Genetic diversity assessment in Portuguese *Olea europea* L. cultivars using the combinatorial β-Tubulin-based polymorphism (cTBP) marker system

Hélia Cardoso*, Alexandre Ferreira*, Elisete Santos Macedo, Augusto Peixe & Birgit Arnholdt-Schmitt[1]

ICAAM, Instituto de Ciências Agrárias e Ambientais Mediterrânicas, IIFA, Instituto de Formação e Investigação Avançada, Universidade de Évora, Ap. 94, 7006-554 Évora, Portugal
[1] Corresponding author at: EU Marie Curie Chair, Laboratório de Biologia Molecular, ICAAM, Universidade de Évora. Email: eu_chair@uevora.pt
*These authors have contributed equally to this work.

Abstract

In *Olea europaea* L. subsp. *europaea* var. *sativa* one of the most outstanding issues for commercial production and product certification lies in clarifying genetic relationships among the highly diverse, cultivated olive material. This report validates the applicability of the combinatorial *β-tubulin* based polymorphism (cTBP) molecular marker system for *Olea europaea* L. genotypes. The method is based on the analysis of length polymorphisms resulting from combined amplification of introns 1 and 2 of the *β-tubulin* genes, followed by fragment resolution in polyacrylamide electrophoresis gel. When compared with other molecular tools the technique is advantageous in terms of low costs and time consumption. The cTBP method was applied on a collection of morphologically classified olive trees (Coleção Nacional de Referência de Cultivares de Oliveira) that are established in the 'Instituto Nacional de Investigação Agrária e Veterinária' (INIAV, Elvas, Portugal). The results show that the cTBP marker system is useful to trace genetic variability and phylogenetic relationships in olive material thus able to help clarifying misleading classifications based on morphological characterization. Therefore, the method here presented can be recommended as an efficient tool for commercial applications and scientific studies that aim to verify genetic relationships in *Olea europaea* L.

Keywords: EPIC-PCR, genotype fingerprinting, intron length polymorphism (ILP), olive, tubulin.

Resumo

Estudo da diversidade genética em cultivares portuguesas de *Olea europea* L. com recurso ao marcador molecular *β-Tubulin based polymorphism* (cTBP).

Um dos temas mais relevantes no quadro da produção comercial e certificação de plantas de *Olea europaea* L. subsp. *europaea* var. *sativa* está relacionado com a clarificação da relação genética existente entre as cultivares de oliveira cultivadas. No presente trabalho é feita a validação da aplicabilidade de um marcador molecular na genotipagem de variedades de *O. europaea*. O método baseia-se na análise da variabilidade genética (polimorfismos) existente ao nível dos genes da *β-tubulina,* combinando a variabilidade identificada ao nível dos intrões 1 e 2, sendo por isso denominado de cTBP (*combinatorial β-tubulin based polymorphism*). A abordagem seguida baseia-se na identificação em gel de poliacrilamida dos fragmentos de tamanhos diferentes, resultantes da amplificação das regiões dos intrões 1 e 2. Quando comparado

com outras ferramentas moleculares o marcador cTBP apresenta-se vantajoso, por implicar um baixo custo e não requerer muito tempo para a execução. O método foi testado em material vegetal proveniente da Coleção Nacional de Referência de Cultivares de Oliveira, existente no Instituto Nacional de Investigação Agrária e Veterinária (INIAV, Elvas, Portugal). Os resultados obtidos mostram que o marcador cTBP pode ser utilizado como uma ferramenta eficiente, que permite avaliar a variabilidade genética e as relações filogenéticas existentes entre plantas de oliveira, podendo assim ajudar a esclarecer dúvidas, ou retificar falsas classificações, previamente feitas com base na aplicação do método tradicional que utiliza a observação de carateres morfológicos. O marcador aqui apresentado pode assim ser recomendado como uma ferramenta eficiente para aplicações comerciais e estudos científicos que visam verificar relações genéticas em *O. europaea* L.

Palavras-chave: EPIC-PCR, genotipagem, *Intron length polymorphism* (ILP), oliveira, tubulina.

