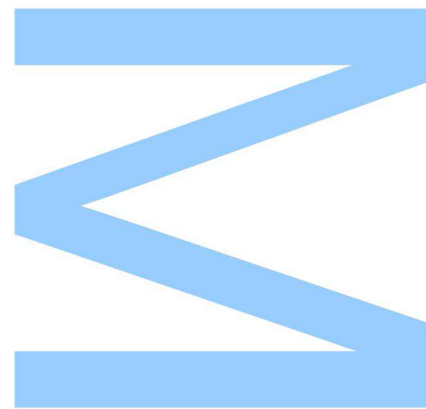
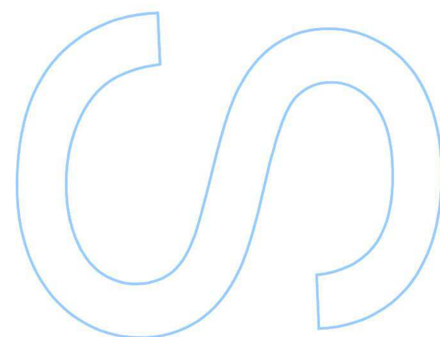
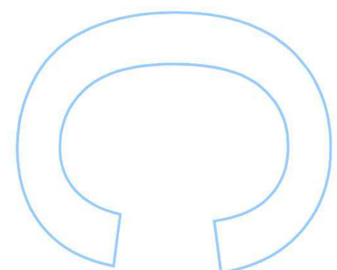
Assessing phylogeographic traits and distribution patterns of *Amanita ponderosa* (Malençon & R. Heim) in Iberian Peninsula







Ricardo Manuel Arraiano Castilho
Biodiversity, Genetics and Evolution
CIBIO - FCUP

Supervisor
Albano Gonçalo Beja Pereira
Investigador Auxiliar – CIBIO-FCUP

Assessing phylogeographic traits and distribution patterns of *Amanita ponderosa* (Malençon & R. Heim) in Iberian Peninsula

Ricardo Manuel Arraiano Castilho

Research Center in Biodiversity and Genetic Resources

Master thesis

Porto 2013

Supervisor

Albano Gonçalo Beja Pereira

Investigador Auxiliar - CIBIO-FCUP

U. PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

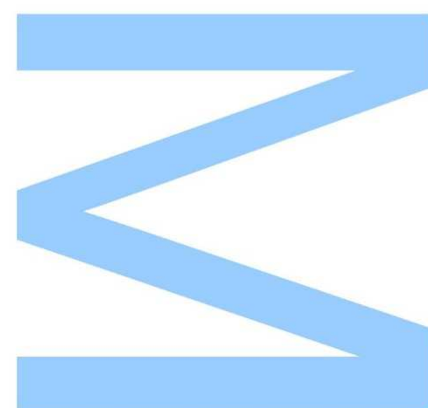
CIBIO
Centro de Investigação em Biodiversidade e Recursos Genéticos

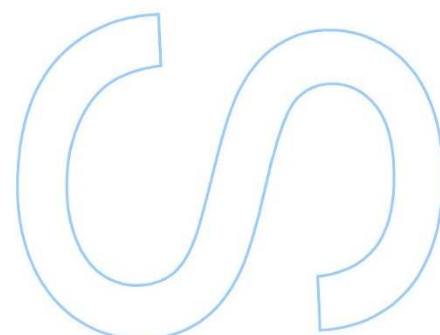
U.PORTO

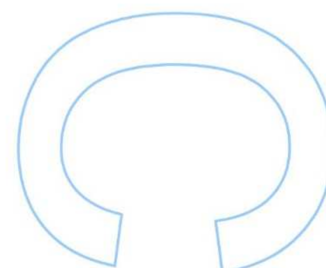
FC FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, ______/______/_________







Esta página foi intencionalmente deixada em branco

Acknowledgements

Behind our personal achievements, besides own considerable effort, hides usually a very large number of contributions, support, suggestions, comments or criticisms from many people. I would like to thank here all those who in some way helped me and contributed to all this became possible. Beginning of course for my supervisor, Dr. Albano Beja Pereira from CIBIO/InBIO who was the first to accept the crazy idea of studying mushrooms giving me the opportunity to proceed the studies in my interest research area and for the enthusiastic, encouragement and useful critiques for this research work. I also would like to thank in particular to Vânia Costa for all advices and lab supervision, Mr. Shanyuan Chen for helping in data Analysis and all members of GENPOP research group for their contribution.

To Professor Michael R. Miller from University of California, Davis for the precious help and for contributing on my ongoing learning process in RAD data analysis.

I would like to express my very great appreciation to Professor Paulo de Oliveira, who always supported and provided me valuable reflections on this unfolding adventure as well as my friends Rui Cardoso, Luis Morgado and Carlos Vila-Viçosa who also gave me a good support, contributions and helped in field work.

To all those who made possible this ambitious journey through the Iberian Peninsula and supported me in some way in fieldwork starting by Duarte Candeias – CEVERN, Mrs Célia and family from São Barnabé, Eng. Gravito Henriques from Castelo Branco, my friend Ana Maria and Mr. Galvão from Erra-Coruche, Mr. Maramelo from Couço, friend Sara Filipe and her father from Tomar, To "Câmara Municipal de Moura" and "Herdade da Contenda", Mr. Fernando Durán, Mr. Filipe Pla and Mr. Celestino Gelpi members of "Sociedad Micológica Extremeña", Mr. Aurelio García Blanco from "Asociación Vallisoletana de Micología", Prof. Luis Ruiz Valenzuela, from Dpto. Biología Animal, Biología Vegetal y Ecología - Universidad de Jaén and Mr. Nicola Amalfi from Sicily. Also thanks to Mrs. Margarita Duñas Curator of Cryptogams Collection from Real Jardín Botánico de Madrid for the cession of some samples and to Professor Rubim Almeida – FCUP by the institutional collaboration.

I would also like to thank to Ana Filipa Ferreira for their patience, work review and moral support in difficult times.

To all "Masterss" people and lab colleagues.

And of course the last but not least to my parents and family for the support and investment in my education and person I am.

Thank you all

Abstract

One of the most emblematic genus of the Fungi Kingdom is the Amanita. This containing not only some of the most toxic species as others gastronomically very appreciated. *Amanita ponderosa* is one of the most charismatic species of the Iberian Peninsula. With high commercial value it suffers a strong pressure of harvest during the fruiting period. Little is yet known about this and other species of mushrooms and which impact can cause the exploitation of this resource. The aims of this study are therefore to carry out an assessment of the fitness of populations of *A. ponderosa* throughout their Iberian distribution through identification of a battery of molecular markers (SNPs) that allow us to access phylogeographic patterns of this species in the Iberian Peninsula; Infer the species complex that compose the group of *A. ponderosa* in order to better understand which patterns allow us to distinguish them among themselves. And finally identify potential areas and environmental factors that allow us to determine the occurrence of the species and locate priority areas for the conservation. Restriction-site associated sequencing (RAD-seq) was used for markers identification. The distribution model of the species followed principle of maximum entropy (Maxent) through 201 occurrences records and 9 predictive variables. 201 SNP were identified in *Amanita ponderosa* and 247 in *Amanita cf pseudovalens.* Through analysis of the ITS rDNA and SNP data was identified a new species in *Lepiotoides* species complex, here dubbed *Amanita cf pseudovalens.* There was a weak structure of populations on the Peninsula showing a strong gene flow between populations. The variables most prevalent in species distribution are related to rainfall, temperature and land cover. The priority areas for the conservation of the species are located in the north and northeast of the peninsula. This study allowed us to increase the knowledge of this species therefore is an important tool to incorporate into future management plans of this resource.

Keywords

Amanita ponderosa, Phylogeography, SNP, RAD-seq, Maxent

Resumo

Um dos géneros mais mediáticos do Reino Fungi é o género Amanita, contendo não só algumas das espécies mais tóxicas como outras extremamente apreciadas gastronomicamente. A *Amanita ponderosa* é uma das espécies mais características da Península Ibérica. Esta tem um elevado valor comercial sofrendo por isso uma forte pressão de colheita. Pouco se sabe ainda sobre esta e outras espécies de cogumelos e qual o impacte causado pela exploração deste recurso. Pretende-se por isso com este estudo realizar uma avaliação do fitness das populações de *A. ponderosa* em toda a sua distribuição Ibérica, através da identificação de uma bateria de marcadores moleculares (SNP) que nos permitam analisar padrões filogeográficos desta espécie na Península Ibérica; inferir sobre a complexidade de espécies que compões o grupo da *A. ponderosa* de forma a melhor compreender quais os padrões que nos permitem distingui-las entre sí; e por fim identificar potenciais factores ambientais que nos permitam determinar a ocorrência da espécie e localizar áreas prioritárias para a conservação da mesma. A técnica de sequenciação utilizada foi a sequenciação massiva de segmentos associados a sítios de restrição, uma técnica de última geração que permite a identificação dos marcadores moleculares. O modelo de distribuição da espécie seguiu o princípio da máxima entropia (Maxent) através de 201 registos de ocorrências e 9 variáveis preditivas. Foram identificados 201 SNP in *Amanita ponderosa e* 247 in *Amanita cf pseudovalens.*. Através da análise do ITS rDNA foi identificado uma nova espécie próxima da *Amanita curtipes*, aqui apelidada de *Amanita cf pseudovalens.* Observou-se uma fraca estruturação das populações na Península evidenciando um forte fluxo genético entre populações. As variáveis mais preponderantes com a distribuição da espécie estão relacionadas com precipitação, temperatura e ocupação do solo. As áreas mais prioritárias para a conservação da espécie são situadas a norte e nordeste da península. Este estudo permitiu aumentar o conhecimento sobre esta espécie sendo por isso uma ferramenta importante para incorporar em futuros planos de gestão do recurso.

Palavras-chave

Amanita ponderosa, filogeografia, SNP, RAD-seq, Maxent

Table of contents

Abstract 6
Resumo 7
List of illustrations and tables 9
List of Manuscripts 10
1. General Introduction 11
1.2 – Lifestyle 12
1.3 – Habitat and Distribution 15
1.4 – Species complex (taxonomy) 17
1.5 – Socio-Economic importance 18
1.6 – State of the art: The role of new approaches 19
1.7 – Objectives 23
2. Manuscript I 24
2.1 Abstract 25
2.2 Introduction 25
2.3 Materials and methods 27
2.4 Results 32
2.6 References 39
3. Manuscript II 42
3.1 Abstract 43
3.2 Introduction 43
3.3 Materials and methods 45
3.4 Results 47
3.5 Discussion 52
3.6 References 54
3.6 Supplementary materials 57
4. Final considerations 60
5. References 61

List of illustrations and tables

Fig 1 *Amanita ponderosa* Malençon & R. Heim 12
Fig 2 Types of mycorrhiza: 13
Fig 3 Mushroom life cycle. 13
Fig 4 *Amanita ponderosa* semi- hypogeous development:... 14
Fig 5 Amanita ponderosa typical habitat 15
Fig 6 A. ponderosa spatial distribution. 16
Fig 7 Lepiotoides species complex 18
Fig 8 Amanita ponderosa selling in Huelva Market 19
Fig 9 Nuclear ribosomal repeat unit 20
Fig 10: Restriction-site associated DNA sequencing. 22
Fig 11 Sampling sites in Iberian Peninsula 29
Fig 12 Phylogenetic tree based on ITS1, 5.8s and ITS2 of rDNA,.............................. 33
Fig 13 PcoA outputs from GeneAlex. 34
Fig 14 STRUCTURE plot for *Amanita pseudovalens*. 35
Fig 15 STRUCTURE plot for *Amanita ponderosa*. 36
Fig 16 Amanita curtipes a) f. curtipes and b) f. pseudovalens with notorious morphological differentiation. 37
Fig 17 Jackknife tests for *Amanita ponderosa* 49
Fig 18 Response curves of *Amanita ponderosa* 50
Fig 19 Model prediction of probability of occurrence of Amanita ponderosa in Iberian Peninsula. 51

Table 1 List of samples used in this study. 28
Table 2 Sampling locations used to predict species distribution model…........….........47
Table 3 Predictive variables selected (P<0.8)and used to model species distribution. 48
Table 4 Contribution of each predictive variables 48

List of Manuscripts

This thesis is based on the following manuscripts:

Manuscript I - Next-generation RAD sequencing: a tool for evaluation of population structure of *Amanita ponderosa* in Iberian Peninsula

Manuscript II - Assessing distribution patterns of *Amanita ponderosa* in Iberian Peninsula

1. General Introduction

1.1 – Fungal Biodiversity: *The Amanita*

The world fungal diversity was estimated conservatively to be at least 1.5 million species (Hawksworth 1991) of which only ~7% have been described (Kirk *et al.* 2008). Of the 1.5 million estimated fungi, 140,000 species produce fruiting bodies of sufficient size and suitable structure to be considered macro fungi, which can be called mushrooms (Hawksworth 2001). The kingdom of fungi encompasses a tremendous biological diversity, with members including a wide array of lifestyles, forms, habitats, and sizes (Branco 2011). They play some essentials and indispensable ecological roles in the balance of the concerned environment, most notably due decomposition processes, but also involved in important symbiotic associations which may be so extreme that they could become parasites (Alexopoulus *et al.* 1996).

Those who we can actually call mushrooms, are organized mainly in two main divisions, Ascomycota and Basidomycota. The latter being representative of the most typical shaped mushrooms with cap (pillius) and stem (stipe) (Kirk *et al.* 2008). One of the most familiar and conspicuous genus belonging to the Basidiomycota is the *Amanita* Pers.. This genus comprising about 400 species worldwide ranging from edible to deadly poisonous fungi (Weiß *et al.* 1998). Many mycologists have contributed to the systematics and taxonomy of the group, since Persoon introduced the genus in 1797, splitting it into smaller genera (Roze 1876; Earle 1909; Gilbert 1940) or suggesting infrageneric classification concepts (Gilbert & Kühner 1928; Konrad & Maublanc 1948; Singer 1951; Garcin 1984). These systems are mainly based on morphological characters such as Corner & Bas (1962) and Bas (1969) that proposed a separation of the group into two subgenera, *Lepidella* and *Amanita* based on spore amyloidity (spore reaction with Melzer's reagent), cap striation, and form of lamellule. As result of this division four sections were recognized within *Lepidella*: *Amidella*, *Validae*, *Phalloideae*, *Lepidella* and two sections within *Amanita*: *Vaginatae*, and *Amanita.* Weiß *et al.* (1998) confirmed this division with the analysis of the large subunit ribosomal rDNA (LSU) of fifty Amanita species suggesting a further division in the section *Vaginatae*, becoming

Ceasareae and *Vaginatae.* Currently the search for the more accurate organization continues in the genus *Amanita* being now represented into two subgenera, four sections, seven subsections and eight series according to Neville & Poumarat (2004).

The genus Amanita is very well represented in Iberian Peninsula, with at least 45 taxa (Castro 1998) since Atlantic to Mediterranean biogeographical region. One of the most typical and characteristic is Amanita ponderosa Malençon & R. Heim (Fig 1), an edible mushroom harvested by locals in all of its distribution area.

Fig 1 *Amanita ponderosa* Malençon & R. Heim

1.2 – Lifestyle

As a consequence of the fungal diversity heretofore known, different life styles can be recognized being the resultant of a long evolutionary path, very well adapted to the environment and to the communities (Tedersoo *et al.* 2010). One of these life styles is the mycorrhizal symbiosis with vascular plants. Mycorrhizas are highly evolved mutualistic associations between soil fungi and plant roots (Smith & Read 1997). The most common associations are vesicular-arbuscular mycorrhizas, (VAM) also known as endomycorrhizas (Fig 2a) The endomycorrhiza occurs when fungi produce arbuscules, hyphae and vesicles within root cortex cells while ectomycorrhizas fungi (EcM) (Fig 2b) form a mantle around roots and between root cells (Brundrett 2009). In this symbiosis, neither the root nor the fungus function independently but form a unit with adapted metabolic pathways and controlled exchange of metabolites. It is a common notion that in this type of symbiosis, mutual benefit between the partners is due to the exchange of plant-derived carbohydrates in exchange for amino acids, nutrients and water supplied by the fungus (Harley & Smith 1983; Smith & Read 1997).

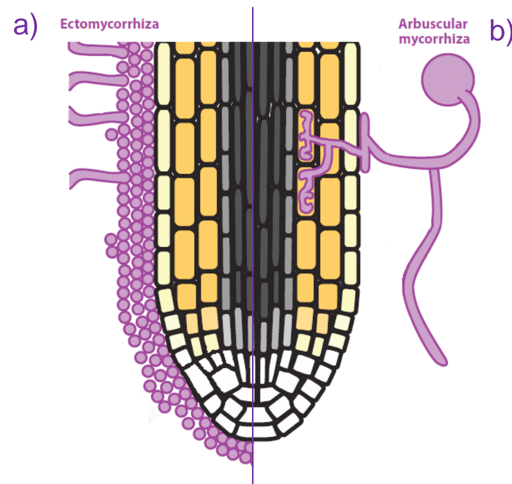


Fig 2 Types of mycorrhiza: (a) Ectomychorriza; (b) Arbuscular mycorrhiza (Bonfante & Anca 2009)

The fungi that produce mushrooms during a part of its life cycle, and in which the genus *Amanita* is included, are mainly ectomycorrhizal (Morris *et al.* 2008; Egli 2011). The mushroom it-self, is just a result of a production of aggregate hyphae, which are the filamentous and vegetative stage of the fungus that produce pseudo tissues with differentiated compartments, developing specialized structures, and eventually differentiate meiotic spores. Summarizing the fruit body of the fungi (mushroom) is just a part of the fungi life cycle (Fig 3) with a principal concern on sexual reproduction, ensuring the spore production (Bon 2004).

