that centralizes genomic information on HSP90 proteins from the Fagaceae family. Its intuitive and user-friendly design enables more efficient analysis of these proteins, facilitating the study of their role in drought stress tolerance. The application serves as a model for developing similar tools in projects that handle genomic information dispersed across multiple databases. By centralizing information and offering download options, HSP90-Fagaceae

improves data accessibility and streamlines bioinformatics workflows.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-026-12697-9>.

- Supplementary Material 1.
- Supplementary Material 2.

Acknowledgements

We thank Pl@ntNet and Natusfera for the images used. We also acknowledge EUFORGEN for making available online the shapefiles containing geographic distribution data of European Fagaceae species, which were used in this study.

Availability and requirements

Project name: HSP90-Fagaceae.
Project home page: <https://github.com/AGRI14molecularBreeding/castanea>.
This tool is also available online at: <https://hsp90.ext.uco.es/>
Operating system (s): Multiplatform.
Programming language: R.
Other requirements (locally): Dependent on R packages and files available in the Github repository.
License: MIT License.
Any restrictions to use by non-academics: None.

Authors' contributions

R.P.-L. contributed to **Software**, **Writing - original draft**, and **Data curation**. D.G.-R. was responsible for **Software**, **Visualization**, and **Deployment**. P.C. and M.A.M. contributed to **Writing - review & editing**, revising the manuscript critically for important intellectual content. J.V.D. contributed to **Conceptualization**, **Funding acquisition**, **Supervision**, and **Writing - review & editing**. All authors read and approved the final manuscript.

Funding

This work is funded by the project ProyExcel_00351 (Proyectos de investigación de Excelencia, Junta de Andalucía). DGR thanks the project of the Universidad de Córdoba and the Andalusian Regional Government with reference PP2F_L1_15. JVD is a Ramón y Cajal postdoctoral fellow supported by the program MCIN/AEI/10.13039/501100011033 (Ref. RYC2019-028188-I).

Data availability

All data and files used to develop the application are available in the following public GitHub repository: [<https://github.com/AGR114molecularBreeding/castanea>] (<https://github.com/AGR114molecularBreeding/castanea>). The repository includes proteome files originally obtained from NCBI, NGDC, Phytozone, and the *Castanea* Genome Database and TreeGenes; geographic shapefiles from EUFORGEN; and plant images sourced from Natusfera and Pl@ntNet.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent of publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 21 May 2025 / Accepted: 23 February 2026