Introduction

Olive (*Olea europaea* L.) is one of the oldest agricultural tree crops worldwide. Its cultivation covers over eight million hectares, being predominantly concentrated in the Mediterranean basin, where 70% of the world olive oil produced is also consumed (Baldoni & Belaj, 2009). The cultivated olive *Olea europaea* subsp. *europaea* var. *sativa* comprises more than 2000 cultivars (Hamman-Khalifa et al., 2007), which are classified mostly considering morphologic and physiological characteristics like the variability in oil content, fruit size, canopy shape and adaptation to the local environmental conditions. However, the morphological characteristics and agronomic traits considered are highly influenced by the environmental conditions, leading to frequent misidentification and/or mislabeling. Thus, the large number of cultivars, added to the many cases of synonym and homonym, makes it particularly difficult to describe and classify the existing olive plant material.

Molecular markers appear as a useful tool to access genetic variability and are now used to supplement and refine the traditional morphological description of cultivars. Different marker systems were applied for phylogenetic studies, DNA fingerprinting of cultivars and detection of intra-cultivar variability, such as RAPD (Bronzini de Caraffa et al., 2002; Cordeiro et al., 2008), AFLP (Owen et al., 2005; Rubio de Casas et al., 2006), SCAR (Busconi et al., 2006), SSR (Baldoni et al., 2009; Fevereiro et al., 2011; Dastkar et al., 2013; Muzzalupo et al., 2014), ISSR (Gemas et al., 2004; Martins-Lopes et al., 2007) SNPs (Macedo et al., 2009; Hakim et al., 2010), ribosomal DNA polymorphism direct sequencing (Besnard et al., 2007; Baldoni et al., 2009), RFLP (Besnard et al., 2001; Besnard et al., 2007) and entire plastome sequencing (Mariotti et al., 2010).

From all markers applied on olive, the most successful were the SSR, due to the high number of loci identified and the availability of respective flanking primer sequences. However, not all combinations of SSR primers give polymorphic alleles, being also possible the occurrence of null alleles. The development of a method which involves low cost and being less time consuming will represent high advantages compared to the methods previously described.

The molecular marker here presented is named combinatorial Tubulin Based Polymorphism (cTBP). It is in fact an upgrade to the Tubulin Based Polymorphism (TBP) method described by Bardini et al. (2004), which is based on the analysis of the

intron 1 length polymorphism (ILP) of the plant *β*-tubulin gene family. The cTBP method allowed combining polymorphism analysis of the introns 1 and 2 of *β*-Tubulin by using gene-specific primers (degenerated across-plant species) designed in the exon boundaries for each of the two introns (Exon Priming Intron Crossing-PCR, EPIC-PCR), and subsequently analyzed in a polyacrylamide gel electrophoresis resulting in a DNA barcode image (Breviario et al., 2007; Braglia et al., 2010). cTBP allows to the identification of different alleles, always in the basis of different length, providing a neutral, co-dominant, stable and specific marker DNA fragments. The combination of the genetic variability in both introns results in a more reliable assessment of species/cultivars/ecotypes relationships (Breviario et al., 2007; Braglia et al., 2010). The discrimination power, of both TBP and cTBP methods, among species, varieties and ecotypes has been largely assessed and discussed by many authors (Bardini et al., 2004; Gavazzi et al., 2012; Braglia et al., 2016). The present study was developed in order to validate the applicability of the cTBP marker for olive material and promote the methodology as an efficient tool for scientific and commercial applications in public and private sectors.

