

PREDICTING THE DISTRIBUTION OF SUITABLE HABITAT OF
ARMILLARIA MELLEA (AGARICALES, PHYSALACRIACEAE) USING
ECOLOGICAL NICHE MODELING

by

Bryce Allen Kerr

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Biology

Middle Tennessee State University

August 2018

Thesis Committee:

Dr. Sarah E. Bergemann, Chair

Dr. R. Stephen Howard

Dr. Ashley B. Morris

ACKNOWLEDGEMENTS

I would like to sincerely thank my adviser, Dr. Sarah Bergemann, for her continued support throughout my time as a graduate student. I thank the other members of my thesis committee, Dr. Ashley Morris and Dr. R. Stephen Howard, for providing additional guidance during my time in this program. I further thank Middle Tennessee State University for providing access to resources and grants, and I thank the other Biology graduate students.

Outside of MTSU, I would like to thank Jay Justice who collected several of the specimens in the southern United States in this study. This project would also not have been possible without the contribution of vouchered specimens from the herbaria mentioned in this thesis.

While they did not contribute directly to this project, I thank my parents Robin and Rebecca Kerr, and my sister, Mandi Phipps. I remember my childhood dog, Gabby, who passed in May of 2016 for being my closest and truest friend. I thank my four pet rats for keeping me relatively entertained. Finally, I express gratitude to my partner, Jess, for tolerating my persistent pessimism over these past three years.

ABSTRACT

Armillaria mellea is an economically and ecologically significant fungal pathogen; however, the distribution of this fungus in eastern North America is not well understood. This study aims to use ecological niche modeling to predict the extent of suitable habitat of the species and identify the environmental predictors that affect its ecogeographic distribution. In this study, herbarium vouchers and mycelia obtained from cultures were selected to document occurrences, which were identified and annotated using a combination of morphological and molecular analyses. The records of all specimens that were confidently identified were georeferenced. Environmental variables were compiled from relevant databases, values of variable importance were estimated given prior constraints, and a smaller subset of abiotic variables were identified that were important for predicting habitat suitability of *A. mellea* in eastern North America. Annual mean temperature had the greatest importance on the predicted suitable habitat of *A. mellea*. Five other variables (annual mean moisture index, mean temperature of the driest quarter, mean diurnal temperature range, annual precipitation, and precipitation seasonality) were identified as significant in contributing to the model. This study assembled a collection of annotated herbarium vouchers that represent the occurrence of *A. mellea* in eastern North America. The results of the present study indicate that niche modeling may be used to understand the suitable habitat of an important pathogenic fungus.

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF APPENDICES.....	vii
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. METHODS	6
2.1. <i>Sample acquisition</i>	6
2.2. <i>DNA Extraction, PCR amplification and sequencing of nuclear loci</i>	7
2.3. <i>Phylogenetic analyses</i>	9
2.4. <i>Restriction fragment length polymorphisms (RFLPs) of nuclear ribosomal intergenic spacer (IGS1)</i>	9
2.5. <i>Morphological analyses</i>	10
2.6. <i>Ecological niche modeling</i>	13
CHAPTER 3. RESULTS	16
CHAPTER 4. DISCUSSION.....	27
REFERENCES	33
APPENDIX.....	40

LIST OF TABLES

Table 1. Variable importance of environmental predictors used for ecological niche modeling of *Armillaria mellea* in eastern North America 15

LIST OF FIGURES

Figure 1. The geographic location of all specimens used to predict the distribution and potential suitable habitat of <i>Armillaria mellea</i> in eastern North America	12
Figure 2. Maximum likelihood phylogram of specimens for species identification of <i>Armillaria mellea</i> using the <i>actin-1</i> gene.....	18
Figure 3. Maximum likelihood phylogram of specimens for species identification of <i>Armillaria mellea</i> using the <i>tefl</i> gene.....	19
Figure 4. Receiver operator characteristic (ROC) curve of the best fitted model in Maxent for predicting the distribution and potential suitable habitat of <i>Armillaria mellea</i> in eastern North America	21
Figure 5. The predicted distribution and potential suitable habitat of <i>Armillaria mellea</i> in eastern North America	22
Figure 6. Jackknife estimates of the variable importance for the model predicting the distribution and potential suitable habitat of <i>Armillaria mellea</i> in eastern North America	24
Figure 7. Response curves of each environmental variable in the best fitted model used to predict the distribution and potential suitable habitat of <i>Armillaria mellea</i> in eastern North America	26

APPENDIX

Appendix. Specimens used to model the distribution and potential suitable habitat of
Armillaria mellea 40

CHAPTER 1. INTRODUCTION

Ecological niche modeling (ENM) uses associations between environmental predictors and the occurrence of a species to identify suitable habitats across the geographic range of a species (Guisan and Zimmermann, 2000; Peterson, 2001; Soberon and Peterson, 2005). Models are often referred to as “ecological niche” (Alvarado-Serrano and Knowles, 2014) and “species distribution” (Pineda and Lobo, 2009); however, the methodological approaches are very similar. First, the study area is modeled as a map composed of grid cells at a specified resolution dependent on the geographic scale of the study. Geographic data representing species’ occurrence, typically in the form of presence or absence data, are selected. The dependent variable of ENM is the distribution of the species and a set of environmental variables is used to describe the environmental characteristics of each cell, and a probability distribution of the species occurrence is generated, given prior constraints to assess the degree to which cells are suitable or unsuitable for the species (Elith et al., 2011; Peterson and Nakazawa, 2008). In the absence of extensive surveys, only presence data are available to indicate the occurrence of the species. In this case, the Maximum entropy (Maxent) algorithm is often used for modeling the species distribution (Baldwin, 2009; Pineda and Lobo, 2009; Warren and Seifert, 2011).