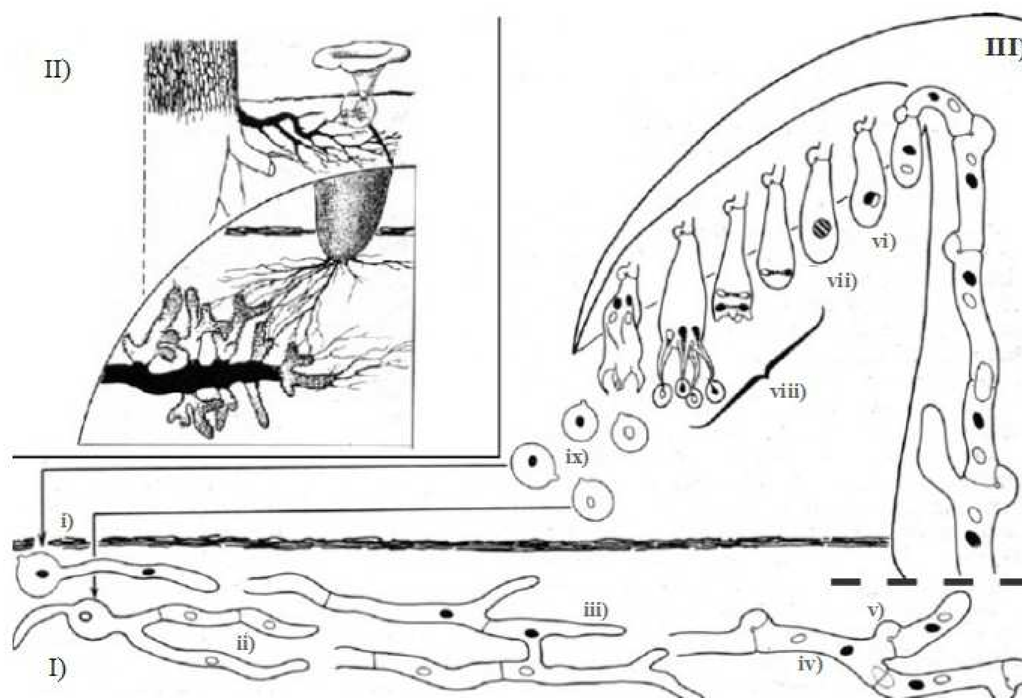


Fig 3 Mushroom life cycle. Stage I) vegetative, II) symbiotic, III) reproductive. i) spore germination (n); ii) hyphal growth (primary mycelium); iii) plasmogamy; iv) mycelium growth (secundary); v) anastomosis; vi) karyogamy; vii) 2n nuclei; viii) meiosis and ix) free spore. (Adapted from Courtecuisse & Duhem 2011)

The EcM fungi life cycle still being discussed (Verbruggen & Kiers 2010; Simard *et al.* 2012) but there are three stages already in consensus: (I) a vegetative stage which corresponds to the hyphal growth in the underground soil ecosystem; (II) a symbiotic stage when the mycorrhizal association is established; and (III) a reproductive stage leading to the organization of fruiting bodies (Murat *et al.* 2008).

In A. ponderosa scenario, the 3rd stage of its life cycle lead to the formation of the semi-hypogeous fruit body, with partial development underground (Fig. 4a) in the beginning of the fruiting resulting in subsequent maturation above ground (Fig. 4b; Courtecuisse & Duhem 2011).

Fig 4 *A. ponderosa* semi- hypogeous development: a) partial development underground; b) maturation above ground

A. ponderosa belong to the symbiotic group of *Amanita sp*, and is described that is usually hosted in Mediterranean forests of *Quercus ilex*, *Q. suber*, Cistaceae shrub type (*Cistus ladanifer, C. crispus, C. monspeliensis* and *C. salvifolius*), *Arbutus unedo*, *Myrtus communis*, *Erica arborea*, *Phyllyrea sp.*, and rarely in forest of *Eucalyptus sp.*, and *Pinus sp.* always together with a cistaceae shrubs (Moreno *et al.* 2007), very typical from the south-west corner of the Iberian Peninsula.

The environmental and climatic variables appear to influence not only their presence or absence, but also their fruit body production. Some of these variables are the precipitation, atmospheric and soil temperature as well as the water availability in the soil (Daza *et al.* 2007). *A. ponderosa* is a spring species and its fruiting period is comprised from January to April, getting the climax production in March and occasionally can occur in October (Santos-Silva *et al.* 2011). The fruiting can occur isolated or more frequently in groups (Tulloss 2005).

Henriques (2010) identified very well these bio-climatic patterns and its influence in *A. ponderosa* fruit-body production along four years in Cabeço de Mouro – Natural Park of International Tagus River (Portugal). He reported that is crucial the occurrence of at least

~15mm of precipitation during the 15 days before fruiting period and the temperature values superior to 12ºC to guarantee a continuous production flow.

1.3 – Habitat and Distribution

A. ponderosa is a typical heliophile species occurring usually in acid pH soils predominantly with a strong component in schist's rocky types. Occasionally it can arise in sandy and deeper soil types. These soil types are usually depleted in organic matter contents (Costa *et al.* 1998). Because of the specificity in the mycorrhizal associations with the species mentioned above (see section 1.2 Life style) the occurrence *A. ponderosa* is mostly reported in the montado ecosystem (Fig. 5a) or in typical Mediterranean Shrub land (Fig 5b) (Azul *et al.* 2009). The montado ecosystem is characterized by agroforestry areas in open woodlands with holm oaks and cork oaks with extensive areas of rock-rose. The extensive regime of management of these systems allied to the high flora diversity turn this habitat of excellence to *A. ponderosa* and also to other mushroom species (Azul *et al.* 2009; Louro *et al.* 2009; Morgado *et al.* 2011; Arraiano-Castilho *et al.* 2013).

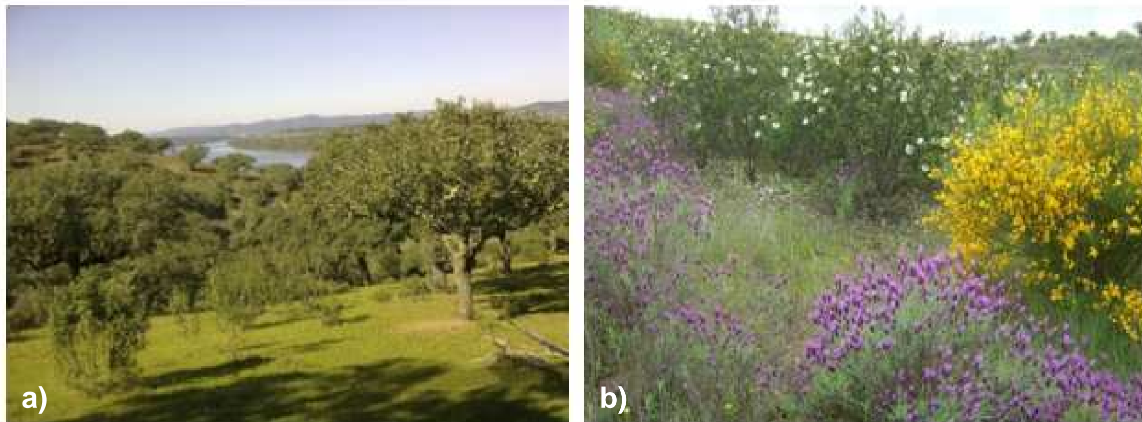


Fig 5 *Amanita ponderosa* typical habitat: a) open woodland montado, b) Mediterranean shrub land

According to Curreli (1994) and Daza *et al.* (2002) the distribution area of A. ponderosa is restricted to southwest of Iberian Peninsula, some regions in North Africa, South France and Italy mainly in Sardinia and Sicily. The same author's classify this species as endemic from these regions (Fig 6) supported by the local harvest records.

Fig 6 *A. ponderosa* spatial distribution. Light green represent the occurrence data from herbarium and historical records; dark green represents where the harvest is most intensively with high socio-economic importance (adapted from Moreno *et al.* 2007)

The distribution of *A. ponderosa* presented nowadays is the result observation of the places with greater harvest tradition. Therefore to better understand the species behaviour and its interaction with the environment it is extremely important determine species current distribution (Austin 2002) that still unknown. The most common technique to predict a species occurrence is the species distribution modelling (SDM) also known as climate envelope-modelling, habitat modelling, or environmental or ecological niche-modelling (Sillero 2011). The most commonly used are the correlative distribution models (Kearney *et al.* 2010). That predict the realized niche of a species (Guisan & Zimmermann 2000; Pearson & Dawson 2003) by associating spatial environmental data with species abundance or presence records (presence-true-absence records; presence-pseudoabsence; presence-only) to provide values of suitability for each point in space (Sillero 2011). Thus there are three categories of correlative models depending on the type of data used: presence-absence, presence-pseudoabsence and pesence-only models.

Presence-absence models relate the presence or the absence of a species with a set of ecogeographical variables modelling the suitable conditions for the presence of a species and the unsuitable conditions for its absence. Presence-pseudoabsence models use only presence records and attribute suitable conditions to areas where a species

record exists and pseudoabsence conditions to areas where no records exist (Phillips *et al.* 2006). Presence-only models uses only presence records and overlap them with maps of the ecogeographical variables deducing the range of conditions suitable for species survival. These models may be advantageous over presence-absence ones because the latter are prompted to introduce errors since absences may have traces of biotic interactions, dispersal constraints and disturbances which may impede the accurate modelling of distributions. Presence-only models may also bring some inaccuracy in species predicted distribution since species may be absent from an area with suitable conditions due to past events that caused local extinctions (Elith *et al.* 2011).

1.4 – Species complex (taxonomy)

The organization in section *Amidella* (Weiß *et al.* 1998) of the genus *Amanita* remain unclear, especially regarding the species that corresponding to the *Lepiotoides* complex. This complex contains *Amanita curtipes* E.J. Gilbert, *Amanita lepiotoides* Barla and *Amanita ponderosa* Malençon & R. Heim (Fig 7). These species were the subject of a study conducted by Pinho-Almeida (1994) where she attempted to characterize them evaluating macro and microscopic characters. According with the author of this study this species complex is very well established by the characteristics of the species which it is composed. However, when one intends to distinguish each species within the complex difficulties arise when trying to set boundaries that distinguish them as species. These difficulties come from the different characterizations from different authors thus making it difficult the correct species identification. The concerns begin more evident when one intends to distinguish between *A. ponderosa* and *A. curtipes* due to the possibility of habitat overlapping and coincidence of the fruiting period. Through the analysis of the complete ITS and D1-D2 regions of ribosomal DNA (rDNA), Moreno *et al.* (2007) show that *A. ponderosa* and *A. curtipes* are clearly two distinct species and discard the controversial treatment of *A. ponderosa* and *A. curtipes* as synonyms or varieties of a single species, as suggested by some authors. Castro (1997) included *A. curtipes* and *A. valens* (J.-E. Gilbert) Bertault (another species of this complex) in a single species*, A. curtipes*, and proposed a new combination for *A. ponderosa* as a variety of *A. curtipes*, *A. curtipes var. ponderosa*. Other authors have suggested that *A. ponderosa* is just a giant form of *A. curtipes* (e.g. Kühner & Romagnesi 1953; Mesplède 1980).

Neville & Poumarat (2004) in his landmark work about Amanitaceae had previously considered *A. ponderosa* and *A. curtipes* as two distinct species defining two different forms in *A. ponderosa*, *Amanita ponderosa f. ponderosa* Malençon & R. Heim and *Amanita ponderosa f. valens* (E.-J. Gilbert) with some differences mainly on the size of fruit body. All these difficulties demonstrated by these authors are very illustrative of the lack of studies and knowledge about these species.

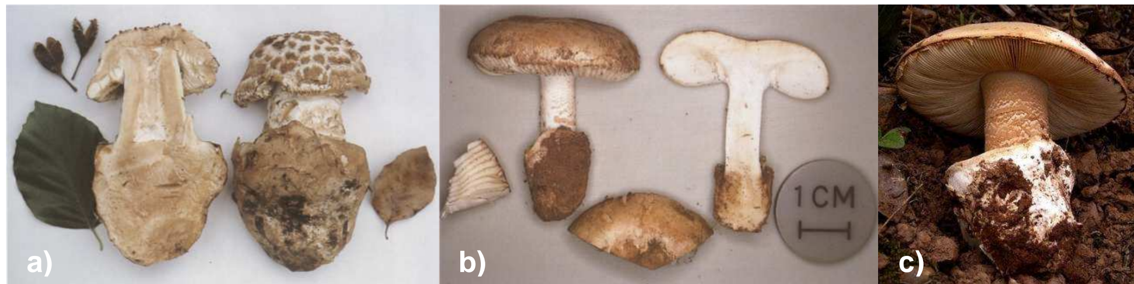


Fig 7 *Lepiotoides* species complex adapted from Neville & Poumarat (2004): a) *Amanita lepiotoides*; b) *Amanita curtipes*; and c) *Amanita ponderosa*

1.5 – Socio-Economic importance

Fungi have been known and used by humans for centuries, but mycology (the scientific study of fungi) traces its beginnings to the 18th century, with the development of the microscope (Ainsworth 1976). Nowadays we know that mycological resource, beyond the important role in the balance of ecosystems, plays others as agents of social and economic development, through the food, biotechnology, tourism and other economic activities (Castro 2009).

A. ponderosa is considered the wild mushroom par excellence in the areas where it occurs. Local people use this mushroom for gastronomic purposes being a delicacy much appreciated. When sold, the price can be very variable, depending on supply and demand, reaching at the beginning of the season values between 25 and 30 € (euros) per kg, decreasing to 10 to 15 € per kg at the middle of the season. An experienced local harvester can reach 15 to 20 kg of *A. ponderosa* per week making it an important source of extra income.

The purchase and sale of wild mushrooms behaves differently in the two countries where the demand of this resource is the highest, Portugal and Spain. In Portugal there is no authority of regulation or certification of this market. Direct sales occur between harvesters and consumers, with the existence, in many cases, of intermediate resellers buying from many small harvesters and selling in both the domestic and the exporting market.

A very first consequence of the absence of regulation, supervision and monitoring of the buying and selling of this species are the constant poisoning cases that occur annually. Morgado *et al.* (2006) studied some cases of poisoning by ingestion of mushrooms that were admitted in the emergency service of the district hospital (Hospital Espírito Santo in Évora, Portugal) and observed that most of the cases were due to confusion between *A. ponderosa* and other toxic species.

In Spain have also been reported some cases of poisoning derived from the confusion of *A. ponderosa* with others macroscopically similar (e.g. *Amanita verna*, *Amanita phalloides var. alba*) (Piqueras-Carrasco 2013). However when it comes to regulation and market certification, Spain positively stands out relatively to Portugal. As this country does not have a specific market classification for wild mushrooms, these are certified and sold in the category of fruits and vegetables (Fig. 8) (*in* Real Decreto 30/2009 – 16/01)

Fig 8 *Amanita ponderosa* selling in Huelva Market

1.6 – State of the art: The role of new approaches

During the last decades research in *A. poderosa*, as well as other mushroom species focused in taxonomic and systematic issues with the propose of catalogue the fungal diversity (Bridge *et al.* 2005). Nowadays mycology is less of a catalogue process than it once was. DNA-based molecular techniques as the polymerase chain reaction (PCR) played a key role in discovery of molecular markers providing new tools for research (Bellemain *et al.* 2010)

The most popular locus for DNA-based mycological studies is the internal transcribed spacer (ITS) region of the nuclear ribosomal repeat unit (Fig.9) (Horton & Bruns 2001; Bridge *et al.* 2005). The ITS region includes the ITS1 and ITS2 regions, separated by the 5.8S gene, and is situated between the 18S (Small subunit-SSU) and 28S (Large subunit-LSU) genes in the rDNA repeat unit. The large number of ITS copies per cell makes the region an appealing target for sequencing when the quantity of DNA present is low (Bellemain *et al.* 2010). This multi-copy, tripartite segment combines the advantages of resolution at various scales: ITS1 rapidly evolving, 5.8S: very conserved, ITS2 moderately rapid to rapid; (Hillis & Dixon 1991; Hershkovitz & Lewis 1996). This locus has been used massively not only with regard to phylogenetics and taxonomy, but also in connection to development of diagnostic strategies for species identification (Nilsson *et al.* 2008; Begerow *et al.* 2010; Conrad *et al.* 2012). Currently, ~172,000 full-length fungal ITS sequences are deposited in GenBank, and 56% are associated with a Latin binominal, representing ~15,500 species and 2,500 genera, derived from ~11,500 scientific studies in ~500 journals (Schoch *et al.* 2012).

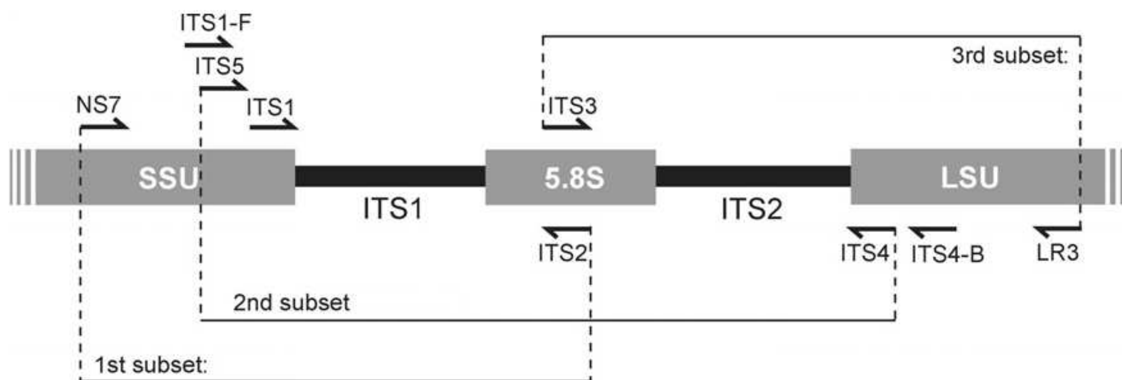


Fig 9 Nuclear ribosomal repeat unit with the most common used set of primers (Bellemain *et al.* 2010)

The other methodology for species identification widely used in fungal research was the PCR-RFLP (Restriction Fragment Length Polymorphism) technique, coupling two known procedures to detect polymorphisms in DNA regions which have been amplified by specific oligonucleotide primers and restricted with different endonucleases where each species shows a specific cutting pattern (Gardes *et al.* 1991; Erland *et al.* 1994; Henrion *et al.* 1994; Farmer & Sylvia 1998; Gomes *et al.* 1999; Glen *et al.* 2001).

Presently, the large improvement of widely use of the Next-generation DNA sequencing techniques triggered the design of new tools that certainly permit a more complete understanding of numerous process that we never been able to work on (Mardis 2008). This great leap forward, not only permitted to solve the old and classical riddles but went beyond encouraging the scientific community to pose new questions.

An excellent example of these new tools is the restriction site associated DNA sequencing (RADseq) (Miller *et al.* 2007), also called Genotyping by Sequencing – GBS (Elshire *et al.* 2011). RAD markers were initially used together with low cost microarray genotyping resources (Miller *et al.* 2007), but the advent in next-generation sequencing technologies and concomitant drop in sequencing costs lead to the integration of short-read sequencing with RAD genotyping (Baird *et al.* 2008). This sequencing technique provides an efficient method to discovery thousands of single nucleotide polymorphisms (SNPs) (Rowe *et al.* 2011) markers that just change a single base in a DNA sequence, with a usual alternative of two possible nucleotides at a given position (Vignal *et al.* 2002). The applications of RAD sequencing technique are numerous divided into investigation of genome organization and population-level studies aimed at understanding the organization of intra-specific variation, identification of population structure, regions experiencing selection in particular environments, migration patterns and speciation (Rowe *et al.* 2011).

The RAD markers discovery is now intrinsically related with the Illumina sequencing approach. Rowe *et al.* (2011) review give us a good perspective how these markers can be generated (Fig. 10). In small steps, RAD-Seq is a complexity reduction system, digesting the genome with a restriction nuclease and attaching a series of adapters to the resulting DNA fragments, thereby large numbers of genetic variations such as SNPs can be readily identified from analysis of high-throughput sequence data.