Material and Methods

Young leaves of nineteen different cultivars of *Olea europaea* L. were used for investigation (Table 1). Field growing olive trees are from a mother-tree collection, belonging to the 'Instituto Nacional de Investigação Agrária' (INIAV), Elvas (Portugal). Genomic DNA (gDNA) extraction was performed with the DNEasy Plant Mini kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. gDNA integrity was analysed by electrophoresis in 0.8% agarose gel (Invitrogen Life technologies, UK) after staining in an ethidium bromide (EtBr) solution (0.2 $ngmL^{-1}$). Gel documentation was performed with the Gene Flash Bio Imaging System (Syngene, Cambridge, UK). gDNA content was measured using a NanoDrop-2000C spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

EPIC-PCRs were performed using degenerated primers for β-Tubulin intron 1 and 2 amplification (Breviario et al., 2007). PCR was carried out in the 2720 Thermocycler (Applied Biosystems, Foster City, USA) following parameters described by Casazza et al. (2011). Ready-To-Go PCR Beads (GE Healthcare, Little Chalfont, England) were used for PCR mix, using 10 $ng\mu L^{-1}$ of gDNA as template and 0.2 µM of each primer. Three independent PCRs were performed per sample. Eight µl of each PCR product were analyzed by electrophoresis in 6% polyacrylamide gel running for 4 h at 100 V. Gel staining and visualization was performed as described above. Two electrophoreses were performed per PCR product.

The genetic distance data between pairs of genotypes was used to construct a matrix of presence/absence of amplified ILP fragments. The FreeTree software was used to compute the distance/similarity matrix according to Nei and Li (1979), to calculate the Unweighted Pair Group Method with Arithmetic Average (UPGMA) dendrogram, and to make the resamplings by bootstrapping with 1000 replications. Dendrogram construction, based on the reference tree obtained with FreeTree software, was performed on MEGA 4 software (Tamura et al., 2007).

Results and Discussion

Figure 1 shows the analysis of cTBP amplification profiles (also named fingerprints, barcodes or polymorphic bands) of 19 *O. europaea* cultivars. Both introns displayed length polymorphisms. The polymorphic bands identified in intron 1 ranged

between 550 - 2000 bp (the majority between 850 - 2000 bp) whereas in intron 2 they ranged between 350 - 1100 bp (the majority between 600 - 1000 bp).

The band pattern in intron 1 was more polymorphic, with 32 ILP bands, 27 of them polymorphic between the cultivars analysed. Intron 2 presented an amplification profile with 11 ILP bands, 6 of them polymorphic among cultivars. In total 43 ILP were identified, with 33 ILP polymorphic bands. This level of variability is in agreement with reports based on other molecular markers, such as the RAPD markers (Gemas et al., 2004; Trujillo et al., 2003; Cordeiro et al., 2008).

Polymorphic bands identified by cTBP method were able to confirm the existence of differences between cultivars usually considered as identical (homonyms), and also to clarify the existence of synonyms, with amplification of the same cTBP profiles in cultivars known by different names.

An example of synonyms identification is the identical cTBP barcode image shared for both introns among the three cultivars 'Manzanilla Carrasqueña de Almendralejo', 'Manzanilla de Jaén' and 'Redondil' (*2 in fig. 1). The total similarity between these cultivars can also be seen in the dendrogram of fig. 2. These data are validated by previous reports based on other molecular markers. Trujillo et al. (2003), describe the same RAPD profile for cvs. 'Manzanilla de Sevilla' (synonym of 'Manzanilha de Jaen' and 'Manzanilha Carrasquenha de Almendralejo') and 'Redondil', and Lopes et al. (2004) identified similar microsatellite loci in both 'Manzanilla Carrasqueña de Almendralejo' and 'Manzanilla de Jaén'. Nevertheless, all those cultivars are considered as different ones on the FAO database (http://www.oleadb.it/olivodb.html).

The cultivars 'Negrinha' and 'Azeiteira', previously classified as synonyms using morphological markers (Cordeiro et al., 2008), revealed different cTBP barcodes. A single polymorphic band in intron 1 distinguished both cultivars (*5 in intron fingerprints of fig. 1), leading to cultivars differentiation with a high level of similarity (98%) (fig. 2). Previous reports based on molecular makers, such as microsatellite (SSR) (Lopes et al., 2004), RAPD and ISSR (Gemas et al., 2004; Martins-Lopes et al., 2007; Cordeiro et al., 2008), also distinguished these two cultivars.