ENM has been performed for a handful of macrofungi (Ponce et al., 2011; Sato et al., 2012; Smith et al., 2016; Wolfe et al., 2010; Wollan et al., 2008). For these fungi, vegetative structures are mostly subterranean, and therefore, largely cryptic, and sexual or asexual structures are sporadically produced and ephemeral. Consequently, accurate

presence data are difficult to obtain. ENM predicts the extent of suitable habitat of a species, which is largely dependent on the accuracy of species identification (Canhos et al., 2004; Loiselle et al., 2008; Smith et al., 2016). For most species of fungi, obtaining occurrence data for use in ENM is challenging since the geographic data are sparse, the collections gathered are typically obtained using opportunistic sampling, and generally, the collections of most macrofungi are underrepresented in surveys. In these cases, vouchers in herbaria provide an additional source of historical occurrence data for use in ENM.

Armillaria (Fr.) Staude is a genus of plant pathogens that cause some of the most destructive forest diseases worldwide (Baumgartner et al., 2011). *Armillaria* are white rot fungi that can efficiently decompose all components of plants (including lignin) and are also necrotrophs; the fungus colonizes plant roots, kills the host tissues, and utilizes the dead tissues as a source of nutrition (Baumgartner et al., 2011). *Armillaria* is often identified by its rhizomorphs, which are hyphal aggregates (1–3 mm in diameter) (Morrison, 2004; Prospero et al., 2006). Similar to mycelia, rhizomorphs form a network in the soil that can reach immense sizes by colonization of multiple hosts and wood (Ferguson et al., 2003; Smith et al., 1992). Root contact by rhizomorphs is the main mode of expansion in forests; however, long distance dispersal is facilitated by forcible discharge of airborne basidiospores from basidiomata (Travadon et al., 2012). Basidiospores can disperse long distances (ca. 2000 km) which allows fungal individuals to colonize new territories (Baumgartner et al., 2010b).

In recent decades, the development of DNA sequence-based analysis has improved the accuracy of species identification for *Armillaria* (Physalacriaceae, Basidiomycota) (Elías-Román et al., 2018; Harrington and Wingfield, 1995; Tsykun et al., 2013). The ribosomal RNA genes (rDNA), including the internal transcribed spacers 1 and 2 (ITS) with the intervening 5.8S (ITS) and the intergenic spacer 1 (IGS1), have been useful in distinguishing some *Armillaria* species but not all (Chillali et al., 1998; Harrington and Wingfield, 1995; White et al., 1998). More recently, low-copy nuclear loci such as the translation elongation factor-1 alpha (*tef1*) region have successfully been used to identify *Armillaria* species (Coetzee et al., 2011; Maphosa et al., 2006). Since DNA sequence-based methods can identify *Armillaria* species with a high level of accuracy, this approach provides a better method for species identification.

One of the more ecologically and economically important species, *A. mellea*, occurs in temperate regions of the northern hemisphere with disjunct populations located in Asia, Europe, and North America (Coetzee et al., 2011). In Europe, it is a highly virulent pathogen that can cause increased levels of mortality in hosts (Guillaumin et al., 1993). In North America, the pathogen occurs in the western US and eastern US and Canada (Baumgartner and Rizzo, 2001a, 2001b; Bruhn et al., 1997, 2000; Harrington and Rizzo, 1993; McLaughlin 2001). Analysis of genetic diversity for populations from North America demonstrate that these are two, geographically isolated populations that are genetically divergent (Baumgartner et al., 2010b).

In eastern North America, *A. mellea* is considered a secondary pathogen that often attacks trees infected by other pathogens or that are affected by environmental stressors

(Bruhn et al., 2000; Haavik et al., 2015; Marçais and Bréda, 2006; Thomas et al., 2002). In all regions, *A. mellea* infects multiple hosts and more often, hardwoods (Banik et al., 1995; Harrington and Rizzo, 1993; McLaughlin 2001). The species is often identified as a causal agent of oak decline (Marçais and Bréda, 2006; Thomas et al., 2002).