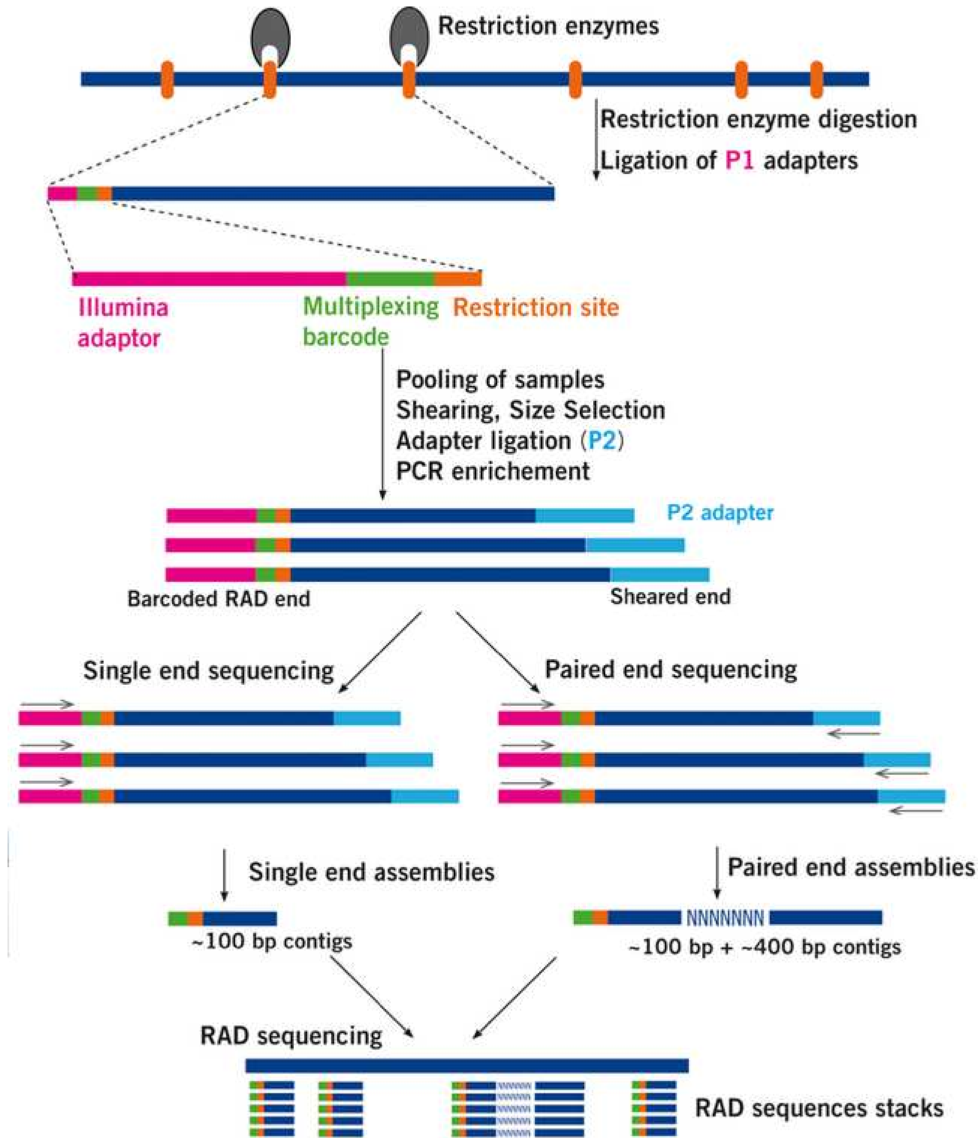


Fig 10: Restriction-site associated DNA sequencing. At the beginning genomic DNA should be digested with specific and very well known endonucleases in small fragments. Comprising that a modified adapter containing the Illumina P1 amplification and sequencing primer and a DNA barcode will be ligated to these fragments. After that, samples are pooled, sheared into 300 to 800bp libraries and ligated to a second adapter P2. Sequencing is performed as single end (one sequence of 36-108 bp per fragment) or paired end (two sequences of 36-108 bp per fragment). In single end assemblies barcoded sequences are assembled into overlapping stacks. Paired-end sequencing allows the assembly of larger contigs on the sheared end of the fragment, whose size depends on the length of the Illumina sequences and the size of the sheared fragments isolated. Adapted from: http://floragenex.com/floragenex-radseq-bioinformatics-genotyping-technology.php

1.7 – Objectives

The lack of knowledge about the *Lepiotoides* species complex still being the principal reason for the contempt of this resource. Therefore, the main goals of this study are:

- Identify a significant number of SNP markers in *Amanita ponderosa;*
- Infer about the species complex to better understand the differences between each species;
- Analyse the phylogeographic patterns in *Amanita ponderosa* using SNP markers;
- Identify the most important factors related to *Amanita ponderosa* occurrence;
- Determine the *Amanita ponderosa* current distribution in Iberian Peninsula;
- Locate priority areas for species conservation.

2. Manuscript I

Next-generation RAD sequencing: a tool for evaluation of population structure of *Amanita ponderosa* in Iberian Peninsula

2.1 Abstract

Amanita ponderosa is the most charismatic fungal species in Iberian Peninsula. Despite its high socio-economic value little is known about this species. Here is presented a study that contemplates a way to better understand the relation with between *A. ponderosa* and sister species and infer some phylogeographic patterns in Iberian Peninsula. For better understand the genetic architecture of this species in Peninsula, a battery of SNP markers were generated through RAD-sequencing. It has found application in wild populations and non-traditional study species, and promises to become an important technology for ecological population genomics. A cryptic species from the *Lepiotoides* species complex was identified and classified as *Amanita pseudovalens* through the analysis of 247 SNP. Two populational clusters were found in that species. *Amanita ponderosa* don't shown any population structure across Iberian Peninsula through the analysis of 201 SNP. A distinct cluster was identified corresponding to *Amanita ponderosa f. valens*. RAD sequencing proved to be a powerful tool to identify loci in that mushroom species. The main cause for the unstructured populations in these species is due to the high spore dispersal leading to great connectivity between all populations.

.

Keywords

Amanita ponderosa; RAD sequencing; Phylogeography

2.2 Introduction

Phylogeography seeks to explain the geographic distribution of genetic lineages along species evolutionary history. The last two decades it has experiencing the boom of the PCR and Sanger sequencing era (Puritz *et al.* 2012). Recent improvements in the speed, cost and accuracy by the Next-Generation sequencing (NGS) and advances in the accompanying bioinformatic tools are revolutionizing the opportunities for generating numerous genetic resources (Helyar *et al.* 2011). These advances in sequencing have radically expanded the reach of genetic studies to non–model organisms and wild populations, making this an exciting area to watch. Now the type and amount of data generated leads us to the genomics era exploring all genome and leaving behind the era

of genetics with high knowledge and a model organism requirements (Nosil & Buerkle 2010). The revolution of the population genomic analyses require multi-locus data sets from multiple populations and allow identify non-neutral or outlier loci by contrasting patterns of population divergence among genetic regions. Generate and analyse these data sets is nowadays a promising field of research (McKenna *et al.* 2010; Nekrutenko & Taylor 2012)

A novel and efficient genotyping approach for marker generation based on Illumina sequencing of libraries is Restriction-site Associated DNA (RAD) (Miller *et al.* 2007; Baird *et al.* 2008). RAD sequencing, a NGS based technique that generates short sequence reads at thousands of regions adjacent to restriction endonuclease recognition sites across the genomes of multiple individuals (Hohenlohe *et al.* 2011). This technique allow us to identify a large number of SNPs across genome and as being used in a broad range of studies in non-model organisms with different proposes, such as hybridization in rainbow trout (Hohenlohe *et al.* 2011) and parallel adaptation in threespine stickleback (Hohenlohe *et al.* 2010).

Despite nowadays already be quite frequent the approach to non-model organisms, it was not always like and relatively little is known about the patterns and distributions of SNPs in these organisms, including most fungi, thus opening up an excellent research opportunity. Here is presented a population genomic analysis in one of the most prominent wild edible mushroom of Iberian Peninsula. *Amanita ponderosa*, also known as silarca, púcarinha, tortulho, gurumelo between others. Endemic from south-west peninsula, some regions of north Africa and west coast Italy (Daza *et al.* 2007) and ectomicorhyzical of cork and holm oaks preferentially, very typical in Montado ecosystems characterized by agroforestry areas in open woodlands (Moreno *et al.* 2007). The fruiting period of this species occurs between January and April reaching the climax in March when precipitation and temperature conditions are more favourable (Henriques 2010). Very few studies has been conducted in this and other mushrooms species constraining our knowledge about the effect of the harvest pressure in the fitness of populations and how this interferes with the balance of the ecosystems. Xu *et al.* (2007) is one of the few that described and discovery 178 SNP sites in *Tricholoma matsutake*, a wild edible mushroom with high expression in East Asia. These SNPs were identified through analysis of a shotgun genomic library in order of evaluate the population structure and shown the effect of harvest and habitat loss in this species. Other studies like the one conducted by Pringle *et al.* (2009) demonstrated the invasive character of the deadly poisonous mushroom *Amanita phalloides* in North America through the analysis of a microsatellite data set providing a strong evidence for a European origin of North American populations. Like others mushroom species, *A.*

ponderosa is highly and intensive harvested by locals due to its large gastronomical value. At present the absence of good harvesting practices practiced by locals may be problematic for species conservation. Here arises a good research opportunity to evaluate the Iberian populations fitness conditions. For this the main goals of this study are: i) identify a set of SNP markers in *Amanita ponderosa* through RAD-sequencing; ii) infer about the *Lepiotoides* species complex to better understand the differences between each species and iii) analyse phylogeographic patterns in *Amanita ponderosa* using SNP markers.

2.3 Materials and methods

2.3.1 Sampling

The sampling was performed in order to cover all known spatial distribution of species in Iberian Peninsula according to Daza *et al.* 2007. Sampled regions were Baixo-Alentejo, Alto Alentejo, Ribatejo and Beira Baixa in Portugal and Extremadura and Andaluzia in Spain (Fig. 11). Two hundred and sixty nine samples were collected between February and April 2013 and preserved partially in 96% ethanol for posterior DNA extraction. All collected samples were dried at lower temperature for posterior herbarium storage. Seventeen external Herbarium specimens were joined to the data set (10 from Herbarium Real Jardin Botânico-Madrid, 4 from Dept. of Animal and Plant Biology and Ecology of University of Jaen and 3 from personal collection of Nicola Amalfi (Sicily - ITA). Samples of two different species were added, one *Amanita ovoidea* and one *Amanita verna* from personal collections for outgroup propose. All samples used in this study are illustrated in Tab 1. To avoid self-sampling of the same individual each sample was collected in a range of 5 meters between them (Taylor *et al.* 2012)

2.3.1 DNA extraction & quantification

DNA was extracted from dried and 96% ethanol preserved samples using the JETquick Tissue DNA Spin Column DNA extraction kit (Genomed GmH; Bad Oeynhausen, Germany) with some modifications. Samples preserved in ethanol were extracted with a digesting period extended to 48 hours in 200µl of extraction buffer T1 + 25µl Proteinase K at 56ºC with a reinforcement of 25µl Proteinase K after 24 hours. The pellet and undigested material were rejected and the supernatant transferred to a new tube. Then 200µl of T2 were added (mixing by gentle inversion) and incubated at 70ºC for 10 min. After 1 minute at room temperature 200µl of ethanol at -20ºC were added and the entire volume transferred to a Spin column to centrifuge at 8000 rpm for 1 minute. The flow

Table 1 List of samples used in this study. *n* represents the number of samples collected by site; *Host* indicates the most probably tree associated; *Land cover* the shrub type; *Soil type*, the main composition of the soil and finally how the samples were stored before the DNA extraction. Data deficient represented by "-"

Location	Cod	*n*	Host	Land cover	Soil type	Storage
Aracena	Ap_arc	16	*Q. ilex*	*C.ladanifer*	Schists	ETOH
Cabeça Gorda	Ap_cbg	16	*Q. suber*	*C.salvifolius*	Schists	ETOH
Erra	Ap_cco_e	19	*Q. ilex*	*C.ladanifer*	Sandy	ETOH
Pavia	Ap_cco_p	5	*Q. ilex*	*C.ladanifer*	Schists	ETOH
Moura (contenda)	Ap_cntd	21	*Q. ilex & P. pinea*	*C.ladanifer*	Schists	ETOH
Malpica	Ap_mlpc	13	*Q. suber*	*C.ladanifer*	Sandy	ETOH
Montargil	Ap_mntg	7	-	-	-	ETOH
Mértola	Ap_mrtl	30	*Q. ilex*	*C.ladanifer*	Schists	ETOH
Odemira	Ap_odm	14	-	-	-	ETOH
Proença-a-Nova	Ap_pan	26	*Q. ilex*	*C.ladanifer*	Granitic	ETOH
Portel	Ap_prl	17	*Q. suber*	*C.ladanifer*	Schists	ETOH
Redondo	Ap_red	6	*Q. ilex*	*C.ladanifer*	Schists	ETOH
São Barnabé	Ap_sbe	28	*Q. suber*	*C.ladanifer*	Schists	ETOH
Talayueda	Ap_tlyd_t	4	*Q. ilex*	-	Sandy	ETOH
Robledillo-de-la-Vera	Ap_tlyd_r	9	*Q. pyrenaica*	-	-	ETOH
Valencia d'Alcántara	Ap_vda	19	*Q. ilex*	*C.ladanifer*	Schists	ETOH
Vilanueva del Fresno	Ap_vnf	10	*Q. ilex*	*C.ladanifer*	Schists	ETOH
Tomar	Ap_tmr	9	*P. pinea & Eucalyptus*	*C.ladanifer & salvifolius*	Schists	ETOH
Huelva	Ap_huel	2	-	-	-	Dried
Cicilia (ITA)	Ap_ita	3	-	-	-	Dried
Jaen	Ap_jaen	2	-	-	-	Dried
Salcé	Ap_zam	1	*Q. ilex*	*C.ladanifer*	-	Dried
Membrio	Ap_mem	1	*Q. ilex*	-	-	Dried
Villa del Prado	Ap_vdp	2	*Q. ilex*	-	Sandy	Dried
Valverde del Camino	Ap_vlvc	1	*Q. ilex & Q. suber*	-	-	Dried
Azuaga	Ap_ba	1	*Q. ilex*	*C.ladanifer*	-	Dried
Salorino	Ap_sal	1	*Q. suber*	-	-	Dried
Sigüenza	Ap_sig	1	-	-	-	Dried
Arenas de San Pedro	Ap_asp	2	*Q. ilex*	-	-	Dried
Portel	Ao_prl	1	*Q.suber*	*C. salvifolius*	Limestone	ETOH
Couço	Av_cço	1	*Q.suber*	*C. salvifolius*	Sandy	ETOH
	Total:	288				

Fig 11 Sampling sites in Iberian Peninsula represented by Blue dots

was discarded and the centrifugation step was repeated with the addiction of 500 µl of standard solution TX at 13000 rpm for 3 minutes. At the end the spin columns were eluted with warm elution buffer (10 mM tris (hydroxymethyl) aminomethan hydrochloride) at 70ºC for 5 min.

Quantification of extracted DNA was performed with Qubit 1.0 (Invitrogen, Carlsbad, CA, USA)

<u>2.3.2 Species Identification</u>

All samples were identified in the field by macro-morphological characteristics according to Pinho-Almeida (1994); Moreno *et al.* (2007). Posterior species confirmation was performed by DNA sequences analysis according to Moreno *et al.* 2007. The ITS region from rDNA (ribosomal DNA) was amplified according the following procedure: Each 20-µl reaction consisted of PCR-water, DNA (2µl), primers (0.6 µM primer forward ITS1-F: CTT GGT CAT TTA GAG GAA GTA A (Gardes & Bruns 1993) and 0.6 µM primer reverse ITS4-B: CAG GAG ACT TGT ACA CGG TCC AG (Gardes & Bruns 1993)), dNTPs (10 mM each), 10× buffer [200 mM Tris–HCl (pH 8.4), 500 mM KCl], BSA (0.5 µg/µl), MgCl2 2 mM and Platinum® Taq DNA Polymerase ((0.3 U); Invitrogen™). PCR reaction was performed in Applied Biosystems Veriti 96 well Thermal cycler according with the following program: 10min of initial denaturation at 94ºC following 35 cycles with 30s

Desnaturation at 94ºC, 30s Hybridation at 57ºC and 30s Extension at 72ºC. Finally 10 min at 72ºC for final extension.

Consensus sequences were adjusted with SeqMan package of DNAStar software (Lasergene) and aligned following Clustal W algorithm in MEGA5 v. 5.2 (Tamura *et al.* 2011). All sequences were blasted against reference sequences in NCBI – GeneBank.

2.3.3 Restriction site associated DNA library preparation

RAD library preparation followed the protocol of Baird *et al.* (2008); Hohenlohe *et al.* (2010) with further modifications. Around 150ng (5ng/ul) of genomic DNA from each individual was eluted in a digestion mix (4.5ul H_2O, 5ul NEBuffer 4 (10x) and 0.5ul of SbfI restriction enzyme) at 37ºC for 60min. The enzyme was inactivated after at 65ºC for 20min. Next, 2ul of the appropriate barcoded SbfI P1 RAD adapter (50 nM) was added to each sample and 8ul of the corresponding ligation mix (5.9ul H_2O, 1ul NEBuffer 4 (10x), 0.6ul rATP (100nM, Fermentas R0441) and 0.5ul T4 DNA ligase (NEB M0202M)). Again incubated at 20ºC for 60min and the enzyme inactivated at 65ºC for 20min. All the samples sequenced together were pooled in the same library according to Miller et al., (2012). The pooled samples were submitted to Bioruptor on high with cycles of 30 seconds on and 1 minute off for 15 minutes to produce fragments with an average size of 500 bp. All the entire sample was loaded in a 1% agarose gel to extract from each pool the fragments corresponding to 400-600bp followed by purification with a Qiagen MinElute column. The column was after eluted with 20ul of elution buffer. To blunt and repair each pool the following reagents were added by the following order, 2.5ul Bluting buffer (10x), dNTP mix (1nM) and Bluting Enzyme Mix (NEB E1201L). Incubation at 20ºC for 60 minutes. The fragments were after purified with Quiagen MinElute PCR Purification Kit and eluted in 43ul of elution buffer. The A-overhangs were added to the fragments by adding the following reagents to the pool in order, 5ul NEBuffer 2 (10x), 1ul dATP (10nm), 2ul Klenow Fragment (3’-> 5’ exo-) (NEB M0212L) and incubated at 37ºC for 60minutes. A new purification process was performed with Quiagen MinElute PCR Purification Kit and eluted with 44ul of elution buffer. The P2 adapter was ligated to the fragments by adding the following reagents to the pool in order, 5ul NEBuffer 2 (10X), 1ul P2 RAD adapter (10uM), 0.5ul rATP (100nM, Fermentas R0441), 0.5ul T4 DNA ligase (NEB M0202M) and incubated at 20ºC for 30minutes. Another purification with Quiagen MinElute kit with final elution of 50ul.