Another example of misclassification is the case of cvs. 'Madural' and 'Cornezuelo', considered as synonyms at the FAO germplasm database. The application of the cTBP differentiated both, sharing only 72% similarity, which is in agreement with a previous report based on SSR (Lopes et al., 2004).

In fig. 2 a complete identity of the cTBP band pattern between two trees coming from clonal propagation of cv. 'Cobrançosa' (named 'Cobrançosa' 1 and 2) is also shown, which confirms the accuracy of the cTBP method for cultivars discrimination. Regarding the conservation of the cTBP pattern within a cultivar, previous experiments performed in the apomitic *Hypericum perforatum* L. showed complete identity among progeny plants and mother plant.

Conclusions

In this study, cTBP marker system can successfully distinguish olive cultivars and can efficiently help to clarify synonym and homonym relations. Phylogenetic analysis based on cTBP data is consistent with the knowledge available for these cultivars. Application of the methodology on a higher number of olive cultivars and related olive taxa will help to a better understanding of olive phylogeny. The results achieved can validate several available reports, based on different molecular marker systems, all applying more complex methodologies. cTBP method is based on basic

PCR technique and polyacrylamide gel electrophoresis, and besides the reliability demonstrated, it's a simple and low costs method.

Acknowledgements

The study was funded by the National Funds through FCT - Foundation for Science and Technology under the project PTDC/AGRGPL/099263/2008 that supported also Alexandre Ferreira, and the project UID/AGR/00115/2013. The authors are grateful to FCT also for the support given to HGC (SFRH/BPD/109849/2015) and want to thank Dr. António Cordeiro from INIA-Elvas (Portugal), who provided the olive plant material. We recognize stimulation of this research and comments given on our work to Diego Breviario.

References

Baldoni, L. & Belaj, A. 2009. Olive. P. 397-421. In: J. Vollmann & I. Rajean (eds.), Oil crops. Handbook of plant breeding, Vol. 4, Springer Science Business Media, New York.

Baldoni, L., Cultrera, N.G., Mariotti, R., Riccioloni, C., Arcioni, S., Vendramin, G.G., Buonamici, A., Porceddu, A., Sarri, V., Ojeda, M.A., Trujillo, I., Rallo, L., Belaj, A., Perri, E., Salimonti, A., Muzzalupo, I., Casagrande, A., Lain, O., Messina, R. & Testolin, R. 2009. A consensus list of microsatellites markers for olive genotyping. Molecular Breeding 24:213-231.

Bardini, M., Lee, D., Donini, P., Mariani, A., Gianì, S., Toschi, M., Lowe, C. & Breviario, D. 2004. Tubulin-based polymorphism (TBP): a new tool, based on functionally relevant sequences, to assess genetic diversity in plant species. Genome 47:281-291.

Besnard, G., Baradat, P.H., Chevalier, D., Tagmount, A. & Bervillé, A. 2001. Genetic differentiation in the olive complex (*Olea europaea*) revealed by RAPDs and RFLPs in the rRNA genes. Genetic Resources and Crop Evolution 48:165-182.

Besnard, G., Casas, R.R. & Vargas, P. 2007. Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea*). Journal of Biogeography 34:736-752.

Braglia, L., Gianì, S., Breviario, D., Gavazzi, F., Mastromauro, F. & Morello, L. 2016. Development and validation of the modular Feed-code method for qualitative and quantitative determination of feed botanical composition. Analytical and Bioanalytical Chemistry 408:8299-8316.

Braglia, L., Manca, A., Mastromauro, F. & Breviario, D. 2010. cTBP: A Successful Intron Length Polymorphism (ILP)-Based Genotyping Method Targeted to Well Defined Experimental Needs. Diversity 2:572-585.

Breviario, D., Baird, W.V., Sangoi, S., Hilu, K., Blumetti, P. & Giani, S. 2007. High polymorphism and resolution in targeted fingerprinting with combined β-tubulin introns. Molecular Breeding 20:249-259.