While prior studies have described the distribution of *Armillaria* for western North America (Baumgartner and Rizzo, 2001b), there is no such description of *A. mellea* in eastern North America. Surveys have provided a basis for understanding local distributions of *Armillaria* species through sampling in federal and state as well as larger regions (e.g. states or provinces) of eastern North America (Banik et al., 1995; Bruhn et al., 1997, 2000; McLaughlin 2001); however, the local sampling methods limit inference of the geographic extent of the population. More extensive sampling for population genetics and phylogenetic studies of *A. mellea* in eastern North America have also provided a wealth of occurrence data (Baumgartner et al., 2010b; Hughes et al., 2013); however, there are no studies that have collated occurrence data from multiple sources to predict the extent of suitable habitat of *A. mellea* in eastern North America.

This study focuses on assessing the distribution and extent of suitable habitat of *A. mellea* in eastern North America. Vouchers from herbaria and mycelia from culture repositories were obtained. Species identification was performed to provide an annotated collection of occurrence data for *A. mellea* that can be used in ENM. The primary objectives of the present study were as follows: i) the collation of an annotated collection of herbarium vouchers and environmental samples of *A. mellea* that describe the known occurrence of the species in eastern North America; ii) to predict the extent of suitable

habitat of *A. mellea* in eastern North America using ecological niche modeling, and; iii) to evaluate the importance of the environmental predictors used to predict suitable habitats of *A. mellea* when compared to optimal growth requirements identified from laboratory studies.

CHAPTER 2. MATERIALS AND METHODS

2.1. Sample acquisition

The ability to reliably model the distribution of suitable habitat of *A. mellea* required accurate identification of all collections. Three methods for species identification were used: i) partial sequences of two low-copy nuclear loci (*actin-1* and *tef1*) were obtained and phylogenetic analyses were performed for each locus separately; ii) in the absence of PCR amplification or successful sequencing, restriction fragment length polymorphisms (RFLPs) of the IGS1 rDNA were employed and; iii) morphological identification based on macroscopic and microscopic characters of the basidiomata was conducted on vouchers that were not identified by sequencing or RFLPs. The herbaria providing vouchered material in this study include: USDA Center for Forest Mycology Research (CFMR), State University of New York at Cortland (CORT), Davis and Elkins College Herbarium (DEWV), Duke University Herbarium (DUKE), Field Museum of Natural History Herbarium (F), William Sherman Turrel Herbarium at Miami University (MU), University of Tennessee Fungal Herbarium (TENN), and Massey Herbarium at Virginia Tech University (VPI). Vouchers of basidiomata from these herbaria ($n = 115$), vouchers collected over the duration of the study ($n = 10$), and cultures obtained from Clemson University, Ontario Forest Research Institute, United States Department of Agriculture (USDA) Forest Products Lab, University of Missouri, and the University of Tennessee ($n = 13$) were identified in this study. The vouchers obtained from herbaria were selected to increase the geographic coverage of occurrence data and specimens were sampled from the last 30 years of

collections to increase the success of DNA extraction, PCR amplification, and sequencing. Information including collector ID, herbarium accession number, herbarium code, GenBank accession number, collector, description of collecting locale, year collected, host, coordinates, and method of species identification were provided for collections identified in this study ($n = 105$) (Appendix).

2.2. DNA Extraction, PCR amplification, and sequencing of nuclear loci

Mycelia or tramal tissues from the pilei or stipes of the basidiomata were placed in 1.5 mL tubes with two, 6 mm glass beads, lyophilized for 30 min in a Labconco Freeze Dryer 8 (Kansas City MO USA) and pulverized for 20 seconds with a QBioGene FastPrep 120 (Carlsbad CA USA). DNA extractions were performed on pulverized tissues by overnight incubation at 65°C with 2× cetyltrimethylammonium bromide (CTAB) extraction buffer. DNAs were further purified by adding phenol-chloroform-isoamyl alcohol (25:24:1) following the procedures in Kluting et al. (2014). The supernatants were transferred to a new 1.5 mL centrifuge tube, and 800 µL of MP Biomedicals Genomic salt solution (Solon OH USA) was added. Supernatants were washed by binding genomic DNA to a glass milk column (MP Biomedicals, Solon OH USA), adding 70% ethanol and washing by centrifugation for 1 min. DNAs were eluted in 35 µL of 0.1× TE buffer and samples were stored at -20°C in a non-frost freezer.

Two, low-copy, partial, nuclear gene regions, actin-1 subunit (*actin-1*) and translation elongation factor subunit-1 alpha (*tef1*) (Baumgartner et al., 2010a; Maphosa et al., 2006) were selected for analysis due to previously documented high levels of

variability and consequential success in distinguishing among *Armillaria* species. PCRs for both regions were performed in 25 μ L reactions using 1 \times PCR buffer, 2 mM of magnesium chloride (MgCl₂), 2mM of each dNTP, 1 μ M of forward and reverse primers, 0.2 mg/mL bovine serum albumin (BSA), 0.5 M betaine, 0.025 U Taq (Promega, Madison WI USA) and 3-5 μ L of template DNA. PCRs of the *actin*-1 subunit were obtained following the PCR amplification and cycling protocols outlined in Baumgartner et al. (2012).