Final a PCR reaction was performed according with the following conditions: 98ºC for 30sec, 14x (98ºC for 10sec, 65ºC for 30 sec, and 72 º C for 30 sec), 72ºC for 5min, hold at 10ºC. The PCR mix prepared was 38ul H_2O, 50ul 2X Phusion Master Mix (NEB F-531L), 4ul P1 adapter primer (10uM), 4ul P2 adapter primer (10uM) and 4ul of purified

pool sample from last step. The PCR product was purified with Quiagen MinElute kit with final elution of 14.5ul and loaded in agarose gel for low molecular weight adapter and primer junk removal by cut. The fragment removed was around 400-600bp. Quiagen Buffer QG was used to dissolve the gel chunk. One last purification procedure was conducted with the Quiagen MinElute column. To the final elution with 14.5ul was added 1.5ul of elution buffer containing 1% Tween-20.

In the end the pool sample was quantified and diluted with elution buffer containing 0.1% Tween-20 in order to obtain a final concentration of 10nM.

2.3.4 Marker generation

First of all the individuals were identified according with specific barcode sequence. The SNP discovery and genotyping was performed using custom Perl scripts and the alignment program Novoalign (Novocraft Technologies, Selangor, Malaysia). Bad quality reads and with potential sequencing errors were filtered running QualityFilter script. After run the quality filter was selected a sub-set of 125K of sequences from each sample, to reduce even more potential sequencing errors. Candidate SNP were selected from a subset of individuals selected based on the frequency histogram (number of sequences/frequency) and applying a cut line of 20 to filter bad quality sequences. Nine individuals were selected from *Amanita sp.* group and 13 individuals from *Amanita ponderosa.* SNP discovery was developed aligning all the selected individuals each other within groups. The posterior operation involves searching for these SNP in remaining individuals of each group aligning them with the previously selected group

2.3.5 Population genetic and phylogenetic analyses

Preliminary phylogenetic analysis based on rDNA was performed in MEGA5 and phylogenetic tree produced according to Maximum Likelihood approach with 1000000 bootstrap replication. Thirteen other sequences present in NCBI-GeneBank from different amanitas were included representing the major sections of the genus *Amanita*. Section Amanita was represented by *Amanita muscaria* EU071957.1 and *Amanita pantherina* AB015701.1, Section Volvatae sub-section Validae with *Amanita rubescens* AJ889922.1 and *Amanita flavoconia* AY325847.1, sub-section Phalloidinae with *Amanita phaloides* GQ221843.1 and the series focus of this study, Section Volvatae, sub-section Validae, series Amidella, represented by *Amanita ponderosa* AY486234.2; AY486233.2 and EF653962, *Amanita curtipes* EF653963.1, AY486235.2, AY486236.1, *Amanita volvata* JF723273.1 and *Amanita avelanosquamosa* AY436447.1. *Limacella glioderma* FJ478086.1 was used to root the tree.

The covariance standardized method of a Principal Coordinate Analysis (PCoA) was conducted in GenAlex (Peakall & Smouse 2006). The analysis of population structure was performed in STRUCTURE 2.3.4 (Pritchard *et al.* 2000) with the following parameters; Burnin sequences 50000 and repeat Monte Carlo Marcov Chain 100000; establish a K=1 to K=6.

2.4 Results

2.4.1 Species Identification

Amplification and sequencing of ITS1, 5.8s and ITS2 was successful for al specimens. The size of DNA fragments varies between 790 and 810bp being last the most common. The Blast analysis in NCBI-Genebank results in 286 confirmations which 271 were *A. ponderosa* matching with AY486234.2; AY486233.2 and EF653962.1 from Moreno *et al*. (2007) and 17 other amanitas with no reference.

2.4.3 Marker generation

A total number of 247 loci was discovered in a total of seventeen *Amanita sp.* individuals and a total of 201 loci in a total of one hundred sixty-three individuals.

2.4.4 Population genomics and phylogenetic analyses

rDNA – ITS1, 5.8s, ITS2 analysis

The preliminary results of phylogenetic analysis (Fig. 12) display clearly the genus *Amanita* as expected according to Neville & Poumarat 2004. The majority of specimens collected for this study cluster together with the reference sequences described in Moreno *et al*. (2007), forming *Amanita ponderosa* cluster. A new clade has shown representing the specimens that did not meet with any other *Amanita sp.* in preliminary BLAST analysis. According with these results this clade is more related with *Amanita curtipes* than *Amanita ponderosa*.

SNP analysis

The first alignment with all individuals performed in Novoalign highlighted the previous observation in ITS analysis. Again the individuals with no matches in NCBI do not aligned with either belonging to *A. ponderosa* clade, forming together a distinct group.

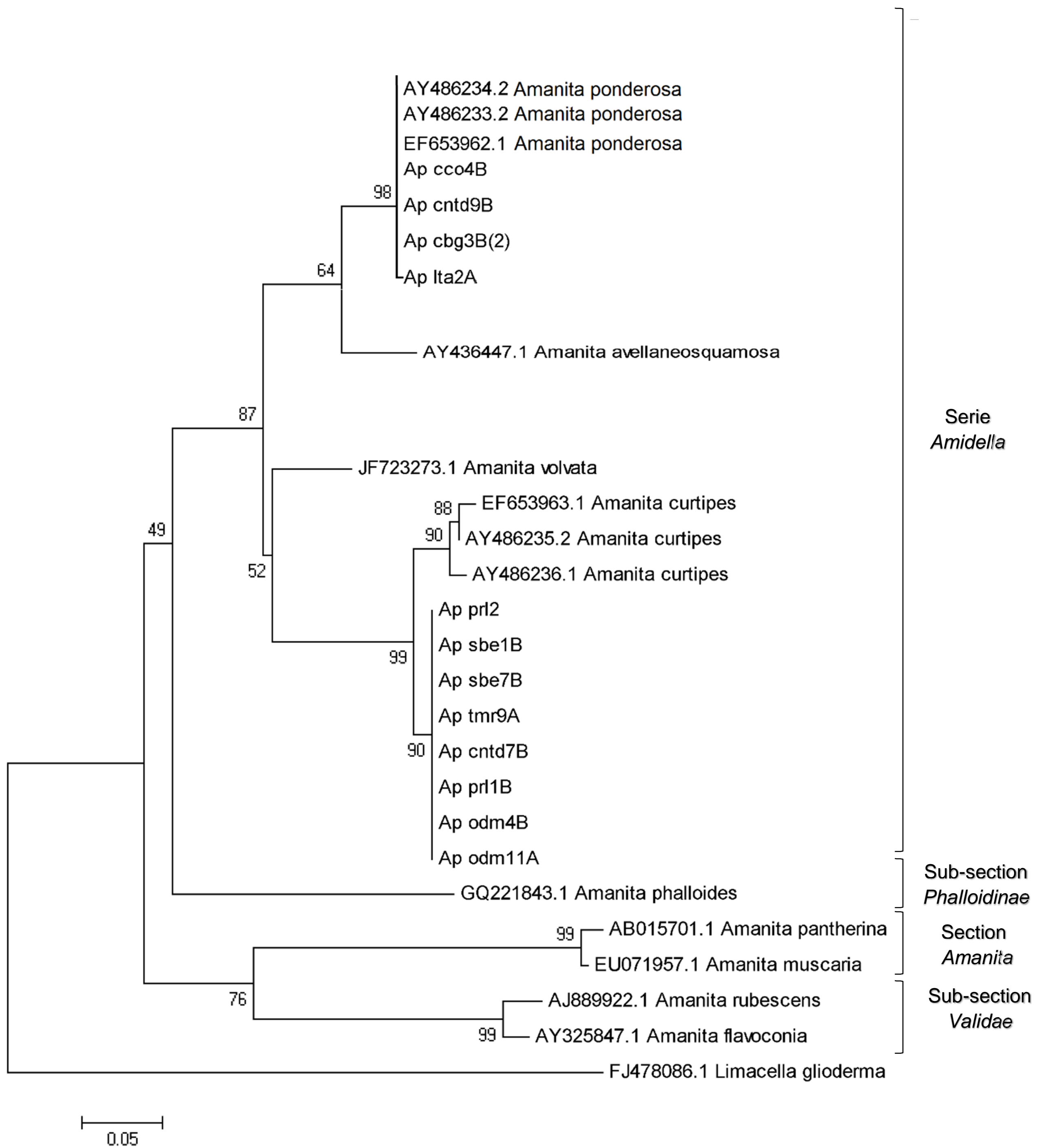


Fig 12 Phylogenetic tree based on ITS1, 5.8s and ITS2 of rDNA, generated in MEGA5 according Maximum likelihood approach, 100000 bootstrap replicates. A sequence of *Limacella glioderma* was used to root the tree. Bootstrap values are indicated above or at the branches. The scale bar indicates the nucleotide substitution rate. Others (274) represent all the others sequences of *A. ponderosa* present in Tab 1.

PCoA

Because the distinction between the two groups the PCoA analysis was performed separately for both groups. Both graphs are plotted below (Fig 13).

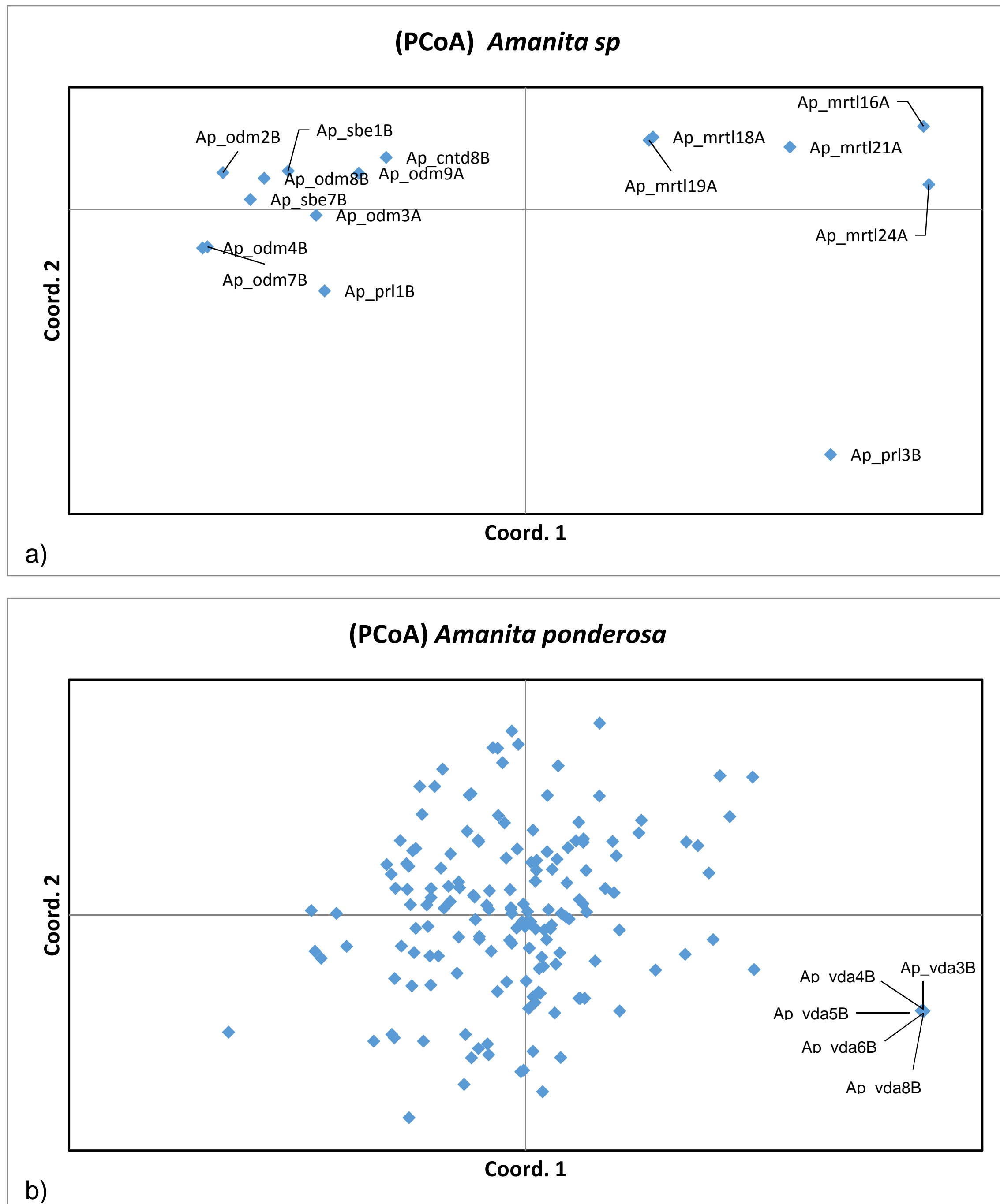


Fig 13 – PcoA outputs from GeneAlex a) *Amanita sp.* and b) *Amanita ponderosa.*

The individual AP_odm11A was removed from the analysis due to low coverage of SNP. PCoA show in the *Amanita sp.* two different groups of individuals and one isolated individual (Ap_prl3B). These two main groups that individuals from Mértola (Ap_mrtl*) are clearly distinct from the others (Ap_odm*, Ap_cntd*, Ap_sbe* and Ap_prl*). The percentage of variance was 33.21%, 17.89%, 12.78% in the 1st, 2nd and 3rd axis.

PCoA from *Amanita ponderosa* illustrates that all the individuals chosen are clustered together except a group of five from Valência de Alcântara (Ap_vda3B, Ap_vda4B, Ap_vda5B, Ap_vda6B and Ap_vda8B). The percentage of variance was 5.92%, 5.21% and 4.89% and it was pretty identical when removing the low coverage loci and individuals.

Structure analysis

The structure analysis was also conduced for diferente species groups. The most likely number of clusters (K) when likelyhood is maximized *Amanita sp.* show two distinct clusters (K=2) (Fig. 14). The indivuals provenient from Mértola (Ap_mrtl) represented by the color red are clearly distinct from the others represented in green. The individual 2 (Ap_prl3B) besides to shown to be an outliers in the previous PCoA analysis shown to be more related to Mértola cluster than the other.

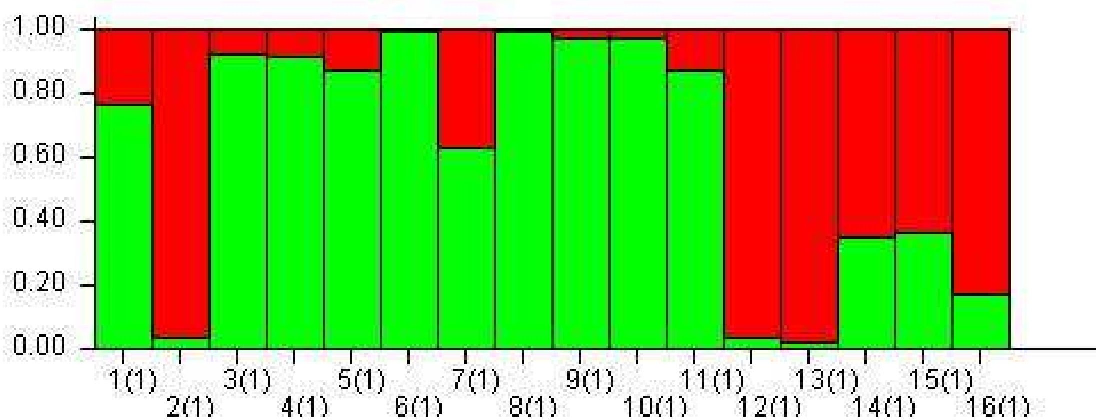


Fig 14 Clustering assignments of all individuals of *Amanita sp.*. Each vertical line represents an individual, and colors represent the different clusters. The amount of color represents the proportion of association to a cluster.

The *Amanita ponderosa* structure analysis was performed with 153 individuals, removing the ones with low coverage (high missing data). The most likely number of clusters (K) when likelyhood is maximized in *Amanita ponderosa* show three distinct clusters (K=3) (Fig. 15). It is easily identified the cluster with the color green corresponding to five individuals belonguing to Valência de alcântara group. They are the same that were presented in the PCoA of the *Amanita ponderosa.* Concernig the rest of the individuals they do not show any pattern or structure having individuals form the same “populations” in both clusters.

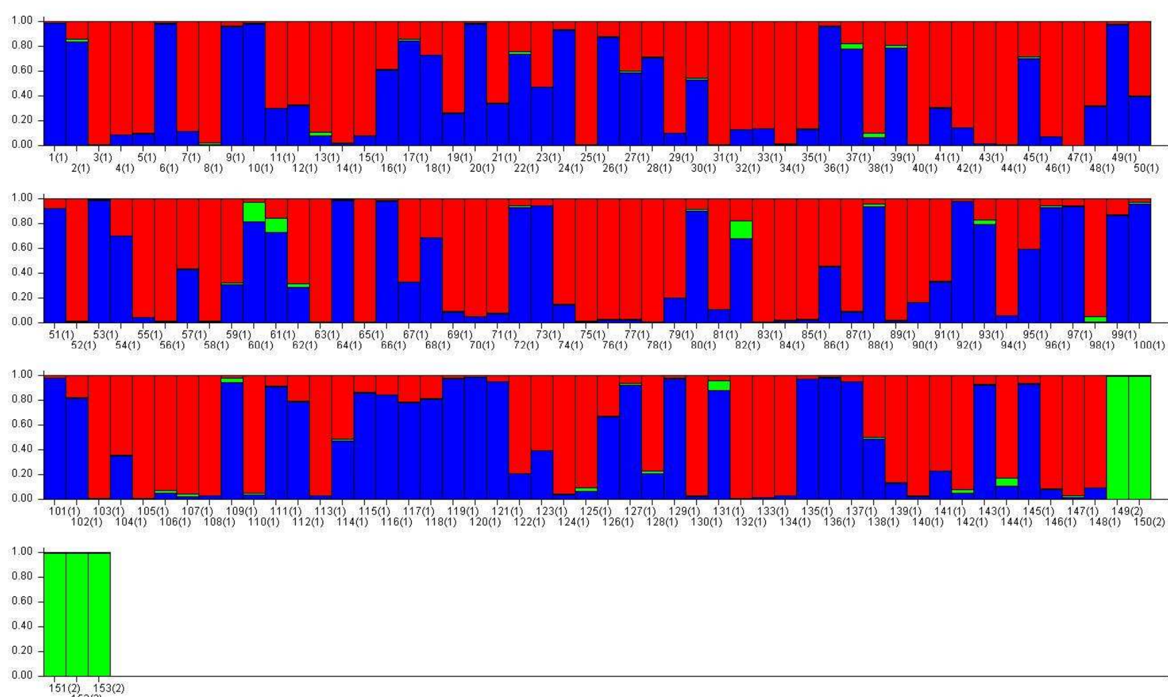


Fig 15 Clustering assignments of all individuals of *Amanita ponderosa.* Each vertical line represents an individual, and colors represent the different clusters. The amount of color represents the proportion of association to a cluster. The individual's correspondence is present in supplementary materials Table S1.