Bronzini de Caraffa, V., Maury, J., Gambotti, C., Breton, C., Berville, A. & Giannettini, J. 2002. Mitochondrial DNA variation and RAPD mark oleasters, olive and feral olive from Western and Eastern Mediterranean. Theoretical and Applied Genetics 104:1209-1216.

Busconi, M., Sebastiani, L. & Fogher, C. 2006. Development of SCAR markers for germplasm characterisation in olive tree (*Olea europaea* L.). Molecular Breeding 17:59-68.

Casazza, A.P., Gavazzi, F., Mastromauro, F., Gianì, S. & Breviario, D. 2011. Certifying the feed to guarantee the quality of traditional food: An easy way to trace plant species in complex mixtures. Food Chemistry 124:685-691.

Cordeiro, A.I., Sanchez-Sevilla, J.F., Alvarez-Tinaut, M.C. & Gomez-Jimenez, M.C. 2008. Genetic diversity assessment in Portugal accessions of *Olea europaea* by RAPD markers. Biologia Plantarum. 52 (4):642-647.

Dastkar, E., Soleimani, A., Jafary, H., & Naghavi, M.R. 2013. Genetic and morphological variation in Iranian olive (*Olea europaea* L.) germplasm. Crop Breeding Journal 3(2):99-106.

Fevereiro, P., Leitão, F., Potes, F., Gemas, V., Alves, M. & Favoretto, P. 2011. The Portuguese olive (*Olea europaea* subsp. *europaea*) germplasm. Acta Horticulturae 924:291-298.

Gavazzi, F., Casazza, A.P., Depedro, C., Mastromauro, F. & Breviario, D. 2012. Technical improvement of the TBP (tubulin-based polymorphism) method for plant species detection, based on capillary electrophoresis. Electrophoresis 33:2840-2851.

Gemas, V.J.V., Almadanim, M.C., Tenreiro, R., Martins, A. & Fevereiro, P. 2004. Genetic diversity in the olive tree (*Olea europaea* L. subsp. *europaea*) cultivated in Portugal revealed by RAPD and ISSR markers. Genetic Resources and Crop Evolution 51:501-511.

Hakim, R.I., Grati-Kammoun, N., Makhloufi, E. & Rebaï, A. 2010. Discovery and potential of SNP markers in characterization of Tunisian olive germplasm. Diversity 2:17-27.

Hamman-Khalifa, A.M., Navajas-Pérez, R., Herrán, R., Rejón, M.R., Garrido-Ramos, M.A., Ruiz Rejón, C. & Rosúa J.L. 2007. Establishing the genetic relationships between the wild and cultivated olives using a nuclear intron from nitrate reductase (nia-i3). Plant Systematics and Evolution 269:63-73.

Lopes, M.S., Mendonça, D., Sefc, K.M., Gil, F.S. & Machado, A.C. 2004. Genetic evidence of intra-cultivar variability within Iberian olive cultivars. Hortscience 39 (7):1562-1565.

Macedo, E.S., Cardoso, H.G., Hernández, A., Peixe, A.A., Polidoros, A., Ferreira, A., Cordeiro, A. & Arnholdt-Schmitt, B. 2009. Physiologic responses and gene diversity indicate olive alternative oxidase as a potential source for markers involved in efficient adventitious root induction. Physiologia Plantarum137:532-552.

Mariotti, R., Cultrera, N.G.M., Díez, C.M., Baldoni, L. & Rubini, A. 2010. Identification of new polymorphic regions and differentiation of cultivated olives (*Olea europaea* L.) through plastome sequence comparison. BMC Plant Biology 10:211-224.

Martins-Lopes, P., Lima-Brito, J., Gomes, S., Meirinhos, J., Santos, L. & Guedes-Pinto, H. 2007. RAPD and ISSR molecular markers in *Olea europaea* L.: Genetic variability and molecular cultivar identification. Genetic Resources and Crop Evolution 54:117-128.