PCRs of the *tef1* region were obtained using two sets of primers. For the initial amplifications, primers designed to amplify *A. mellea* (Ef-AMf and Ef-AM785r) were selected (Baumgartner et al., 2010a). Cycling protocols followed Baumgartner et al. (2010a). In the absence of positive amplification of the *tef1* gene using primers specifically designed to amplify *A. mellea*, PCR reactions were generated using degenerative primers, EF595F and EF1160R (Maphosa et al., 2006). PCRs employed PCR amplification and cycling protocols described in Maphosa et al. (2006) after reducing the annealing temperature of the cycling protocol to 54°C.

PCRs were purified using Exo-SAP-IT (GE Healthcare, Little Chalfont UK) following the protocols detailed in Kluting et al. (2014). Forward and reverse sequences were generated using the protocol in Kluting et al. (2014) and obtained by bi-directional sequencing using an Applied Biosystems 3130xl Genetic Analyzer at Middle Tennessee State University or GenHunter Corporation (Nashville, TN USA). Sequencher 4.8.0 (GeneCodes Corporation, Ann Arbor MI USA) was used to manually edit chromatograms for use in phylogenetic analyses. Alignments were generated for each

region separately using multiple sequence alignment with the default parameters in MUSCLE 3.8.31 (Edgar, 2004) and subsequently manually aligned in Mesquite 3.31 (Maddison and Maddison, 2018). Additional sequences of *Armillaria*, *Desarmillaria*, and outgroups selected from genera with prior use in *Armillaria* identification (Baumgartner et al., 2012) from GenBank were also used in phylogenetic analyses.

2.3. Phylogenetic analyses

For *actin-1* and *tef1*, introns were delimited using Augustus 3.2 and excluded before performing phylogenetic analyses (Stanke et al., 2004). Phylogenetic analyses for each locus were analyzed individually using Maximum Likelihood (ML) analysis in RAxML-HPC 8.2.10 (Stamatakis, 2006) with the CIPRES Gateway (Miller et al., 2010). Phylogenies from each locus were inferred by 1000 ML replicate analyses and assumed a GTR model with a gamma distribution. Levels of support were inferred with ML bootstrap (MLBS) values. MLBS analyses were employed in RAxML with 1000 multiparametric bootstrap replicates. The RAxML manual states the GTR model is most appropriate for this software, thus the GTRCAT model of nucleotide substitution was selected. Phylograms were viewed with FigTree 1.4.3 (Rambaut, 2012) and edited in Adobe Illustrator 22.1 (San Jose CA USA).

2.4. Restriction fragment length polymorphisms (RFLPs) analysis of nuclear ribosomal intergenic spacer (IGS1)

If PCR amplification or direct sequences for either *actin-1* or *tef1* was not

successful, restriction fragment length polymorphisms (RFLPs) of the nuclear ribosomal intergenic spacer 1 (IGS1) were performed. Harrington and Wingfield (1995) found that by PCR amplification of the IGS1 region and then restricting with *Alu* I endonuclease, most species of *Armillaria* in North America were differentiated by RFLP patterns. Here, the nuclear ribosomal IGS1 region of *Armillaria* spp. common to eastern North America (*A. calvescens*, *A. gallica*, *A. gemina*, *A. mellea*, *A. sinapina*, and *Desarmillaria tabescens* (formerly *A. tabescens*)) and vouchers ($n = 28$) for which sequences were not obtained were PCR amplified and digested with *Alu* I. PCR reactions were performed in 25 μ L reactions using the primers LR12R (Veldman et al., 1981) and O-1 (Duchesne and Anderson, 1990) following the procedures of Harrington and Wingfield (1995). PCR cycling protocols were employed at 95°C for 3 min, with 34 cycles at 94°C for 30 s, 60°C for 40 s, and 72°C for 1 min, followed by 10 min at 72°C. Restriction digests were prepared in 35 μ L reactions with 16 μ L PCR product, 2 units of *Alu* I enzyme and 3 μ L of buffer. Digests were incubated at 37°C for 2 hours. Products were visualized in agarose gels.

2.5. Morphological analyses

When PCR amplification, sequencing of low-copy nuclear loci or RFLPs were unsuccessful, more extensive morphological examination of vouchered specimens was employed ($n = 14$) (see Appendix for vouchers used). First, vouchered specimens were screened for macroscopic features following the characters for species identification as described by Watling et al. (1982). Since many macroscopic characters are difficult to

distinguish on dried basidiomata, microscopic characters were also used. Small sections (ca. 2 mm²) of dried lamellar tissue were rehydrated in 95% ethanol for 1 min and submerged in dH₂O. Tissues were mounted in 3% potassium hydroxide (KOH) and 1% congo red stain. The absence of clamp connections at the base of the basidia was used to distinguish between *A. mellea* and other *Armillaria* species (Korhonen, 1980; Motta and Korhonen, 1986; Ullrich and Anderson, 1978). Microscopic features were examined at 1000× under an oil immersion lens with a Zeiss Axio Scope A.1 (Oberkochen BW DE).