2.5 Discussion

The interesting results obtained in the analysis of rDNA in fact show that there is still much to deepen concerning the *Lepiotoides* species complex. In addition to *Amanita lepiotoides, Amanita ponderosa* and *Amanita curtipes* Neville & Poumarat (2004) make reference to other two taxa intrinsically connected to *Amanita curtipes.* They are *Amanita curtipes f. curtipes* and *Amanita curtipes f. pseudovalens. Amanita curitpes f. curtipes* is illustrated as a typical form of *Amanita curtipes* with slight differences of the *pseudovalens* form. The differences between them are in terms of morphological dimensions in stipe and pilleus diameter and shape, pilleus coloration and smell as we can see in Fig 13. *Amanita curtipes f. pseudovalens* is reported with a pilleus and stipe more plumb and robust than the *f. curtipes.* The absence of brownish tonalities and washed out looks also link to *f. pseudovalens.* It will be important to deepen the taxonomic characterization of these individuals, with special attention to microscopic characters as spore and basidia sizes as well as the pileipellis (hypha organization from the mushroom cap surface) structures. Neville & Poumarat (2004) observed a distinct cutis (horizontal hypha organization with gelatinized tissue) in *f. curtipes* that is not present in *f. pseudovalens*, the latter presenting an ixocutis (horizontal hypha organization without gelatinized tissue).

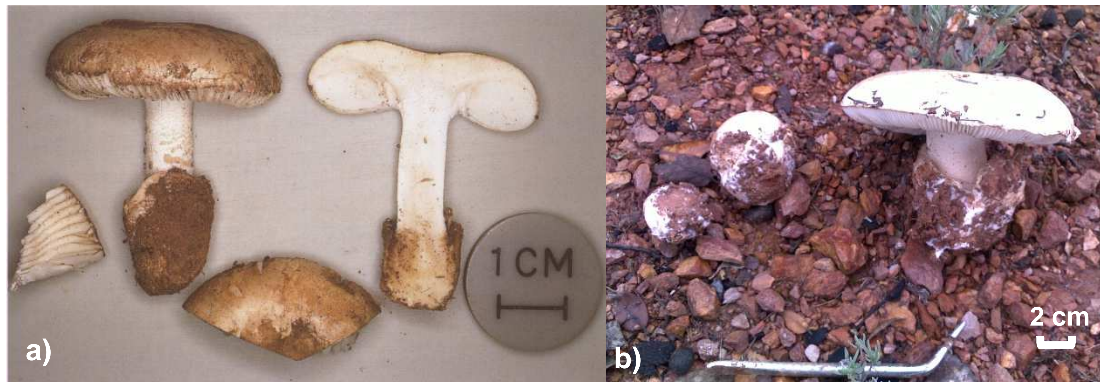


Fig 16 *Amanita curtipes* a) *f. curtipes* and b) *f. pseudovalens* with notorious morphological differentiation.

Besides that, the strong evidence of two very distinct morphotypes and genotypes, could directly reveal the presence of a cryptic species within the complex lepiotoides. For this it is proposed for that new specie a designation of *Amanita pseudovalens*, regarding the previous macroscopic and microscopic characterization, the divergence to *Amanita ponderosa* and the affinity to *Amanita curtipes* demonstrated by the ITS and SNP analysis. The opportunity to realize this study with all of these distinct species may have been a matter of chance. The factor that can major contribute for the occurrence of this species in the data set could be the seasonality. All the samples were collected between the months of February to April being the ones that shown the *Amanita pseudovalens* genotype collected in the beginning of the field work (February). None was presented in the "populations" sampled during the rest sampling campaign. This could reveal a seasonal fruiting pattern of this species complex. Thus the *Amanita pseudovalens* tends to fruiting in the colder months anticipating spring wile *Amanita ponderosa* tends to be more termophilic fruiting in the warmer months during the spring wet season. This pattern was already empirical learned by the local harvesters that traditionally consume that species. There are reports that this species occurring in the first spring months (locals call this *Janeirinhas*, from January) has a different taste but yet it is edible and is consumed. The results that come out from the PCoA and STRUCTURE shown a potential population structure but the data collected could be insufficient to infer the populations structure of this species. Besides that even with just a few individuals of *Amanita pseudovalens* in the dataset it is observed that Mértola (Ap_mrtl*) group all together, results also shown by STRUCTURE. It will be necessary to increase the number of individuals to validate this cluster.

The PCoA in the *Amanita ponderosa* show a clearly distinct out-group composed only by five individuals from Valência de Alcántara. This could raise the possibility of some population structure but a more detailed analysis of the STRUCTURE output results show that are unrelated individuals from the same "population" in other clusters. In fact, the individuals that are present in the green cluster don't have any kind of admixture with

other clusters, which can evidence significant differences between this groups. The most likely occurrence that can explain this differentiation could be a presence of a distinct form of *Amanita ponderosa.* Regarding the SNP analysis even this cluster totally distinct from the others, it seems to be more closely related with *Amanita ponderosa* than *Amanita curipes* or even *Amanita pseudovalens.* One of the two forms proposed by Neville & Poumarat (2004) to *Amanita ponderosa* seems to fit perfectly in the morphotype of these individuals, the *Amanita ponderosa f. valens.* Having the same ecological restrictions and sharing the same habitat the only macroscopic difference between them is the size, being the *f. valens* much more small than the *f. ponderosa.* Some individuals out of this cluster seems to have some proportions of signature with the cluster of the *Amanita ponderosa f. valens.* This could evidence the closer relation between these two forms enabling the cross between them.

The absence of a pattern of population structure could be related with high spore dispersal rate between all populations sampled. This seems to be the most contributive factor for gene flow in this organisms as Vincenot *et al.* (2012) shown in this survey in *Laccaria amenthystina* gene flow across Eurasian.

2.6 References

Baird NA, Etter PD, Tressa, SA, Currey MC, Shiver AL, Zachary AL, Eric US, Cresko, WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inference of population history. *Trends Ecol Evol* **18**, 249–256.

Daza A, Camacho M, Romero de la Osa L, Manjón JL, Moreno G, Santamaría C. (2007). Distribución espacial de la frutificación del hongo ECM comestible Amanita ponderosa Malençon & R. Heim durante seis años consecutivos en un encinar adehesado de la Sierra de Aracena (Huelva). *Investigación Agraria: Sistemas Y Recursos Florestales* **16**: 89-94

Earl DA, vonHoldt, BM (2012) Structure harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Cons. Genetic Res.* **4**, 359-361.

Gardes M, Bruns TD. (1993). ITS primers with enhanced specifity for Basidiomycetes: application to identification of mycorrhizae and rusts. *Mol Ecol* **2**:113-118.

Helyar, SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Nielsen, EE (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular ecology resources*, **11** Suppl 1, 123–36.

Henriques JLG. (2010) Produção de tortulhos *(Amanita ponderosa)* na área do Parque Natural do Tejo Internacional. Avaliação da capacidade produtiva do campo de Cabeço de Mouro (Rosmaninhal – Idanha a Nova). DRAPC, 68 pág

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson E, Cresko WA (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS genetics*, **6**(2),

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular ecology resources*, **11** Suppl 1, 117–22.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, DePristo, M. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, **20**(9), 1297–303.

Miller MR, Brunelli JP, Wheeler PA, Liu S, Rexroad CE, Palti Y, Thorgaard GH (2012). A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology*, ***21***(2), 237–249.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res* **17**: 240–248.

Moreno G, Platas G, Peláez F, Bernedo M, Vargas A, Daza A, Santamaría C (2007) Molecular phylogenetic analysis shows that Amanita ponderosa and A. curtipes are distinct species. *Mycological Progress*, **7**(1), 41–47.

Nekrutenko A, Taylor J (2012). Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nature reviews. Genetics*, **13**(9), 667–72.

Neville P, Poumarat S (2004). Amaniteae. 1. *Fungi Europaei* **9**: 1-1119

Nosil P, Buerkle A (2010) Population Genomics. *Nature Education Knowledge* **3**(10):8

Peakall R, Smouse P (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288-295

Pinho-Almeida F (1994) Estudos taxonómicos do género Amanita. Secção Amidella – complexo Lepiotoides: A. curtipes Gilbert, A. lepiotoides Barla e A. ponderosa Malençon & Heim. *Revista Biol.* (Lisboa) **15**: 131-151

Pringle A, Adams RI, Cross HB, Bruns TD (2009). The ectomycorrhizal fungus Amanita phalloides was introduced and is expanding its range on the west coast of North America. *Molecular ecology*, **18**(5), 817–33.

Pritchard JK, Stephens M, Donelly P (2000) Inference of population structure using multicolus genotype data. *Genetics* **155**, 945-959.

Puritz JB, Addison JA, Toonen RJ (2012). Next-generation phylogeography: a targeted approach for multilocus sequencing of non-model organisms. *PloS one*, **7**(3)

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Method. *Molecular Biology and Evolution* **28**: 2731-2739

Taylor P, Bagley SJ, Orlovich DA, Zealand N, Orlovich DA (2012) Genet size and distribution of Amanita muscaria in a suburban park , Dunedin , New Zealand. *New Zealand Journal of Botany* (September), 37–41

Vincenot L, Nara K, Sthultz C, Labbé J, Dubois MP, Tedersoo L, Selosse MA 2012 Extensive gene flow over Europe and possible speciation over Eurasia in the ectomycorrhizal basidiomycete Laccaria amethystina complex. *Molecular ecology*, **21**(2), 281–99.

Xu J (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Mol Ecol* **15**, 1713–1731.

Xu J, Guo H, Yang Z (2007) Single nucleotide polymorphisms in the ectomycorrhizal mushroom Tricholoma matsutake, *Microbiology*, **153**: 2002–2012.

Table S1 Individual correspondence for *Amanita ponderosa* Structure Plot

Plot Id	Individual	Plot Id	Individual	Plot id	Individual	Plot Id	Individual
1	Ap_arc4B	41	Ap_tlyd5B	81	Ap_arc15A	121	Ap_mlpc2A
2	Ap_sbe3B	42	Ap_cbg1	82	Ap_mlpc12A	122	Ap_mlpc9A
3	Ap_cco4B	43	Ap_cbg10	83	Ap_pan13A	123	Ap_tmr7B
4	Ap_mrtl3A	44	Ap_pan2B	84	Ap_pan8B	124	Ap_mntg3A
5	Ap_tlyd1B	45	Ap_arc1B	85	Ap_mlpc5A	125	Ap_Jaen
6	Ap_tlyd9A	46	Ap_arc9B	86	Ap_tmr3B	126	Ap_pan22A
7	Ap_cbg6	47	Ap_cco1A	87	Ap_vnf2A	127	Ap_mrtl13A
8	Ap_cntd4	48	Ap_cco10B	88	Ap_ita1	128	Ap_mrtl29A
9	Ap_vda2B	49	Ap_mrtl8A	89	Ap_mntg6A	129	Ap_pan18A
10	Ap_arc5B	50	Ap_tlyd6B	90	Ap_red3A	130	Ap_mlpc3A
11	Ap_sbe4B	51	Ap_cbg2	91	Ap_arc16A	131	Ap_mlpc10A
12	Ap_cco5B	52	Ap_cntd1B	92	Ap_mlpc13A	132	Ap_tmr8B
13	Ap_mrtl4	53	Ap_cntd9B	93	Ap_mrtl17A	133	Ap_mntg4A
14	Ap_tlyd2B	54	Ap_vda7B	94	Ap_mrtl25A	134	Ap_red0A
15	Ap_tlyd10B	55	Ap_pan3A	95	Ap_pan14A	135	Ap_mrtl14A
16	Ap_cbg7	56	Ap_arc2B	96	Ap_pan9B	136	Ap_mrtl30A
17	Ap_cntd5B	57	Ap_arc10B	97	Ap_mlpc6A	137	Ap_pan19A
18	Ap_prl4B	58	Ap_cco2A	98	Ap_tmr4B	138	Ap_mlpc4A
19	Ap_arc6B	59	Ap_mrtl1A	99	Ap_vnf3A	139	Ap_tmr1B
20	Ap_sbe5B	60	Ap_mrtl9A	100	Ap_ita2	140	Ap_tmr9A
21	Ap_cco6B	61	Ap_tlyd7B	101	Ap_mntg7A	141	Ap_vnf9A
22	Ap_mrtl5B	62	Ap_cbg4	102	Ap_tlyd13B	142	Ap_mntg5A
23	Ap_tlyd3B	63	Ap_cntd2	103	Ap_mrtl26A	143	Ap_arc14
24	Ap_tlyd11B	64	Ap_cntd10B	104	Ap_pan15A	144	Ap_mlpc11B
25	Ap_cbg8	65	Ap_pan4A	105	Ap_pan10	145	Ap_mrtl15A
26	Ap_cntd6B	66	Ap_arc3B	106	Ap_mlpc7A	146	Ap_mrtl23A
27	Ap_prl12B	67	Ap_cco3B	107	Ap_tmr5A	147	Ap_pan11A
28	Ap_arc7B	68	Ap_mrtl2A	108	Ap_vnf4A	148	Ap_pan20A
29	Ap_cco7B	69	Ap_mrtl10B	109	Ap_ita3	149	Ap_vda3B
30	Ap_mrtl6B	70	Ap_cbg5	110	Ap_huel1A	150	Ap_vda4B
31	Ap_tlyd4B	71	Ap_vda1B	111	Ap_mrtl11A	151	Ap_vda5B
32	Ap_tlyd12B	72	Ap_vda9B	112	Ap_mrtl27A	152	Ap_vda6B
33	Ap_cbg9	73	Ap_pan5	113	Ap_mlpc1A	153	Ap_vda8B
34	Ap_cntd7B	74	Ap_pan6B	114	Ap_tmr6A		
35	Ap_S1	75	Ap_pan7B	115	Ap_vnf5A		
36	Ap_prl24A	76	Ap_tmr2A	116	Ap_mntg1A		
37	Ap_arc8B	77	Ap_vnf1A	117	Ap_huel2A		
38	Ap_sbe8B	78	Ap_vnf10A	118	Ap_tlyd15A		
39	Ap_cco8A	79	Ap_pan21A	119	Ap_mrtl12A		
40	Ap_mrtl7B	80	Ap_red2B	120	Ap_mrtl28A		

3. Manuscript II

Assessing distribution patterns of *Amanita ponderosa* in Iberian Peninsula

3.1 Abstract

Environmental factors constrain the distribution of species and their interactions with the environment. Determining the most influencing factors is important to understand their range limitations. Here, a Species Distribution Model is produced following the maximum entropy approach in order to determine the current distribution of *Amanita ponderosa*, one of the most emblematic fungal species in Iberian Peninsula. A total of nine predictive variables were used (six climatic, two topographic and one land uses) and 201 species presence records. The most important variables conditioning the distribution of *Amanita ponderosa* were the precipitation, temperature and land use. The distribution model predicted a higher occurrence of *A. ponderosa* mostly in the southwest corner of Iberian Peninsula with some occasional occurrences in north and northeast. This study contributes to a better understanding of the ecology of a high valuable mushroom species, which can be integrated in a resource management plan. Nevertheless model can be improved with the increase of the sampling records and also with the integration of more predictive variables.

Keywords

Amanita ponderosa, Maxent, Species Distribution Model, Ibeiran Peninsula

3.2 Introduction

Predicting species distributions has become an important component of conservation planning in recent years, and a wide variety of modelling techniques have been developed for this purpose (Guisan & Thuiller 2005). These are called Species distribution models (SDM) and commonly combine associations between environmental variables and known species occurrence/absence records to identify for instance, suitable habitat and environmental conditions (Elith & Leathwick 2007), the effects of climate changes on species distribution, the biogeographic relationships between other species, the presence of hotspot areas and high diversity areas (Guisan & Zimmermann 2000; Peterson 2006; Sillero *et al.* 2009). The spatial distribution of environments that are suitable for the species can then be estimated across a study region (Pearson 2007). This approach has proven to be valuable for generating biogeographical information that

can be applied across a broad range of fields, including conservation biology, ecology and evolutionary biology.

SDM's have been widely used in a wide range of taxonomic groups such as animals (Ferreira *et al.* 2013; Brito *et al.* 1999) and plants (Elith 2002), but still rarely used in fungi. In fact there are just a few studies regarding fungi spatial distribution and requirements (e.g. Bendiksen *et al.* 2004; Mathiassen & Økland 2007). More recently Wollan *et al.* (2008) tested if the temperature and related parameters were the primary determinants of the regional distribution of macrofungi in Norway. In general the lack of ecological biogeographical studies in macrofungi is probably caused mainly by practical difficulties related to their largely cryptic life cycles. Their fruiting bodies are normally ephemeral and thus their detection and collection are difficult. Also, a correct identification relies mostly on microscopic analyses. The maximum entropy model (Maxent; Phillips *et al.* 2006) show promising results in predicting the distribution of species based in presence-only data (Elith *et al.* 2006). Because of that, this method is particular beneficial for fungi since that the obtaining of reliable absence data is largely impracticable due to their largely cryptic life cycles.

Amanita ponderosa has one of the most cryptic life cycle due to its semi-hypogeous initial development, therefore very difficult to detect. Beside this it is one of the most harvest edible mushrooms in Iberian Peninsula. According to Curreli (1994) and Daza *et al.* (2002) the distribution area of *A. ponderosa* in peninsula is restricted to southwest corner, with some occasional and very restricted occurrences in others places with the same bioclimatic conditions. *A. ponderosa* is therefore classified as thermophilic due to their tendency to grow in areas with hot and dry climate, with an average annual precipitation of 500 mm concentrated during the cold rainy season. It withstands large diurnal and annual temperature ranges, with cold winters and hot summers (Santos-Silva *et al.* 2011). It is now generally agreed that *A. ponderosa* as Ectomicorryzal species is also restricted to their host geographical distribution in the Continental Mediterranean climate parameters characterized by temperature ranges very well defined, hot summers and severe winters with some precipitation (Costa *et al.* 1998). Despite the potential regional distribution of this species is generally known, have become imperative specify the current distribution in order to better understand the relation between species and the environment. This is particularly relevant in a species that is subjected to a severe harvest pressure and without any management measures aiming species conservation.

According with this scenario the following study was conducted towards to i) better understand which climatic variables contribute more for *A. ponderosa* distribution; ii)

define the environmental limits in which this species occurs; and iii) determine the current distribution of this species in the Iberian Peninsula.