Muzzalupo, I., Vendramin, G.G. & Chiappetta, A. 2014. Genetic Biodiversity of Italian Olives (Olea europaea) Germplasm Analyzed by SSR Markers. The Scientific World Journal, Article ID 296590. doi.org/10.1155/2014/296590.

Nei, M. & Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences 76:5269-5273.

Owen, C.A., Bita, E.C., Banilas, G., Hajjar, S.E., Sellinakis, V., Aksoy, U., Hepaksoy, S., Chamoun, R., Talhook, S.N., Metzidakis, I., Hatzopoulos, P. & Kalaitzis, P. 2005.

AFLP reveals structural details of genetic deiversity within cultivated olive germoplasm from the Eastern Mediterranean. Theoretical and Applied Genetics 110:1169-1176.

Rubio de Casas, R., Besnard, G., Schönswetter, P., Balaguer, L. & Vargas, P. 2006. Extensive gene flow blurs phylogeographic but not phylogenetic signal in *Olea europaea* L. Theoretical and Applied Genetics. 113:575-583.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biologyand Evolution 24 (8):1596-9.

Trujillo, I., Barranco, D., Belaj, A. & Rallo, L. 2003. Identificación y catalogación del banco de germoplasma mundial de olivo de Córdoba mediante marcadores morfológicos y moleculares (Proyecto CAO98-001). http://www.academia.edu/23368017/Identificaci%C3%B3n_y_Catalogaci%C3%B3n_Del_Banco_De_Germoplasma_Mundial_De_Olivo_De_C%C3%B3rdoba_Mediante_Marcadores_Morfol%C3%B3gicos_y_Moleculares_Proyecto_CAO98-001.

Table 1 - Cultivars of *Olea europaea* L. subsp. *europaea* var. *sativa* used and its origin (country-region). The names are in accordance with the FAO nomenclature.

Cultivar Name	Country	Region
'Arbequina'	Spain	Catalunha
'Azeiteira'	Portugal	Beira Interior
'Carrasquenha'	Portugal	Alentejo
'Cobrançosa'	Portugal	Trás-os-Montes
'Conserva d'Elvas'	Portugal	Alto-Alentejo
'Cordovil Castelo Branco'	Portugal	Beira Interior
'Cordovil de Elvas'	Portugal	Alto Alentejo
'Cordovil de Serpa'	Portugal	Baixo Alentejo
'Cornezuelo'	Portugal	Castilla La Mancha
'Galega'	Portugal	unknown
'Madural'	Portugal	Trás-os-Montes
'Mançanilha de Jaén'	Spain	Andaluzia
'Mançanilha de Tavira'	Portugal	Algarve
'Manzanilla Carrasqueña de Almendralejo'	Spain	Extremadura
'Negrinha'	Portugal	Trás-os-Montes
'Redondil'	Portugal	Alto Alentejo
'Verde Verdelho'	Portugal	unknown
'Verdeal de Serpa'	Portugal	Baixo Alentejo
'Verdeal Transmontana'	Portugal	Trás-os-Montes