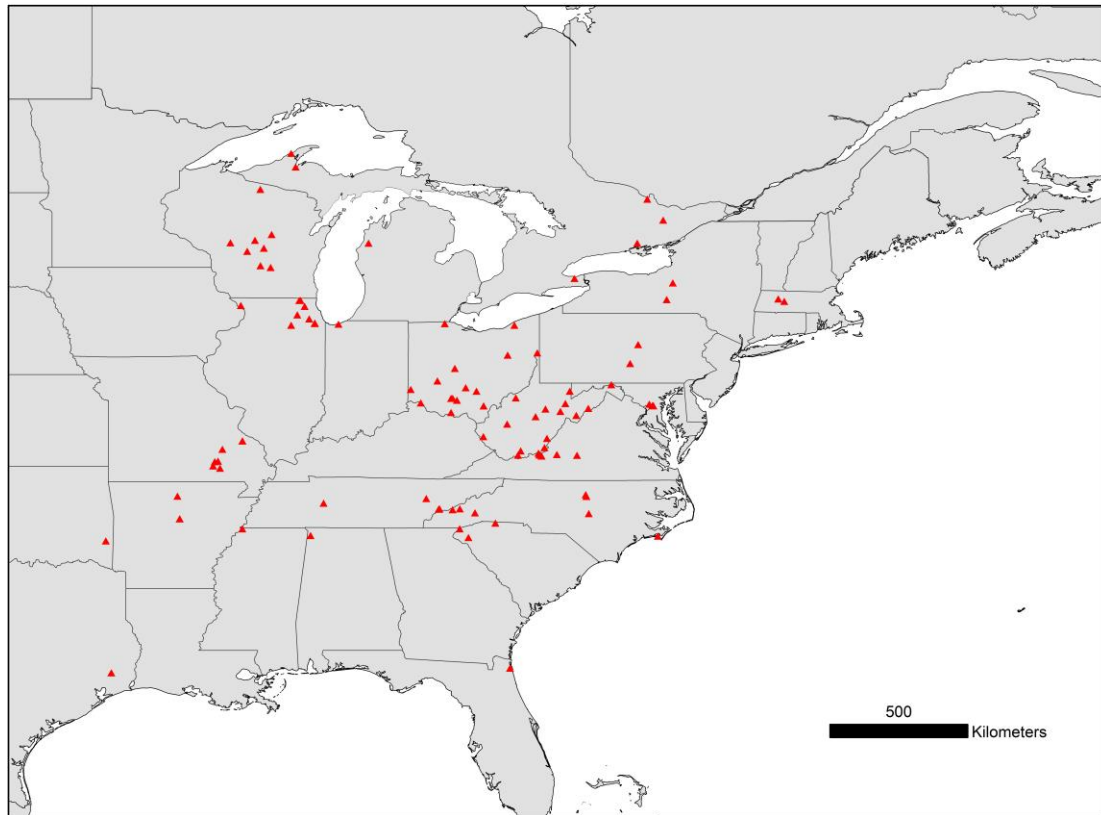


Figure 1. The geographic location of all specimens used to predict the distribution and potential suitable habitat of *Armillaria mellea* in eastern North America.

2.6. Ecological niche modeling

Maxent 3.4.1 (Phillips et al., 2017) was used to model the distribution and potential suitable habitat of *A. mellea*. A raster data file containing variable information from georeferenced locales within grids of equal size represented each environmental variable. Raster cell size was scaled to the range of the study (10 arc-minutes). Maxent was used to predict the distribution of *A. mellea* by utilizing the cell values that contained occurrence points to generate a distribution of probable species occurrence (Phillips et al., 2006).

All location data for *A. mellea* were georeferenced using GEOLocate 2.0 (Tulane University, New Orleans LA USA) and transferred to ArcMap 10.0 (ESRI, Redlands CA USA). To avoid spatial autocorrelation, presence data were projected for the raster resolution and points were selectively removed until a single locale occurred ($n = 102$) in each cell (Dormann et al., 2007; Segurado et al., 2006) (Fig. 1). Environmental data were imported from the CliMond database (Kriticos et al., 2012).

Once the initial model was generated, a subset of presence data was treated as “testing” data to assess model accuracy; the remaining distribution points were used as “training” data to generate the model (Jiménez-Valverde and Lobo, 2006; Pearce and Boyce, 2006). Model performance was evaluated with a receiver operator characteristic (ROC) curve by comparing the model obtained from testing data to the probabilities of occurrence that were generated from training and testing data. The area under the curve (AUC) was calculated to provide estimates of model accuracy. A regularization modifier was applied to multiply each cell by a fixed value ($\times 1.5$) to reduce overfitting due to

sampling bias. After employing a regularization modifier, an AUC value of ≥ 0.85 was considered acceptable for model performance (Warren and Seifert, 2011).

The models were generated without equal weighting of environmental variables, by analyzing the contribution of each layer for the predicted occurrence (Phillips et al., 2006). Maxent also estimates logistic response curves of the impact of each variable and heuristic and jackknife estimates of variable importance. These estimates were made by determining the AUC, the testing gain and training gain for each variable independently, and then by excluding each variable in the model (van Gils et al., 2014). Jackknife estimations were used to generate a set of models for each variable and all other variables excluding the variable under investigation (Pearson et al., 2007).