3.3 Materials and methods

3.3.1 Study area and sampling

The study was conducted in south-west corner of Iberian Peninsula according to local harvest records. Sampling regions were Baixo-Alentejo, Alto Alentejo, Ribatejo and Beira Baixa in Portugal and Extremadura and Andaluzia in Spain. A total of 201 records were used where 136 were collected in season of 2013 and 65 were from personal collection 2009 to 2012 (Table 2). All the records were obtained with a Global Positioning System (Garmin Oregon 450), represented on the WGS84 grid in a georeferenced database and displayed using ArcMap 9.3 GIS (ESRI 2009).

Table 2 Sampling locations used to predict species distribution model

Location	Cod	*n*	Host	Land cover	Soil type
Aracena	Ap_arc	3	*Q. ilex*	*C.ladanifer*	Schists
Cabeça Gorda	Ap_cbg	16	*Q. suber*	*C.salvifolius*	Schists
Couço	Ap_cco	21	*Q. suber*	*C.ladanifer*	Sandy
Moura (contenda)	Ap_cntd	18	*Q. ilex & P. pinea*	*C.ladanifer*	Schists
Malpica	Ap_mlpc	13	*Q. suber*	*C.ladanifer*	Sandy
Montargil	Ap_mntg	1	*Q. suber*	*C.ladanifer*	-
Odemira	Ap_odm	1	*Q. suber*	-	Schists
Proença-a-Nova	Ap_pan	1	*Q. ilex*	*C.ladanifer*	Granitic
Portel	Ap_prl	48	*Q. suber*	*C.ladanifer*	Schists
Redondo	Ap_red	3	*Q. ilex*	*C.ladanifer*	Schists
São Barnabé	Ap_sbe	28	*Q. suber*	*C.ladanifer*	Schists
Talayueda	Ap_tlyd	21	*Q. ilex*	-	Sandy
Valencia d'Alcántara	Ap_vda	7	*Q. ilex*	*C.ladanifer*	Schists
Vilanueva del Fresno	Ap_vnf	15	*Q. ilex*	*C.ladanifer*	Schists
Tomar	Ap_tmr	2	*P. pinea & Eucalyptus*	*C.ladanifer & salvifolius*	Schists
Huelva	Ap_huel	2	-	-	-
Jaen	Ap_jaen	1	-	-	-
	Total:	201			

3.3.2 Environmental factors

Because of the absence of scientific information about factors related to *A. ponderosa* (and Fungi in general) species range, distribution types and ecological preferences, a wide array of possible predictive variables (PV), 36 in total (Table S1), were selected for modelling potential species distribution (Santos-Silva *et al*. 2011; Daza *et al*. 2002 & 2007; Moreno *et al*. 2007; Henriques 2010; Wollan *et al*. 2008). Climatic variables with a

resolution of 30 arc-seconds (~1 km spatial resolution) were obtained from the WorldClim data base in the datum WGS84 (Hijmans *et al.* 2005; http://www.worldclim.org/). A Digital elevation model (DEM) was obtained from Shuttle Radar Topography Mission, SRTM (SRTM 2010) and used to derive others topological parameters as slope, aspect and solar direct radiation. Land uses information was obtained from Corine Land Cover, v.2006 (EEA 2010). Evapotranspiration values were obtained from International Water Management Institute (IWMI; Zomer *et al.* 2006). All values of PV in each species location were obtained by interception with Xtools (DataEast) in GIS environment ArcMap 9.3. The principle of parsimony applied to statistical modelling recommends a model utilization with k-1 predictive variables instead a model with k variables (Crawley 2007). Thus a correlation test was performed to avoid high correlation values between all potential predictive variables in sampled sites. Pearson correlation coefficient was calculated with SPSS Statistics v.21 (IBM Corp.) between all variables (Snelder & Lamouroux 2010). All correlations with absolute values above 0.8 was excluded (cf. Snelder & Lamouroux 2010)

3.3.3 Species distribution model

Species distribution model for current conditions was generated using Maxent software version 3.3.3k available at http://www.cs.princeton.edu/~schapire/maxent/. This software calculates the realised niche of species (Sillero 2010) by using the environmental data from the only-presence records and the background sample (finite number of points from the landscape to which values of climatic variables are associated) to estimate the ratio "probability density of covariates across locations within landscape where the species is present"/"probability density of covariates across landscape" (Elith *et al.* 2011). This is done by choosing the statistical model with the maximum entropy (Phillips *et al.* 2004; Phillips *et al.* 2006; Phillips 2012). Pixels with no presence records are not treated as absences. Maxent makes a post-transformation of the raw output providing a logistic output that makes assumptions about prevalence and sampling effort to achieve a better estimate of the suitability of the environment for a species presence rather than the probability of occurrence (Phillips 2012). The final model was the average of 20 replicates which were run with random seed (i.e a different random train was made and a different random subset of the background points was used for each run). A random test percentage of 30% was chosen so that the program sets aside 30% of the sample records for testing (n=60) and uses 70% as training (n=141) data set on each run. The test data set is used to evaluate the performance of the model. The sampling technique employed was bootstrap. This is a Monte Carlo statistical method that takes sampling with replacement from the original sample, hence being more accurate than sampling without replacement in terms of simulating chance as it samples the impacts

of the real sample size (Wintle *et al.* 2005). Each sample is created independently from the other samples and the number of samples is equal to the total number of presence points. Model was run with auto features and recommended default values and options were used (Phillips *et al.* 2006). The output was in logistic format which gives an estimate of suitability between 0 and 1.

Model was tested with the area under the curve (AUC) of the receiver operating characteristics (ROC) plot as it is a measure of individual model fit (Fielding & Bell 1997). A ROC plot is produced by relating the proportion of presences correctly predicted (sensitivity) with the proportion of pseudo-absences incorrectly predicted (1-specificity). Higher values of AUC may translate the proportion between the total size of the distribution area of the species and the size of the study area: the larger the proportion the larger the value of AUC (Lobo *et al.* 2008). The relevance of each variable to the Maxent prediction was determined by Jackknife analyses of the training and test gain and of AUC. In Jackknife analysis each PV is excluded in turn and a model is created with the remaining variables; then another model is created using each variable in isolation; and finally, a model is created using all variables. The value of the gain for each PV indicates if the variable contributes more in predicting the suitability than a uniform distribution (which would have zero gain) (Miller 1974; Pearson *et al.* 2006). Response curves of each PV were performed to also assess the importance of each PV in predicting the suitability of habitat. The values of the 20 replicates of each PV were averaged and represented by a tendency curve adjusted by a polynomial equation of sixth degree. These curves give the probability of contribution of PV for the raw prediction of the models which indicate the values of the PV suitable for each species (following Martínez-Freiría *et al.* 2008).

3.4 Results

From 36 PV previously selected, 9 were included (Pearson<0.8) (Table 3) to predict species distribution in Iberian Peninsula. The variable with the higher contribution for model prediction was the Land Cover while Aspect was the one that less contributes. All the percentage contributions of the PV are expressed in Table 4. Together Land Cover and Mean Temperature contribute to more than a half of the prediction. The results of Jackknife analyses corroborate these findings and were concordant (Fig 14). Jackknife analyses shown that removing the Land Cover (clc) from the model it will decrease the training and test gain (Fig 14 a) and b)) which indicates that clc contain the most contributive information that is not present in other variables. The proportion of

presences correctly predicted by the model show promising results with the AUC values closer to 1, and the Training and Test AUC values being of 0.991 and 0.988 respectively.

Table 3 Predictive variables selected (P<0.8) and used to model species distribution.

Variable	cod
Aspect	aspect
Land Cover	clc
Elevation	dem
Evapotranspiration	evapo
Annual Mean Temperature	meantemp
Precipitation September	presep
Precipitation wettest Quarter	prewetq
Temperature January	tempjan
Mean temperature wettest quarter	tmeanwetq

Table 4 Contribution of each predictive variables (%) and mean training and test AUC for the 20 Maxent models of *Amanita ponderosa* in Iberian Peninsula. Minimum and maximum values within brackets.

		Average	Min - Max
Predictive variables	aspect	1.354	(0.485 - 3.306)
	clc	30.190	(26.390 - 33.922)
	dem	11.376	(3.561 - 21.268)
	evapo	9.044	(4.499 - 15.592)
	meantemp	22.573	(6.88 - 29.129)
	presep	4.313	(0.961 - 11.843)
	prewetq	2.103	(0.741 - 4.498)
	tempjan	4.933	(0.320 - 15.337)
	tmeanwetq	14.110	(9.688 - 18.727)
Test	Training AUC	0.991	(0.988 - 0.994)
	Test AUC	0.988	(0.981 - 0.995)
	Test Gain	3.826	(3.459 - 4.173)

Response curves (Fig 15) generated by the model show that *A. ponderosa* preferentially occurs in south exposed areas (i.e aspect values between 100º and 230º) and preferentially at low altitudes, peaking at 100 to 200 m. Concerning the temperature variables, *A. ponderosa* distribution area is mostly characterized by annual mean temperature of 15-17ºC, mean temperature of the wettest quarter 10-12ºC and mean temperature of January 9-10ºC. The precipitation values not discloses a harmonious pattern throughout the gradient and can lead to some difficulties of interpretation. Precipitation of September reaches the maximum response at 25mm corresponding to 65% of probability of presence while the mean precipitation of the wettest quarter shows values between 200 and 300mm corresponding to ~60% probability of presence.

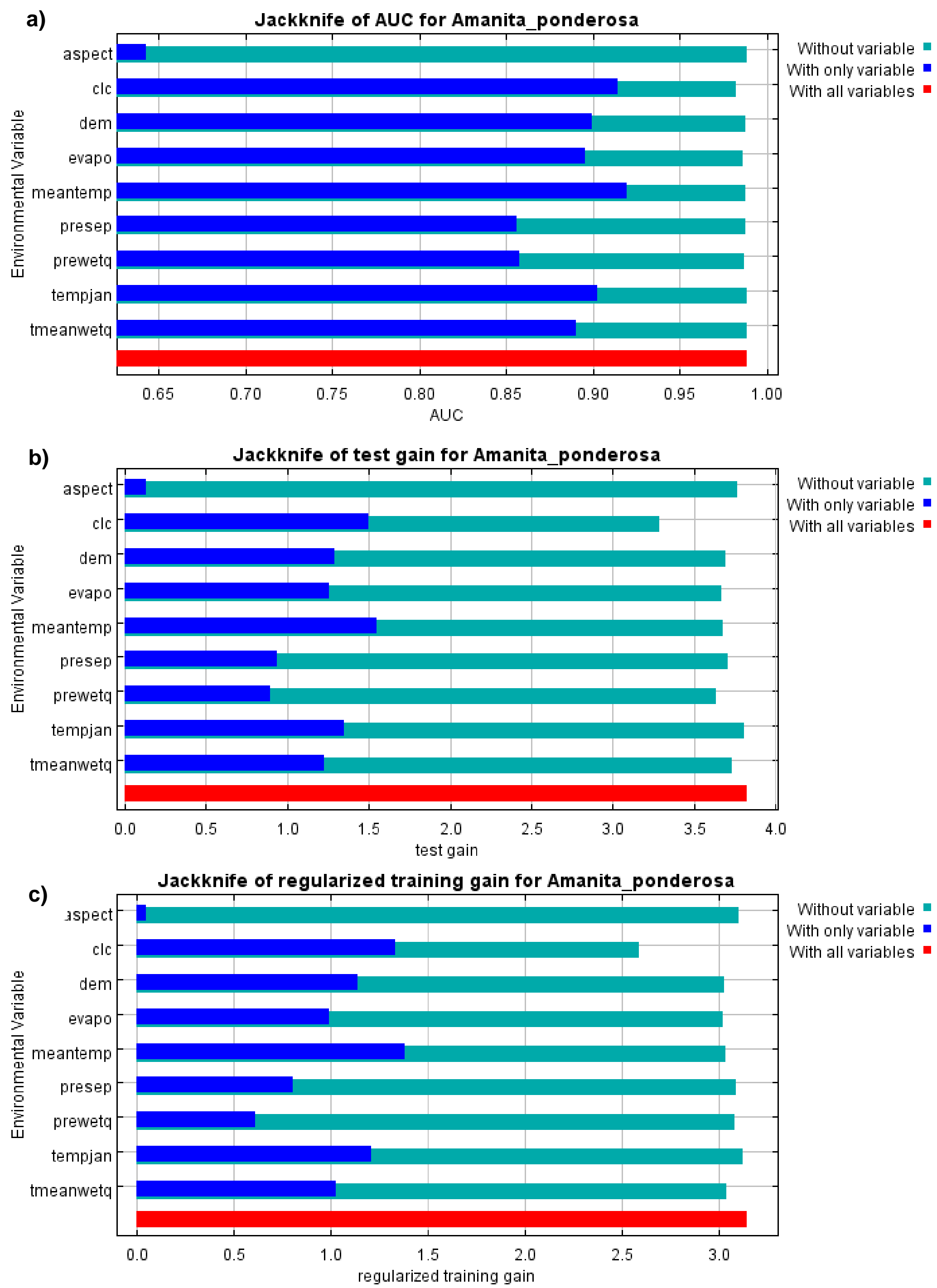


Fig 17 Jackknife of AUC (a) Regularized Test Gain (b) and (c) Training of *Amanita ponderosa* model distribution

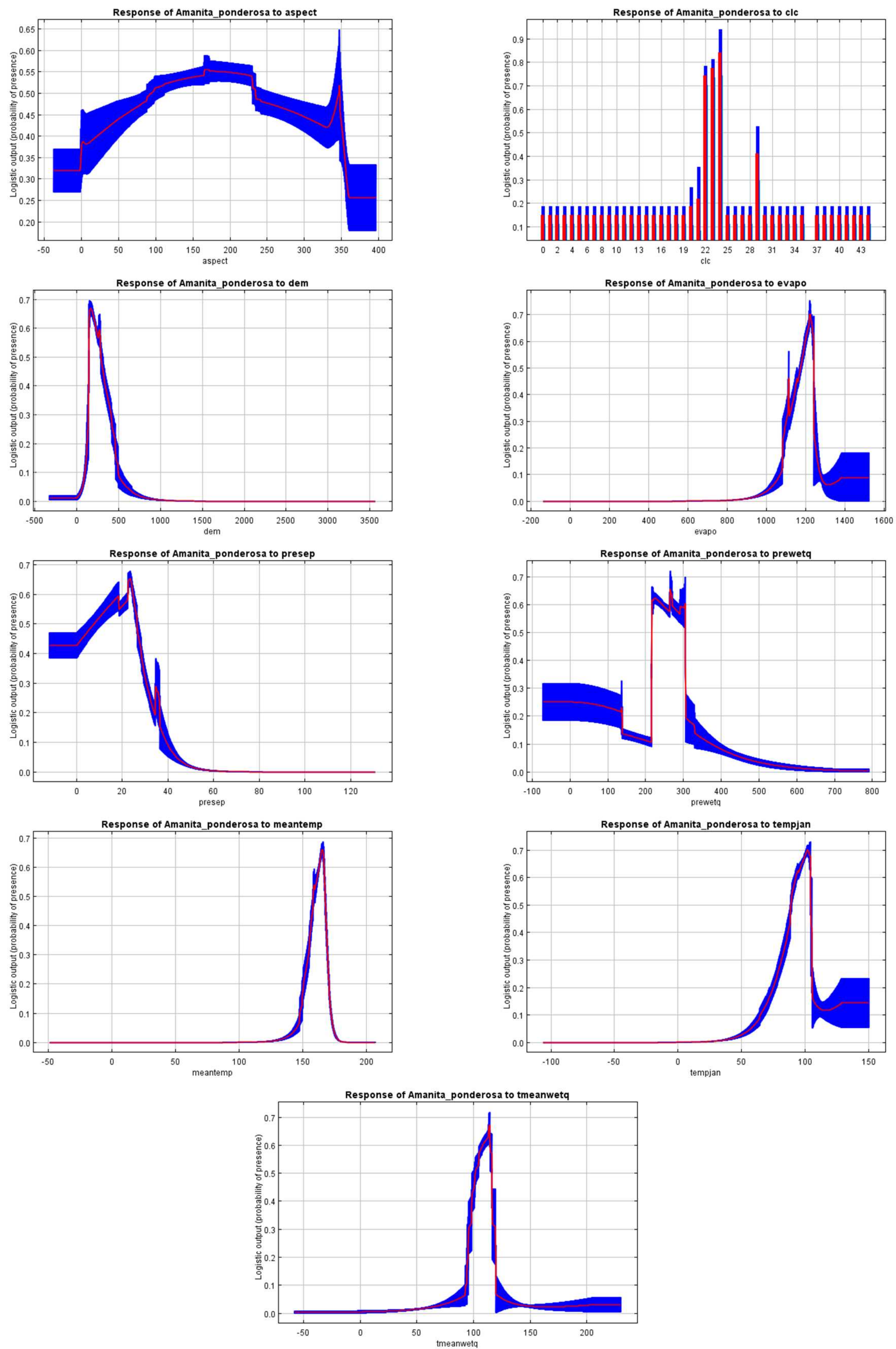


Fig 18 Response curves of each predictive variable to the *Amanita ponderosa* distribution model. The curves show the mean response of the 20 replicate Maxent runs (red) and the mean +/- one standard deviation (blue)

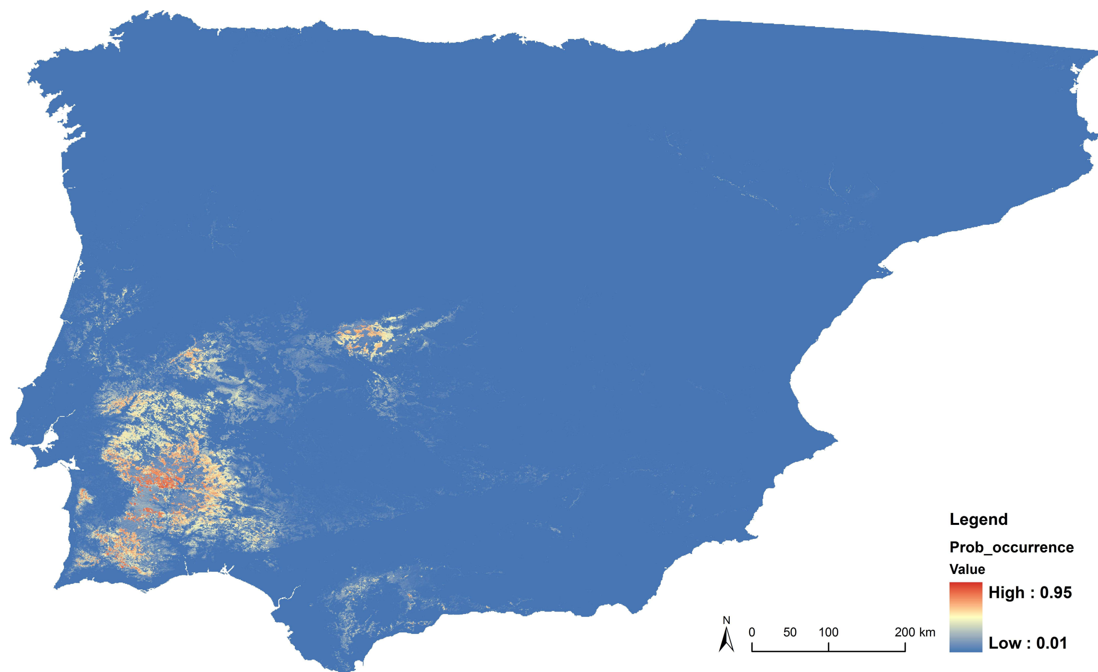


Fig 19 Model prediction of probability of occurrence of *Amanita ponderosa* in Iberian Peninsula.

