

Identification of the geographical origin of Spanish beech nuts, *Fagus sylvatica* L., using a chemometric supervised approach

Sandra Lamas,¹ Nuno Rodrigues,¹ Baudilio Herrero,² Susana Casal,³ Rebeca Cruz,³ José Alberto Pereira,¹ António M. Peres,¹

¹CIMO, LA SusTEC, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Bragança, 5300-253, Portugal; sandra.lamas@ipb.pt, nunorodrigues@ipb.pt, jpereira@ipb.pt, peres@ipb.pt

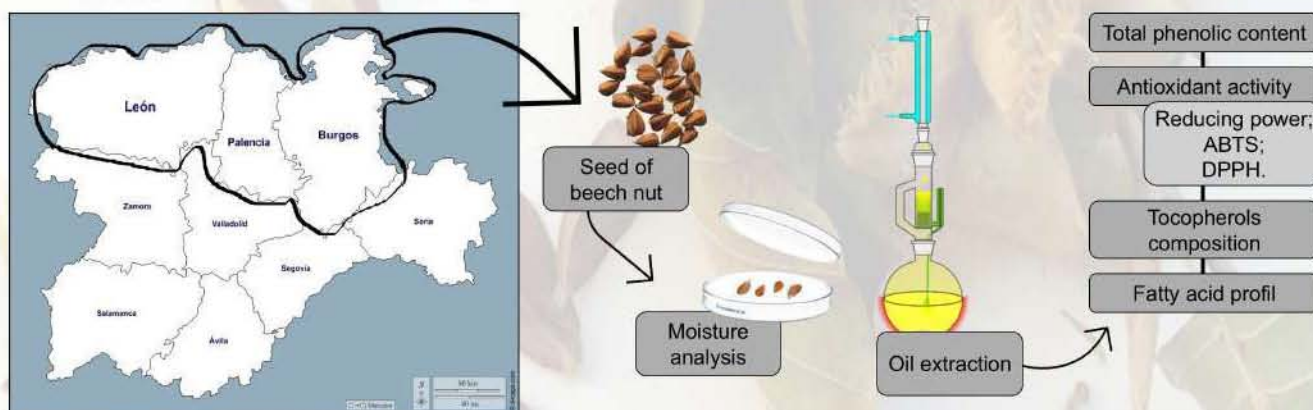
² Universidad de Valladolid, Dpto. de Ciencias Agroforestales, Avda. de Madrid, 57, Palencia, 34004, Spain; baudilio.herrero@uva.es

³ LAQV/REQUIMTE, Laboratory of Bromatology and Hydrology, Faculty of Pharmacy, University of Porto, Rua de Jorge Viterbo Ferreira, 228, Porto, 4050-313, Portugal; sucasal@ff.up.pt, rcruz@ff.up.pt

Introduction and objectives

In general, dry nuts, such as hazelnuts, almonds, chestnuts and walnuts, composition has been investigated and are well characterized. Nevertheless, other less known nuts, such as beech nuts have little information. Beech nuts proceed from the *Fagus sylvatica* L. tree, which is prevalent in central and southern Europe, and is commonly used for oil extraction.¹ A study carried out by Obranovic et al.¹ found that the chemical composition of beech nut oil varies significantly depending on the geographical origin. Siger et al.² characterized the oil obtained from beech nut seeds, highlighting its high content in linoleic acid. However, scientific data regarding the physicochemical properties and composition of beech nuts are scarce. In this context, this work evaluated whether there are differences in terms of the physicochemical data and composition of the beech nut between regions, by applying a chemometric procedure.

Material and methods



Results and Discussion

A linear discriminant analysis (LDA) was performed, establishing a multivariate classification model based on six variables selected as the most informative predictors, by the simulated annealing (SA) algorithm: moisture content and the relative abundance of C16:1, C18:3, C20:2, C22:0 and C22:1. The LDA-SA model allowed 100% of correct classifications for training (Figure 1).

Furthermore, when using leave-one-out cross-validation (LOO-CV), the model kept 100% of sensitivity, highlighting its robustness. A repeated K-fold CV was further performed to verify the accuracy of the proposed methodology, showing an average correct classification of 97.5% of the samples studied when 25% of the dataset was kept aside for internal validation.

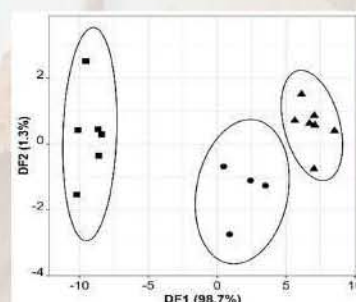


Figure 1: Supervised discrimination of the Spanish beech nuts according to geographical origin: (■ Burgos, ● León and ▲ Palencia).

Acknowledgements

The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support by national funds FCT/MCTES (PIDDAC) to CIMO (UIDB/00690/2020, <https://doi.org/10.54499/UIDB/00690/2020>; and UIDP/00690/2020, <https://doi.org/10.54499/UIDP/00690/2020>) CEB (UIDB/04469/2020) units as well as to the Associate Laboratory SusTEC (LAP/007/2020). National funding by FCT- Foundation for Science and Technology, through the institutional scientific employment program-contract with Nuno Rodrigues. Sandra Lamas acknowledges the Ph.D. research grant (2022.10070.BD) provided by FCT.

References

1. M. Obranovic, K. Kraljic, D. Skevin, S. Balbino, K. Tomljanovic, *Nutraceuticals* 4 (2024) 94-103.
2. A. Singer, K. Dwiecki, W. Borzyskowski, M. Turski, M. Rudzinska, M. Nogala-Kalucka, *Food Chem.* 225 (2017) 239-245.

Conclusion

The satisfactory classification results achieved allow the conclusion that the six variables selected may be seen as geographical biomarkers and can be used to develop a robust tool for identifying the geographical origin of beech nuts.