Correlation between environmental variables can influence ENM output by overestimating the impact of environmental variables. This study employed a step-wise variable removal process (van Gils et al. 2014), which used correlation coefficients and the generated values of variable importance to remove highly correlated variables. After removing correlated variables, six variables identified by iterant models (AUC < 0.85) were used to model the distribution and habitat suitability of *A. mellea* (Table 1).

Table 1.

Variable importance of environmental predictors used for ecological niche modeling of *Armillaria mellea* in eastern North America.

Variable	Source ID	Percent Contribution	Permutation Importance	Testing gain	Training gain	AUC
annual mean temperature	Bio01	63	51	0.745	1.263	0.955
annual mean moisture index	Bio28	18	27	0.251	0.166	0.655
mean temperature of the driest quarter	Bio09	7	3	0.470	0.776	0.853
mean diurnal temperature range	Bio02	6	12	0.180	0.121	0.601
annual precipitation	Bio12	4	1	0.247	0.302	0.603
precipitation seasonality	Bio15	1	6	0.323	0.409	0.748

CHAPTER 3. RESULTS

Identification of *A. mellea* collections employed three main methods: maximum likelihood analyses of partial sequences from two, nuclear loci ($n = 91$ attempted), RFLP analyses of nuclear IGS1 rDNA ($n = 28$ attempted), or morphological identification ($n = 12$ attempted). For the phylogenetic analyses performed, 61 *actin-1* and 59 *tef1* sequences were generated for the purposes of this study (Appendix). From the *actin-1* and *tef1* sequences, 88 of the 91 collections (97%) were identified as *A. mellea*. In both phylogenetic analyses, *A. mellea* was monophyletic and recovered with strong support (*actin-1*, MLBS = 99; *tef1*, MLBS = 95) (Figs. 2, 3). Trees of both gene regions supported the separation of *A. mellea* into monophyletic groups consistent with their geographic origins; however, the relationships among the monophyletic groups were, for the most part, was unsupported (Figs. 2, 3). In this case, all specimens collected from eastern North America were separated from collections from western North America, Asia, and Europe (Figs. 2, 3). All samples used for ENM analysis were grouped with the eastern North American clade (Figs. 2, 3).

To identify the isolates based on previously described RFLPs (Harrington and Wingfield, 1995), the IGS1 region was PCR amplified and restricted with *Alu I* for vouchers when the *actin-1* nor *tef1* genes could not be PCR amplified or sequenced. In the RFLP analysis, the patterns produced by digests resulted in 8 distinct RFLP patterns. Since not all RFLP patterns were recognizable with sample DNAs from *Armillaria* species, only isolates with RFLPs identical to *A. mellea* were scored. Of the 28 vouchers included, 39% ($n = 11$) were positively identified as *A. mellea*. Among the vouchers

identified as *A. mellea*, two patterns were present, referred to as types A (490, 180 bp) and B (320, 155 bp) by Harrington and Wingfield (1995), with type A being the most common of the two RFLP patterns observed in the vouchered specimens.