The response curve corresponding to Land Cover shows the predominant occurrence of this species is in the agro-forestry areas (clc code 22), broad-leaved forest (clc code 23), coniferous forest (clc code 24) and transitional woodland-shrub (clc code 29). Higher values of evapotranspiration (~1200mm) increase the probability of occurrence of *A. ponderosa*.

The mean model prediction evidence clearly the probability of occurrence of *A. ponderosa* in southwest Iberian Peninsula (Fig 16). The areas with higher probability values (>60%) in Portugal were the Algarve mountain range of Caldeirão and Monchique, the Guadiana river valley, the mountain range of Portel, Ossa and Grândula, some areas in Sorraia river basin and finally the region delimited by Campina de Idanha-a-Nova converging with the Spanish San Pedro mountain range. In Spain the higher probability of occurrence is represented throughout the bordering area with Portugal in the regions of Estremadura and Andalusia. In Estremadura is also very evident the high values in the southern part of the Iberian central mountain range corresponding to the Sierra de Gredos. Lower values of probability of occurrence (<60%) are represented in west side of Penibético mountain system corresponding to Natural Park of Sierra de Grazalema and in mountain range of Sierra Morena. The model prediction evidence also some remote areas with probability values inferior to 30%. These areas were the International Douro River Natural Park and the Ebro Valley.

3.5 Discussion

Areas with high probability of *A. ponderosa* occurrence in Iberian Peninsula follow the general distribution patterns previously identified by others authors Curreli (1994); Daza *et al.* (2002) and Moreno *et al.* (2007). Nevertheless, this model allowed the definition of accurate suitable habitats for species that were previously considered widespread from all southwest peninsula. The model was developed considering a restricted range environmental conditions in relation to the total distribution of *A. ponderosa*, which may have induced spatial biases in estimates of occurrence probability. It is therefore necessary increase the sampling area to better estimate the current species distribution. Nevertheless the model prediction and the value of AUC (0.99) shows a very good fit for the model. The results obtained in this work corroborate the previous studies concluding that distribution patterns of fungal species are closely related to Temperature and Precipitation (Wollan *et al.* 2008). Ulf *et al.* (2012) also linked climate variability to mushroom productivity and phenology, showing the importance of climate not only in distribution but also in abundance. Their results shown that precipitation amounts and temperature means determined fungal activity. *A. ponderosa* distribution scenario in Iberian Peninsula revealed the importance of Mean Temperature and Mean Temperature of the wettest quarter in terms of most important climatic variables. Despite some variables have been rejected in the model in consequence of the correlation tests, we can infer, through the detailed analysis of the climatic parameters in the Mediterranean Temperate Climate, that areas with higher mean temperatures are also areas with low mean precipitation rates (Baldi *et al.* 2013). The topography of the southwest corner of the Iberian Peninsula is also a very important parameter regulating the majority of the climatic variables. Due to this it is known that the increase of the precipitation values and the decrease of temperature are highly positive correlated with the elevation (Fig S1). This assumption clearly determines the thermophilic pattern of this species (Santos-Silva *et al.* 2011) occurring only in areas with higher annual mean temperatures and low precipitation rates.

The character of highly specification of this species forming mycorrhizal associations with a restrict number of plant species was printed by the influence of the Land Cover (clc) in model prediction. The four most selected categories by the model (Agro-forestry areas (22), Broad-leaved forest (23), Coniferous forest (24) and Transitional woodland-shrub (29)), fits very well on the species distribution with all previous observations. Agro-forestry areas (22) are characterized by annual crops or grazing land under the wooded cover of forestry species matching with typical Montado ecosystems as well as Broad-leaved forest (23) (EEA 2010). Montado is an artificial system highly intervened by human activities which sometimes due to intensification and overexploitation presents

substantial losses in soil quality. Despite Coniferous forest (24) having expressed high values in model contribution, it is unlikely that *A. ponderosa* was associated to this host. All the coniferous forest surveyed for this study were dominated by Mediterranean shrubs like the one belonging to *Cistaceae* family and this is reported by Parra & Domínguez (2012) as the most likely host species. The *Cistaceae* are also the most likely shrub in the Transitional woodland-shrub (29) what makes this habitat also suitable for the occurrence of this species.

The model predictions with lower probability of occurrence, like in areas like Natural Park of International Douro River and Ebro Valley, should require more attention in future works. It is known that *A. ponderosa* is present on these areas, although in small isolated islands (Personal records). Due to its small size, these islands cannot be predicted in this model. It is very important collect more information about these areas in order to better understand which variables are promoting the species occurrence. Besides this the maximum entropy model revealed to be a powerful method to evaluate the current distribution of this species skirting the difficulties related with sampling species with cryptic life cycles.

Hereafter, in addiction of the increment of the sampling area (described above) other issues should be taken into account in future predictions as the relation fungus-host, regarding the distribution of potential *A. ponderosa* host. It is crucial add geological information (as soil types and composition) and similarly derived as chemical composition to better understand the distribution patterns of this species.

3.6 References

Baldi M, Dalu D, Bartolini G, Grifoni D, Maracchi G, Gozzini B, Dalu G, 2013 Seasonal precipitations in the Central Mediterranean basin and large scale climate patterns, *Geophysical Research Abstracts* **15**

Bendiksen E, Økland RH, Høiland K, Eilertsen O, Bakkestuen V (2004) Relationships between macrofungi, plants and environmental factors in boreal coniferous forests in the Solhomfjell area, Gjerstad, S. Norway. *Sommerfeltia*, **30**, 1–128.

Brito C, Crespo EG, Paulo OS (1999) Modelling wildlife distributions: Logistic Multiple Regression vs Overlap Analysis. *Ecography*, ***22***(3), 251–260.

Büntgen U, Kauserud H, and Egli S, 2012 Linking climate variability to mushroom productivity and phenology. Frontiers in Ecology and the Environment **10**: 14–19

Costa JC, Aguiar C, Capelo J, Lousã M, Neto C, (1998) Biogeografia de Portugal Continental. *Quercetea*. ISSN 0874-5250, p. 5-56

Crawley MJ (2007) The R Book. John Wiley & Sons, Ltd. 950pp.

Curreli S (1994) Amanite rare in Sardegna: Amanita heterocroma S. Curreli nom. nov. e Amanita ponderosa Malençon & Heim. *Micologia Italiana* **3**, 61-65

Daza A, Romero L, Santamaría C, Camacho M, Aguilar A, Mejías M, Moreno G, Pérez M, Manjón JL (2002) El gurumelo, una seta de primavera. *Quercus* **193**, 20-24

Daza A, Camacho M, Romero de la Osa L, Manjón JL, Moreno G, Santamaría C (2007) Distribución espacial de la frutificación del hongo ECM comestible Amanita ponderosa Malençon & R. Heim durante seis años consecutivos en un encinar adehesado de la Sierra de Aracena (Huelva). *Investigación Agraria: Sistemas Y Recursos Florestales* **16**: 89-94

EEA, (2010) European Environmental Agency. Corine Land Cover 2000. In: http://www.eea.europa.eu/themes/landuse/interactive/clc-download. Access: 15-10-2013.

Elith J (2002) Quantitative methods for modeling species habitat: Comparative performance and an application to Australian plants. In S. Ferson and M. Burgman (Eds.), *Quantitative methods for conservation biology*, 39–58

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JMcC, Peterson AT, Phillips SJ, Richardson KS, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz, MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. **29**: 129-151.

Elith J, Leathwick JR (2007) Predicting species' distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines, Diversity and *Distributions*, **13**: 165-175.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**: 43–57

ESRI®, 2009. ArcMap 9.3.1.

Ferreira AF, Quintella BR, Maia C, Mateus CS, Alexandre CM, Capinha C, Almeida PR (2013) Influence of macrohabitat preferences on the distribution of European brook and river lampreys: Implications for conservation and management. *Biological Conservation*, **159**(0), 175–186.

Fielding AH, Bell JF (1997) A Review Of Methods For The Assessment Of Prediction Errors In Conservation Presence/Absence Models. *Environmental Conservation*, **1**(24), 38–49

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**(9), 993–1009.

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**: 147–186.

Henriques JG (2010) Produção de tortulhos (Amanita ponderosa) na área do Parque Natural do Tejo Internacional. Avaliação da capacidade produtiva do campo de Cabeço de Mouro (Rosmaninhal – Idanha a Nova). DRAPC, 68 pág

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978. Retrieved June 14, 2012, from http://www.worldclim.org/

Kühner R & Romagnesi H, (1953) Flore analytique des Champignons supérieurs, 556 pp

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr*, **17**: 145–151.

Martínez-Freiría F, Sillero N, Lizana M, Brito JC (2008) GIS-based niche models identify environmental correlates sustaining a contact zone between three species of European vipers. *Diversity and Distributions*, **14:** 452–461.

Mathiassen GH, Økland RH, (2007) Relative importance of host tree species and environmental gradients for epiphytic species composition, exemplified by pyrenomycetes s. lat. (Ascomycota) on Salix in central north Scandinavia. *Ecography* **30**: 251-263

Miller RG (1974) The Jackknife - A Review*. Biometrika*, **61**: 1–15.

Moreno G, Platas G, Peláez F, Bernedo M, Vargas A, Daza A, Santamaría C (2007) Molecular phylogenetic analysis shows that Amanita ponderosa and A. curtipes are distinct species. *Mycological Progress*, **7**(1), 41–47.

Parra BM, Dominguéz ER (2012) Aportaciones al conocimiento de la micoflora de la Serranía de Ronda (Málaga, España)ii*; Acta Botanica Malacitana* **37**: 167-170

Pearson, R.G., Raxworthy, C. J., Nakamura, M., and Peterson, A. T. (2006). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography, **34:** 102–117

Pearson RG (2007) Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org in 14-08-2013

Peterson AT (2006) Uses and Requirements of Ecological Niche Models and Related Distributional Models. *Biodiversity Informatics*, **3**: 59–72.

Phillips SJ, Dudík M, Schapire RE (2004) A Maximum Entropy Approach to Species Distribution Modeling. In, Proceedings of the Twenty-First *International Conference on Machine Learning*. Banff, Canada, 655–662.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**: 231–259.

Phillips S (2012) A Brief Tutorial on Maxent. Retrieved July 9, 2013, from http:www.cs.princeton.edu

Santos-Silva C, Machado H, Guinberteau J (2011) Ecologie de Amanita ponderosa et gestion des écosystèmes. In Ecologie de Amanita ponderosa et gestion des écosystèmes, 114 - 126. *Micosylva*: Jean Rondet et Fernando Martin

Sillero N (2010) Modelling new suitable areas for Hyla meridionalis in a current and future expansion scenario. *Amphibia-Reptilia*, **1**(31), 37–50.

Sillero N, Brito JC, Skidmore AK, Toxopeus AG (2009) Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. *Amphibia-Reptilia*, **30**: 185–206

Shuttle Radar Topography Mission (SRTM). 2010. US Geological Survey. US Department of the Interior. Retirado de: http://www2.jpl.nasa.gov/srtm/spain.htm access in: 06-06-2013.

Snelder TH, Lamouroux N (2010) Co-variation of fish assemblages, flow regimes and other habitat factors in French rivers. *Freshwater Biology*. **55** (4): 881–892.

Wintle BA, Elith J, Potts JM (2005) Fauna habitat modelling and mapping: A review and case study in the Lower Hunter Central Coast region of NSW. *Austral Ecology*, **30**: 719–73

Wollan AK, Bakkestuen V, Kauserud H, Gulden G, Halvorsen R (2008) Modelling and predicting fungal distribution patterns using herbarium data. *Journal of Biogeography*, **35**(12), 2298–2310

Zomer RJ, Trabucco A, van Straaten O, Bossio DA (2006) Carbon, land and water: A global analysis of the hydrologic dimensions of climate change mitigation through afforestation/reforestation. Colombo, Sri Lanka: *International Water Management Institute*. 44p. (IWMI Research Report 101)

3.6 Supplementary materials

Table S2 Predictive variables selected to modelling *Amanita ponderosa* in Iberian Peninsula.

Variable	Cod	Resolution	Description	Source
Elevation (m)	dem	1x1km	Altitude relative to the mean sea level	SRTM, 2010
Slope (º)	slope	1x1km	Maximum variation of the altitude concerning its 8 neighbouring cells	-
Aspect (º)	aspect	1x1km	Downslope direction of the maximum rate of change in value from each cell to its neighbours	-
Radiation (MJ/m2)	rad	1x1km	Represents the insolation at specific locations.	-
Temperature January (ºC)	tempjan	1x1km		Worldclim
Temperature February (ºC)	tempfeb	1x1km		Worldclim
Temperature March(ºC)	tempmar	1x1km		Worldclim
Temperature April (ºC)	tempapr	1x1km		Worldclim
Temperature May (ºC)	tempmay	1x1km		Worldclim
Temperature June (ºC)	tempjun	1x1km	Average of the maximum temperature registered by month during the period of 1950 to 2000	Worldclim
Temperature July (ºC)	tempjul	1x1km		Worldclim
Temperature August (ºC)	tempaug	1x1km		Worldclim
Temperature September (ºC)	tempsep	1x1km		Worldclim
Temperature October (ºC)	tempoct	1x1km		Worldclim
Temperature November (ºC)	tempnov	1x1km		Worldclim
Temperature December (ºC)	tempdec	1x1km		Worldclim
Annual Mean Temperature(ºC)	meantemp	1x1km	Annual mean temperature registered during the period of 1950 to 2000	Worldclim
Minimum Temperature of coldest month (ºC)	tmincoldm	1x1km	Average of the minimum temperature in the coldest month during the period of 1950 to 2000	Worldclim
Mean Temperature of the Wettest Quarter (ºC)	tmeanwetq	1x1km	Average of the mean temperature in the wettest quarter of the year during the period of 1950 to 2000	Worldclim
Precipitation January (mm)	prejan	1x1km		Worldclim
Precipitation February (mm)	prefeb	1x1km		Worldclim
Precipitation March (mm)	premar	1x1km		Worldclim
Precipitation April (mm)	preapr	1x1km		Worldclim
Precipitation May (mm)	premay	1x1km		Worldclim
Precipitation June (mm)	prejun	1x1km	Average of the maximum precipitation registered by month during the period of 1950 to 2000	Worldclim
Precipitation July (mm)	prejul	1x1km		Worldclim
Precipitation August (mm)	preaug	1x1km		Worldclim
Precipitation September (mm)	presep	1x1km		Worldclim
Precipitation October (mm)	preoct	1x1km		Worldclim
Precipitation November (mm)	prenov	1x1km		Worldclim
Precipitation December (mm)	predec	1x1km		Worldclim
Annual Mean Precipitation (mm)	meanpre	1x1km	Annual mean precipitation registered during the period of 1950 to 2000	Worldclim
Precipitation of the driest month (mm)	predrim	1x1km	Average of the precipitation in the driest month during the period of 1950 to 2000	Worldclim
Mean Precipitation Wettest Quarter (mm)	prewetq	1x1km	Average of the mean precipitation in the wettest quarter of the year during the period of 1950 to 2000	Worldclim
Land Cover	clc	1x1km	Land uses and occupation by natural and non-natural areas.	EEA
Evapotranspiration (mm)	evapo	1x1km	Potential evapotranspiration is the sum of evaporation and plant transpiration from the Earth's land surface to atmosphere.	IWMI (Zomer et al 2006)

Table S3 Land cover adapted from CORINE Land Cover EEA, 2010

Group	Grid Code	CLC code	Sub-group	Type
Artificial surfaces	1	111	Urban fabric	Continuous urban fabric
	2	112	Urban fabric	Discontinuous urban fabric
	3	121	Industrial, commercial and transport units	Industrial or commercial units
	4	122	Industrial, commercial and transport units	Road and rail networks and associated land
	5	123	Industrial, commercial and transport units	Port areas
	6	124	Industrial, commercial and transport units	Airports
	7	131	Mine, dump and construction sites	Mineral extraction sites
	8	132	Mine, dump and construction sites	Dump sites
	9	133	Mine, dump and construction sites	Construction sites
	10	141	Artificial, non-agricultural vegetated areas	Green urban areas
	11	142	Artificial, non-agricultural vegetated areas	Sport and leisure facilities
Agricultural areas	12	211	Arable land	Non-irrigated arable land
	13	212	Arable land	Permanently irrigated land
	14	213	Arable land	Rice fields
	15	221	Permanent crops	Vineyards
	16	222	Permanent crops	Fruit trees and berry plantations
	17	223	Permanent crops	Olive groves
	18	231	Pastures	Pastures
	19	241	Heterogeneous agricultural areas	Annual crops associated with permanent crops
	20	242	Heterogeneous agricultural areas	Complex cultivation patterns
	21	243	Heterogeneous agricultural areas	Land principally occupied by agriculture, with significant areas of natural vegetation
	22	244	Heterogeneous agricultural areas	Agro-forestry areas
Forest and semi natural areas	23	311	Forests	Broad-leaved forest
	24	312	Forests	Coniferous forest
	25	313	Forests	Mixed forest
	26	321	Scrub and/or herbaceous vegetation associations	Natural grasslands
	27	322	Scrub and/or herbaceous vegetation associations	Moors and heathland
	28	323	Scrub and/or herbaceous vegetation associations	Sclerophyllous vegetation
	29	324	Scrub and/or herbaceous vegetation associations	Transitional woodland-shrub
	30	331	Open spaces with little or no vegetation	Beaches, dunes, sands
	31	332	Open spaces with little or no vegetation	Bare rocks
	32	333	Open spaces with little or no vegetation	Sparsely vegetated areas
	33	334	Open spaces with little or no vegetation	Burnt areas
	34	335	Open spaces with little or no vegetation	Glaciers and perpetual snow
Wetlands	35	411	Inland wetlands	Inland marshes
	36	412	Inland wetlands	Peat bogs
	37	421	Maritime wetlands	Salt marshes
	38	422	Maritime wetlands	Salines
	39	423	Maritime wetlands	Intertidal flats
Water bodies	40	511	Inland waters	Water courses
	41	512	Inland waters	Water bodies
	42	521	Marine waters	Coastal lagoons
	43	522	Marine waters	Estuaries
	44	523	Marine waters	Sea and ocean
No Data	48	999	No data	No data
Unclassified	49	990	Unclassified terrestrial surfaces	Unclassified terrestrial surfaces
	50	995	Unclassified Water bodies	Unclassified Water bodies
	255	990	Unclassified	Unclassified