Published online: 04 March 2026

References

- Xofis P, Kefalas G, Poirazidis K. Biodiversity and conservation of forests. *Forests*. 2023;14(9):1871.
- DeSoto L, Cailleret M, Sterck F, Jansen S, Kramer K, Robert EMR, et al. Low growth resilience to drought is related to future mortality risk in trees. *Nat Commun*. 2020;11(1):545.
- McDowell NG, Allen CD, Anderson-Teixeira K, Aukema BH, Bond-Lamberty B, Chini L, et al. Pervasive shifts in forest dynamics in a changing world. *Science*. 2020;368(6494):eaaz9463.
- Tian F, Hu X-L, Yao T, Yang X, Chen J-G, Lu M-Z, et al. Recent advances in the roles of hsf s and hsp s in heat stress response in woody plants. *Front Plant Sci*. 2021;9(12):704905.
- Ul-Haq S, Khan A, Ali M, Khattak AM, Gai W-X, Zhang H-X, et al. Heat shock proteins: dynamic biomolecules to counter plant biotic and abiotic stresses. *Int J Mol Sci*. 2019;20(21):5321.
- Tiwari LD, Kumar R, Sharma V, Sahu AK, Sahu B, Naithani SC, et al. Stress and development phenotyping of Hsp101 and diverse other Hsp mutants of *Arabidopsis thaliana*. *J Plant Biochem Biotechnol*. 2021;30(4):889–905.
- Wei Y, Liu W, Hu W, Yan Y, Shi H. The chaperone MeHSP90 recruits MeWRKY20 and MeCatalase1 to regulate drought stress resistance in cassava. *New Phytol*. 2020;226(2):476–91.
- Madritsch S, Wischnitzki E, Kotrade P, Ashoub A, Burg A, Fluch S, et al. Elucidating drought stress tolerance in European oaks through cross-species transcriptomics. *G3 (Bethesda)*. 2019;9(10):3181–99.
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*. 2012;40(D1):D1178–86.
- Singh P, Unik B, Puri A, Nagpal G, Singh B, Gautam A, et al. HSPMdb: a computational repository of heat shock protein modulators. *Database (Oxford)*. 2020;2020:baaa003.
- Sayers EW, Beck J, Bolton EE, Brister JR, Chan J, Connor R, et al. Database resources of the National Center for Biotechnology Information in 2025. *Nucleic Acids Res*. 2025;53(D1):D20–9.
- Ma Y, Zhao X, Jia Y, Han Z, Yu C, Fan Z, et al. The updated genome warehouse: enhancing data value, security, and usability to address data expansion. *Genomics Proteomics Bioinformatics*. 2025;23(1):qzaf010.
- Fan J, Zhang Y, Nie X, Liu Y, Wei S, Peng H, et al. Comprehensive curation and validation of genomic datasets for chestnut. *Sci Data*. 2025;12(1):860.
- Wegrzyn JL, Staton MA, Street NR, Main D, Grau E, Herndon N, et al. Cyber-infrastructure to improve forest health and productivity: the role of tree databases in connecting genomes, phenomes, and the environment. *Front Plant Sci*. 2019;25(10):813.
- Perez-Rial A, Castro P, Solà A, Martín MA, Die JV. Decoding drought tolerance from a genomic approach in *Castanea sativa* Mill. *Plant Genome*. 2025;18(4):e70116.
- Rey M-D, Labella-Ortega M, Guerrero-Sánchez VM, Carleial R, Castillejo MÁ, Ruggieri V, et al. A first draft genome of holm oak (*Quercus ilex* subsp. *ballota*), the most representative species of the Mediterranean forest and the Spanish agrosilvopastoral ecosystem "dehesa." *Front Mol Biosci*. 2023;10:1242943.
- Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21(18):3674–6.
- Thumhuri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res*. 2022;50(W1):W228–34.
- European Forest Genetic Resources Programme. EUFORGEN. 2024. Available from: <https://www.euforgen.org/>. Accessed 7 May 2024.
- Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, et al. shiny: Web Application Framework for R. The R Foundation. 2024. R package version 1.11.0. Available from: <https://CRAN.R-project.org/package=shiny>.
- Sievert C, Cheng J, Aden-Buie G. bslib: Custom "Bootstrap""Sass"Themes for "shiny" and "rmarkdown." The R Foundation. 2024. R package version 0.9.0. Available from: <https://CRAN.R-project.org/package=bslib>.
- Cheng J, Schloerke B, Karambelkar B, Xie Y. leaflet: Create Interactive Web Maps with the JavaScript "Leaflet" Library. The R Foundation. 2024. R package version 2.2.2. Available from: <https://CRAN.R-project.org/package=leaflet>.
- Xie Y, Cheng J, Tan X. DT: A Wrapper of the JavaScript Library "DataTables." The R Foundation. 2024. R package version 0.33. Available from: <https://CRAN.R-project.org/package=DT>.
- Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. msa: an R package for multiple sequence alignment. *Bioinformatics*. 2015;31(24):3997–9.
- Rauscher B, Charlop-Powers Z. msaR: Multiple Sequence Alignment for R Shiny. The Comprehensive R Archive Network (CRAN); 2021. R package version 0.6. Available from: <https://CRAN.R-project.org/package=msaR>.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;15(10):421.

Publisher's Note

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