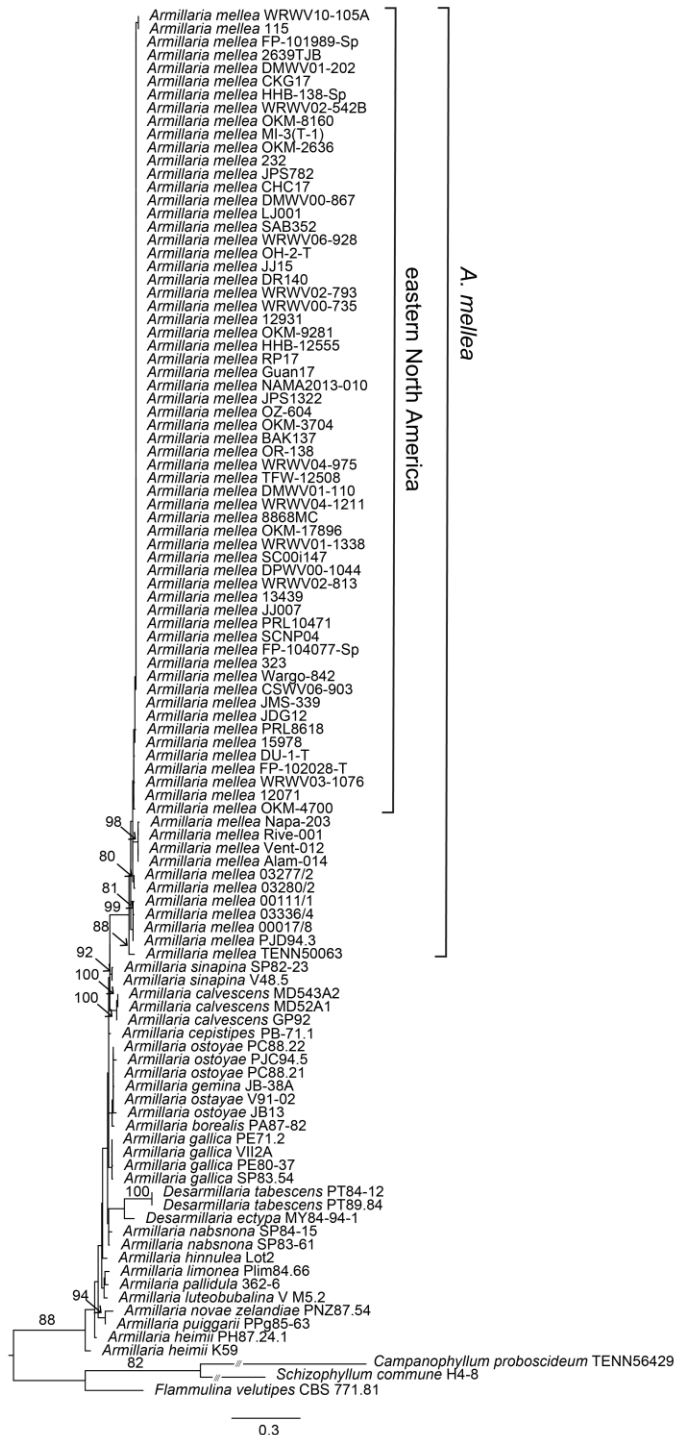


Figure 2. Maximum likelihood phylogram of specimens for species identification of *Armillaria mellea* using the *actin-1* gene. Maximum likelihood bootstrap (MLBS) values are reported for branches (MLBS \geq 80); MLBS for *Armillaria mellea* = 99. Truncated branches are denoted with three slanted lines.