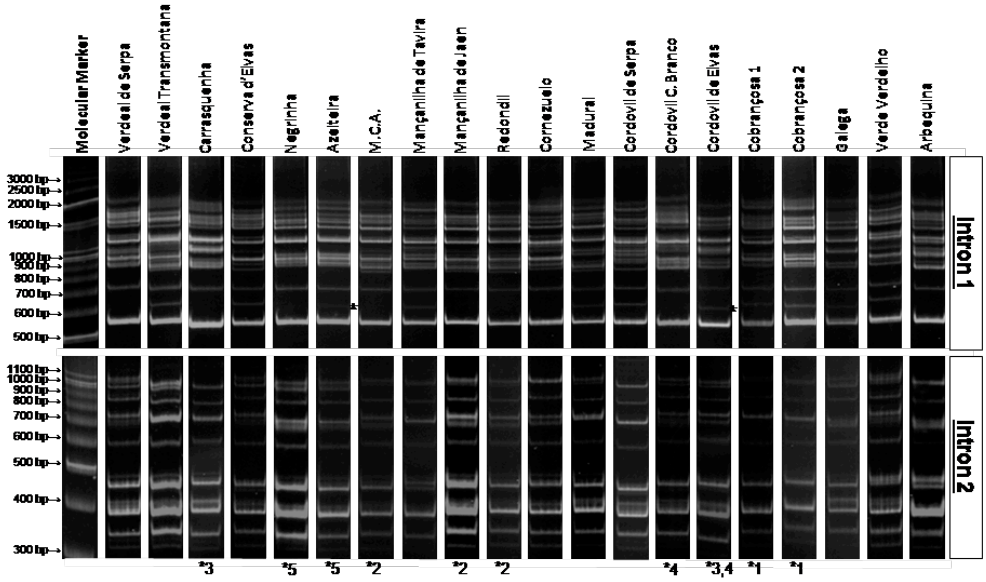

Figure 1 - Genomic fingerprints based on the cTBP amplification profile for intron 1 (top) and intron 2 (bottom) of 20 different *Olea europaea* subp. *europaea* var. *sativa*, belonging to 19 cultivars (cultivars names on top of each cTBP fingerprint). At the bottom of figure, asterisks with identical numbers represents identical (case of *1 and *2) or close related fingerprints (case of *3, *4 and *5). Molecular marker: MassRuler^{TM} DNA Ladder (Fermentas, Ontario Canada); M.C.A.: Manzanilla Carrasqueña de Almendralejo.

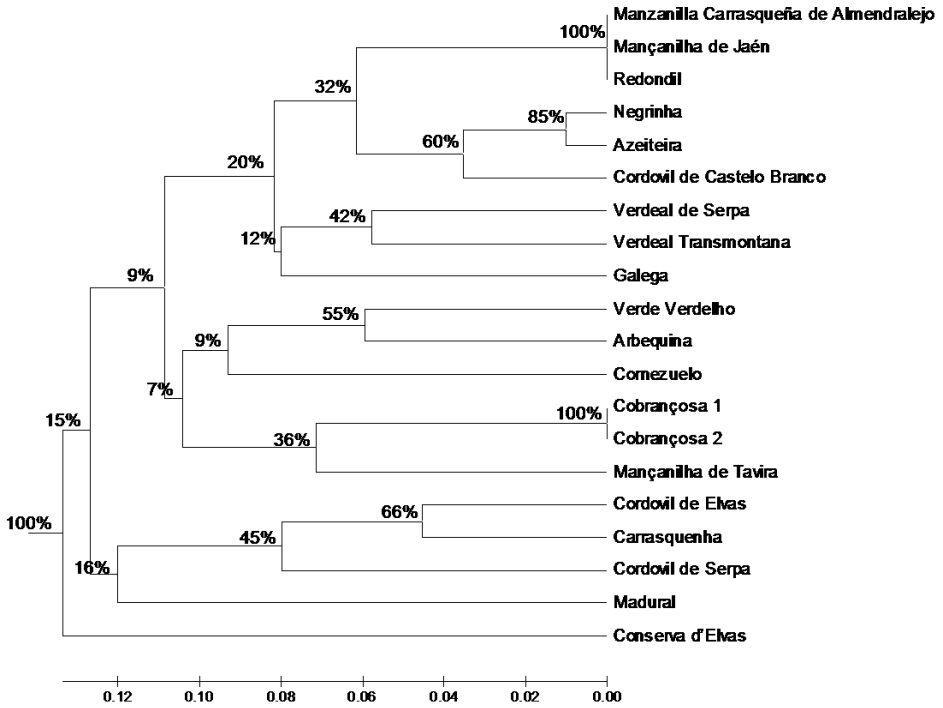

Figure 2 - UPGMA dendrogram constructed based on the cTBP genomic profile of 20 *Olea europaea* subp. *europaea* var. *sativa* genotypes, belonging to 19 different cultivars. Genetic distances are represented by branch length (distance coefficient). Values at nodes refer to distance coefficient. Percentages on the branches represent bootstrap values of 1000 replicates.