Modelling molecular and inorganic data of *Amanita ponderosa* mushrooms using artificial neural networks

Cátia Salvador[1,2], M. Rosário Martins[1,3], Henrique Vicente[1,2], José Neves[4], José M. Arteiro[1,2] and A. Teresa Caldeira[1,2]

[1]*Chemistry Department, University of Évora, 7000 - 671 Évora, Portugal*
[2]*CQE, University of Évora, 7000 - 671 Évora, Portugal*
[3]*ICAAM, University of Évora, 7000 - 671 Évora Portugal*
[4]*Informatics Department, University of Minho, Braga, Portugal*

Keywords: *Amanita ponderosa*, mycorrhizal edible mushrooms, M13-PCR, artificial neural networks.

Amanita ponderosa are wild mushroom eatable, growing spontaneously in some Mediterranean microclimates, namely in Alentejo and Andaluzia, in the Iberian Peninsula, due to its Mediterranean characteristics. It is characterized by a large and robust basidiomata, with a cap 8-17 cm in diameter, a hemispheric morphology when young and plane-convex in maturity, and a slight depression in the centre (Fig. 1).

Figure 1: Some macroscopic Aspects of Eatable *Amanita ponderosa* mushrooms in maturity.

The nutritional values of these fungi make them highly exportable. Due to the wide diversity of mushrooms in nature, it is essential to differentiate and to identify the various edible species. Mushrooms can accumulate high concentrations of some elements, namely toxic metals, because the symbiotic relation between these macrofungi and some plants in its habitats.

The aim of this study was to evaluate inorganic composition of mycorrhizal *Amanita ponderosa* collected from different regions of the southwest of the Iberian Peninsula and to access molecular biomarkers using artificial neural networks.

Artificial neural networks (ANNs) are widely accepted as a tool that offers an alternative way to tackle complex problems. They can learn from examples, are fault tolerant in the sense that they are able to handle noisy and incomplete data, are able to deal with non-linear problems, and once trained can perform prediction and generalization at high speed. Predictive modelling is a process used in predictive analytics to create a statistical model of future behaviour. Predictive analytics is the area of Data Mining concerned with forecasting probabilities and trends. On the other hand, Artificial Intelligence (AI) concerns itself with intelligent behaviour, i.e. the things that make us look intelligent. Following this process of thinking, the aim of this is to assess the impact of using AI based tools for the development of intelligent predictive models, in particular those that may be used to predict the M13-PCR DNA band profile from *Amanita ponderosa*. Indeed, the prediction of phylogenetic profiles is a complex and highly nonlinear problem for which there are no known methods for a direct and accurate prediction.

Fruiting bodies of the *A. ponderosa* mushrooms were collected in Spring from different locations area, in the southwest of the Iberian Peninsula. Three individuals were sampled per location. The material was weighed and placed in sterile bags for its inorganic study and molecular characterization. The inorganic contents of *A. ponderosa* samples were analyzed according with Moreno-Rojas *et al*. (2004).

For the Molecular analyses, DNA amplification was carried out with M13 primer, as described by Caldeira *et al*. (2009). Phylogenetic tree was generated by UPGMA method, using the Dice coefficient of similarity. In this molecular study, the *A. ponderosa* strains were compared with others Basidiomycetes and with one Ascomycete strain.

The inorganic analyses showed that mineral composition of these mushrooms depends on the ecosystem where they grow. Ca, Na and Mg were the elements that mainly contribute to the heterogenic inorganic profile between studied strains. Levels of trace metals are considerably lower, acceptable to human consumption at nutritional and low toxic levels.
Molecular approach using the microsatellite primer M13-PCR allowed to distinguish the mushrooms at specie level and to differentiate the *A. ponderosa* strains according to their location. Data mining tools were used in order to correlate inorganic and molecular results. In order to obtain the best prediction of the M13 PCR DNA band profile, different network structures and architectures were elaborated and evaluated. The optimum number of hidden layers and the optimum number of nodes in each of these was found by trial and error. The methodology adopted starts with a small network (two hidden layers, each with two nodes) and continued by adding nodes to improve performance i.e. in order to minimize an error metric. In the present work the error metric used was the mean squared error.
The neural network selected for modelling the data has a 6-7-14 topology, i.e. an input layer with six nodes, a hidden layer with seven nodes and a fourteen nodes output layer. A good match between the experimental and predicted values can be observed.
The results show that it may be possible to correlate the molecular and inorganic data. The present findings are wide potential application and both health and economical benefits arise from this study.

Caldeira, A.T., Salvador, C., Pinto, F., Arteiro, J.M., Martins, M.R. (2009). "MSP-PCR and RAPD Molecular Biomarkers to characterize *Amanita ponderosa* mushrooms". *Annals of Microbiology*, 59 (3), 1–6.

Caldeira, A.T., Vicente, H., Arteiro, J.M., Roseiro, J.C. e Neves, J. (2011) "An Artificial Intelligence Approach to *Bacillus amyloliquefaciens* CCMI 1051 Cultures: Application to the Production of Antifungal Compounds" *Bioresour. Technol*., 102, 1496–1502.

Moreno-Rojas, R, Díaz-Valverde, A, Moreno-Arroyo, B, González, T, Capote, C (2004). "Mineral content of gurumelo (*Amanita ponderosa*)". *Food Chemistry,* 85, 325–330.