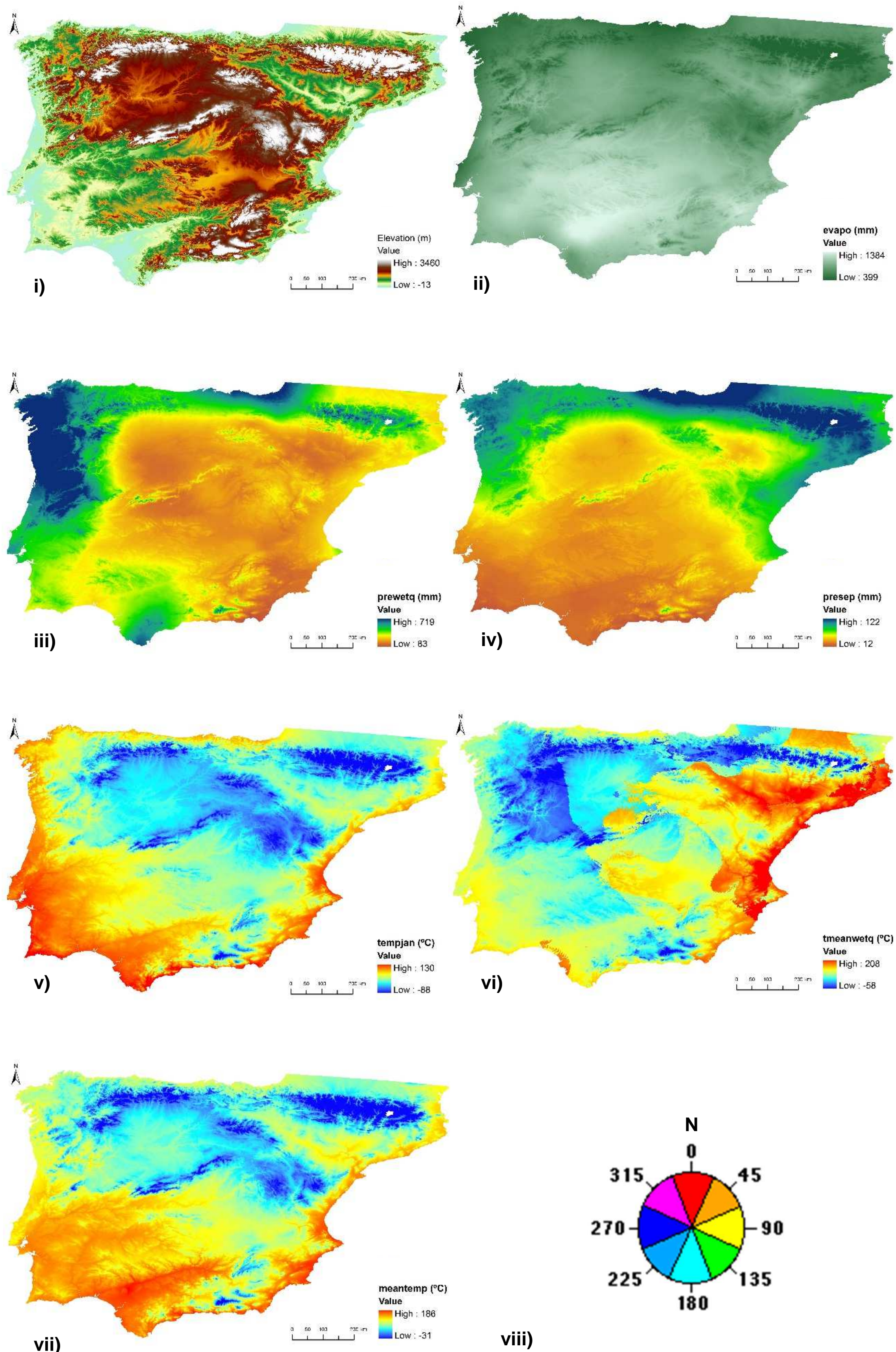


Fig S1 Graphical representation of the predictive variables used in species distribution model. i) Digital elevation model; ii) Evapotranspiration; iii) Precipitation of the wettest quarter; iv) Precipitation of September; v) Temperature of January; vi) Mean temperature of the wettest quarter; vii) Annual mean temperature and viii) Aspect.

4. Final considerations

The results of this study reveal very well some of the difficulties that have arisen nowadays and are intrinsically related with the distinction of some cryptic species complex in the fungi kingdom. Here I will summarise the main conclusions of this work and there relative importance.

The revolution caused by the emergence of genomics in the research field of non-model organisms seems to play nowadays a very important role. The tools that we have now allow us to infer questions than heretofore was completely disabled. Providing a genome "complexity-reduction" system, RADSeq allied to improvement of sequencing techniques makes possible population genomic studies with unprecedented depth and complexity. The number of SNP discovered here allow us to deepen knowledge and cross the barrier of the mysterious world of fungi and learn more about the organization and the relationships between them. Even talking about small genome organisms like fungi, the number of markers seems to be reasonable. These SNP data provide the first genetic evidence for the taxonomic validity of the two endemic species *Amanita ponderosa* and *Amanita pseudovalens* in Iberian Peninsula. At the same time, it was highlighted the power of next generation sequencing technologies to resolve old questions in mycology.

The *Lepiotoides* species complex still being one of the most difficult to interpret, but with the conduction of this study some issues seems to get more sense. The casual harvest of individuals from the difference species here focused was rather contributing to the success of this work. So it is important to note that prior to this study little was known about the relationship between these species and now it was shown that are other species conducting an important role in the local harvest tradition.

This work proved to be a very important tool for the development of a management plan mycological resource, combining genetic and ecological data. This approach seemed to be very interesting reconciling different perspectives and fields of research for better interpretation of the results.

Much work remains to be done to better understand the relationship between these species and their habitat. Besides the sampling for this study propose have been carried out carefully some other issues should be taken into account, as related with micro habitat. It is important to devise a sampling strategy throughout the fruiting period in order to understand better the seasonality of each species. Since this is a highly economically profitable resource is recommended to continue the research to ensure its sustainability.

5. References

Ainsworth GC (1976) Introduction to the History of Mycology. *Cambrigde University Press*. Cambridge, UK.

Alexopoulos CJ, Mims CW, Blackwell M (1996) Introductory Mycology. 4th edition. *John Wiley and Sons*, Inc. New York, USA

Arraiano-Castilho R, Vila-Viçosa C, Mendes P, Morgado L, Pinto-Gomes C (2013) Reconhecimento do Potencial Micológico - Propriedades: Perímetro Florestal Salvada - Cabeça Gorda Casa Velha – Mértola – Relatório Final. Grupo de Trabalho em Micologia do CEAI. (unpublished)

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* **157**: 101-18.

Azul AM, Castro P, Sousa JP, Freitas H (2009) Diversity and fruiting patterns of ectomycorrhizal and saprobic fungi as indicators of land-use severity in managed woodlands dominated by Quercus suber — a case study from southern Portugal. *Canadian Journal of Forest Research*, **39**(12), 2404–2417.

Baird NA, Etter PD, Tressa SA, Currey MC, Shiver AL, Zachary AL, Eric US, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

Bas C (1969) Morphology and subdivision of Amanita and a monograph of its section Lepidella. *Persoonia*, **5**: 285–579.

Begerow D, Nilsson H, Unterseher M, Maier W (2010) Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Applied microbiology and biotechnology*, **87**(1), 99–108.

Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC microbiology*, **10**: 189.

Bon M (2004) Champignons de France et d'Europe Occidentale. Flammarion eds, Paris

Bonfante P, Anca IA (2009) Plants, mycorrhizal fungi, and bacteria: a network of interactions. *Annual review of microbiology*, **63:** 363–83.

Branco S (2011) Fungal Diversity – An Overview, The Dynamical Processes of Biodiversity - Case Studies of Evolution and Spatial Distribution, PhD. Oscar Grillo (Ed.), ISBN: 978-953-307-772-7, InTech, Available from: http://www.intechopen.com/books/thedynamicalprocessesofbiodiversitycasestudiesofevolutionandspatialdistribution/fungaldiversity-an-overview

Bridge PD, Spooner BM, Roberts PJ (2005) The impact of molecular data in fungal systematics. *Adv. Bot. Res.*,**42**:33–67.

Brundrett M (2009) Mycorrhizas in Natural Ecosystems. *Advances in ecological research*, **21**: 171-313

Castro ML (1998) Annotated check-list of the Amanitaceae (Agaricales, Basidiomycotina) of the Iberian Peninsula and Balearic islands. *Mycotaxon* **67**: 227-245

Castro SC (2009) Micoturismo: Enquadramento Estratégico em Áreas Protegidas; Dissertação para obtenção do grau de Mestre em gestão e conservação dos recursoso naturais*; Instituto Superior de Agronomia*, Lisboa; 94 pag

Conrad LS, Keith AS, Sabine, Vincent R, John LS, André L, Wen C, and Fungal Barcoding Consortium (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *PNAS* **109** (16) 6241-6246

Corner EJ, Bas C (1962) The genus Amanita in Singapore and Malaya. *Persoonia*, **2**: 241–304.

Costa JC, Aguiar C, Capelo J, Lousã M, Neto C (1998) Biogeografia de Portugal Continental. *Quercetea*. ISSN 0874-5250, 5-56

Courtecuisse R, Duhem B, (2011) Guide des champignons de France et d'Europe. *Delachaux et Niestlé*

Curreli S (1994) Amanite rare in Sardegna: Amanita heterocroma S. Curreli nom. nov. e Amanita ponderosa Malençon & Heim. *Micologia Italiana,* **3**: 61-65

Daza A, Romero L, Santamaría C, Camacho M, Aguilar A, Mejías M, Moreno G, Pérez M, Manjón JL (2002) El gurumelo, una seta de primavera. *Quercus* **193**, 20-24

Daza A, Camacho M, Romero de la Osa L, Manjón JL, Moreno G, Santamaría C. (2007) Distribución espacial de la frutificación del hongo ECM comestible Amanita ponderosa Malençon & R. Heim durante seis años consecutivos en un encinar adehesado de la Sierra de Aracena (Huelva). *Investigación Agraria: Sistemas Y Recursos Florestales* **16**: 89-94

Earle FS (1909) The genera of the North American gill fungi. Bull. *N.Y. Bot. Gard*. **5**: 373–451.

Egli S (2011) Mycorrhizal mushroom diversity and productivity - an indicator of forest health? *Annals of Forest Science*, **68**(1), 81–88.

Elith J, Steven JP, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, **6**(5), e19379.

Erland S, Henrion B, Martin F, Glover LA and Alexander IJ (1994) Identification of the ectomycorrhizal basidiomycete Tylospora fibrillosa Donk by RFLP analysis of the PCRamplified ITS and IGS regions of ribosomal DNA. *New Phytol* , 0**126**:525-532.

Farmer DJ, Sylvia DM (1998) Variation in the ribosomalDNA internal transcribed spacer of a diverse collection of ectomycorrhizal fungi. *Mycol Res,* **102**:859-865

Garcin R (1984) Les Amanites européennes. Fédération mycologique Dauphiné-Savoie, Marlioz.

Gardes M, White TJ, Fortin JA, Bruns TD, Taylor JW (1991) Identification of indigenous and introduced symbiotic fungi in ectomycorrhizae by amplification of nuclear and mitochondrial ribosomal DNA. *Can J Bot,* **69**:180-190

Gilbert EJ (1940) Amanitaceae. *Iconogr. Mycol*. **27**, 1(1), 1–200.

Gilbert EJ, Kühner R (1928) Recherches sur les spores des Amanites. Bull. *Soc. Mycol. Fr.* **44**: 149–154.

Glen M, Tommerup IC, Bougher NL, O'Brien PA (2001) Interspecific and intraspecific variation of ectomycorrhizal fungi associated with Eucalyptus ecosystems as revealed by ribosomal DNA PCR-RFLP. *Mycol Res,* **105**:843-858.

Gomes EA, Barros EG, Kasuya MCM, Araújo EF (1999) Molecular characterization of Pisolithus spp. isolates by rDNA PCR-RFLP. *Mycorrhiza* **8**:197-202.

Harley JL, Smith SE (1983) Mycorrhizal Symbiosis. *Academic Press*, London.

Hawksworth DL (1991) The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycological Research*, **95**, 641–655

Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological Research* **105**(12), 1422–1432.

Henrion B, Chevalier G, Martin F (1994) Typing truffle by PCR amplification of the ribosomal DNA spacers. *Mycol Res* **98**:37-43.

Henriques JL (2010) Produção de tortulhos (Amanita ponderosa) na área do Parque Natural do Tejo Internacional. Avaliação da capacidade produtiva do campo de Cabeço de Mouro (Rosmaninhal – Idanha a Nova). *DRAPC*, 68 pág

Hershkovitz MA, Lewis LA (1996) Deep-level diagnostic value of the rDNA-ITS region. *Mol. Biol. Evol.*, **13**:1276–95

Hillis DM, Dixon MT (1991) Ribosomal DNA: Molecular evolution and phylogenetic inference. *Q. Rev. Biol.*, **66**:411–53.

Horton TR, Bruns TD (2001) The molecular revolution in ectomycorrhizal ecology: peeking into the black-box. *Molecular ecology*, **10**(8), 1855–71. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11555231

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics, *Genet. Sel. Evol.* **34**, 275–305. doi:10.1051/gse

Kirk PM, Cannon PF, Minter DW, Stalpers JA (2008) Dictionary of the Fungi, 10th Edition. *CABI Publishing,Wallingford*, UK.

Konrad P, Maublanc A (1948) Les Agaricales. *Paul Lechevalier*, Paris, France.

Louro R, Calado M, Pinto B, Santos-Silva C (2009) Epigeous Macrofungi of the Parque de Natureza de Noudar in Alentejo (Portugal). *Mycotaxon* **107**:49–52

Mardis ER (2008) Next-generation DNA sequencing methods. *Annual review of genomics and human genetics*, **9**, 387–402.

Miller MR, Dunham JP, Amores A, Cresko W, Johnson E (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome research*, **17**(2), 240–8.

Moreno G, Platas G, Peláez F, Bernedo M, Vargas A, Daza A, Santamaría C (2007) Molecular phylogenetic analysis shows that Amanita ponderosa and A. curtipes are distinct species. *Mycological Progress*, **7**(1), 41–47.

Morgado L, Martins L, Gonçalves H, Oliveira P (2006) Estudo de intoxicações causadas por ingestão de macrofungos na região do Alto Alentejo. *Anais da Associação Micológica a Pantorra*, **6**: 65-74

Morgado L, Vila-Viçosa C, Mendes P, Castilho R, Pinto-Gomes C (2011) Reconhecimento do Potencial Micológico de Barrancos – Relatório Final. Grupo de Trabalho em Micologia do CEAI. (unpublished)

Morris M, Smith ME, Rizzo DM, Rejmanek M, Blesdsoe S (2008) Contrasting ectomycorrhizal fungal communities on the roots of co-occurring oaks (Quercus spp.) in a California woodland. *New Phytologist* **178**: 167-176.

Murat C, Mello A, Abbà S, Vizzini A, Bonfante P (2008) Edible Mycorrhizal Fungi: Identification, Life Cycle and Morphogenesis; Mycorrhiza 3ªEdition, *Springer-Verlag*, Berlin Heidelberg

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca G, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**(6772), 853–858.

Neville P, S Poumarat (2004) Amaniteae-1. *Fungi Europaei,* **9**: 1-1119

Nilsson RH, Kristiansson E, Ryberg M, Hallenberg N (2008) Intraspecifi c ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evolutionary Bioinformatics*, **4**: 193–201.

Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?. *Global Ecology and Biogeography*, **5**, 361–371

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**: 231–259

Pinho-Almeida F (1994) Estudos taxonómicos do género Amanita. Secção Amidella – complexo Lepiotoides: A. curtipes Gilbert, A. lepiotoides Barla e A. ponderosa Malençon & Heim. *Revista Biol.* (Lisboa) **15**: 131-151

Piqueras-Carrasco J (2013) Intoxicaciones por setas, una actualización. *Revista Española de Medicina Legal,* (0). (in press) Available online 17 July 2013

Real Decreto 30/2009, de 16 de enero, por el que se establecen las condiciones sanitarias para la comercialización de setas para uso alimentario. Boletín Oficial del Estado nº 20, 23 de enero de 2009, Sec I. Pág. 7861-7871

Rowe HC, Renaut S, Guggisberg A (2011) RAD in the realm of next-generation sequencing technologies. *Molecular ecology*, **20**(17), 3499–502.

Roze E (1876) Essai d'une nouvelle classification des Agaricacées. Bull. *Soc. Bot. Fr.* **23**: 45–54.

Santos-Silva C, Machado H, Guinberteau J (2011) Ecologie de Amanita ponderosa et gestion des écosystèmes. In Ecologie de Amanita ponderosa et gestion des écosystèmes, 114 - 126. *Micosylva*: Jean Rondet et Fernando Martin

Schoch CL, Seifert K, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(16), 6241–6.

Sillero N (2011) What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, **222:** 1343–1346

Simard SW, Beiler KJ, Bingham MA, Deslippe JR, Philip LJ, Teste FP. 2012. Mycorrhizal networks: mechanisms, ecology and modelling. *Fungal Biology Reviews* **26**: 39-60

Singer R (1951) The Agaricales (mushrooms) in modern taxonomy. *Lilloa*, **22**: 5–832

Smith JE, Read DJ (1997) Mycorrhizal symbiosis. *Academic Press*, London.

Soberón J, Peterson AT (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**: 1–10.

Tedersoo L, May T, Smith M (2010) Ectomycorrhizal lifestyle in fungi: global diversity, distribution, and evolution of phylogenetic lineages. *Mycorrhiza*, **20**(4), 217–263.

Tulloss RE (2005) Amaniteae: Amanita, Limacella and Torrendia. *Mycotaxon,* **92**: 474-484

Verbruggen E, Kiers ET (2010) Evolutionary ecology of mycorrhizal functional diversity in agricultural systems. *Evolutionary Applications,* **3**: 547-560

Weiß M, Yang ZL, Oberwinkler F (1998) Molecular phylogenetic studies in the genus Amanita. *Canadian Journal of Botany*, **76**(7), 1170–1179.

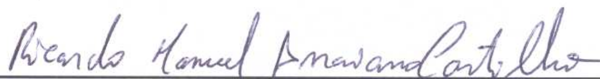
Declaração

Declara-se para todos os efeitos que os resultados apresentados nesta dissertação de mestrado não têm efeito para publicação.

Statement

It is declared for all purposes that the results presented in this dissertation have no effect for publication

O autor / The author



(Ricardo Manuel Arraiano Castilho)

Porto, 3 de Abril de 2014