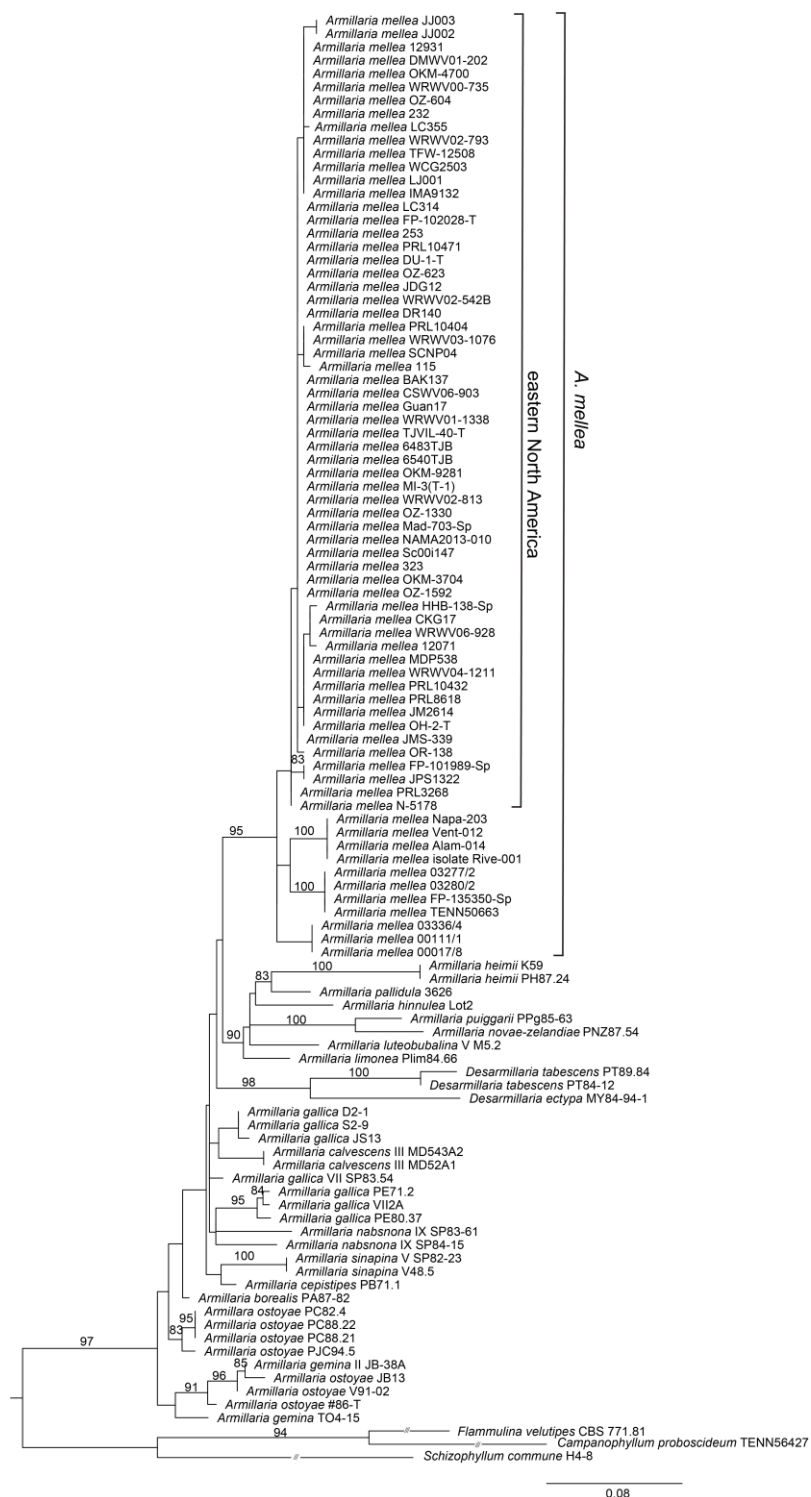


Figure 3. Maximum likelihood phylogram of specimens for species identification of *Armillaria mellea* using the *tef1* gene. Maximum likelihood bootstrap (MLBS) values are reported for branches (MLBS \geq 80); MLBS for *Armillaria mellea* = 99. Truncated branches are denoted with three slanted lines.

The remaining vouchers ($n = 12$) were identified based on morphological characters. Ten of the vouchers were identified as *A. mellea* based on the macroscopic morphology of basidiomata and the absence of clamp connections at the base of the basidia.

Georeferenced locales of *A. mellea* ($n = 102$) were used to predict the probabilities of occurrence based on the environmental predictors and included only those occurrence data that represented a unique geographic location. In total, six environmental predictors were identified that contributed to model fit of all replicate analyses and the conservation of variable importance across training, testing, and AUC (Table 1). The best fitted model across ten replicate runs showed a training gain of AUC of 0.899 (testing AUC = 0.956, testing SD = 0.013) (Fig. 4).

Using Maxent, a predictive map of occurrence probability for *A. mellea* in eastern North America was generated (Fig. 5). In this model, the distribution of *A. mellea* was constrained at 46°N, and regions of moderate to high suitability were predicted from 33-40°N (Fig. 5). In the northern extent of the range, the absence of suitable habitat was apparent to the east at longitudes > 68°W and to the west at longitudes > 97°W with moderate habitat suitability from 72-92°W (Fig 5). The geographic extent of the highest predicted probabilities (> 0.7) occurred in areas of the southeast, further west in the Ozark Mountains and northeastern US (40-44°N) including the Appalachian Mountains, the Adirondack Mountains, and the southern extent of the White Mountains (Fig. 5).

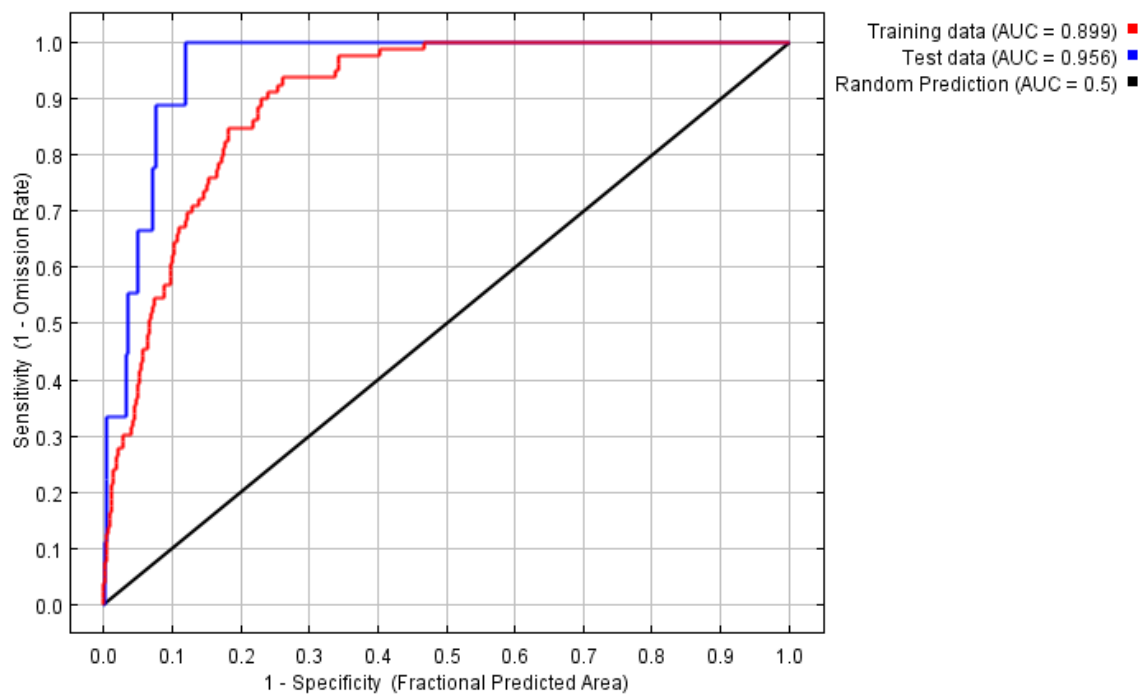


Figure 4. Receiver operator characteristic (ROC) curve of the best fitted model in Maxent for predicting the distribution and potential suitable habitat of *Armillaria mellea* in eastern North America.

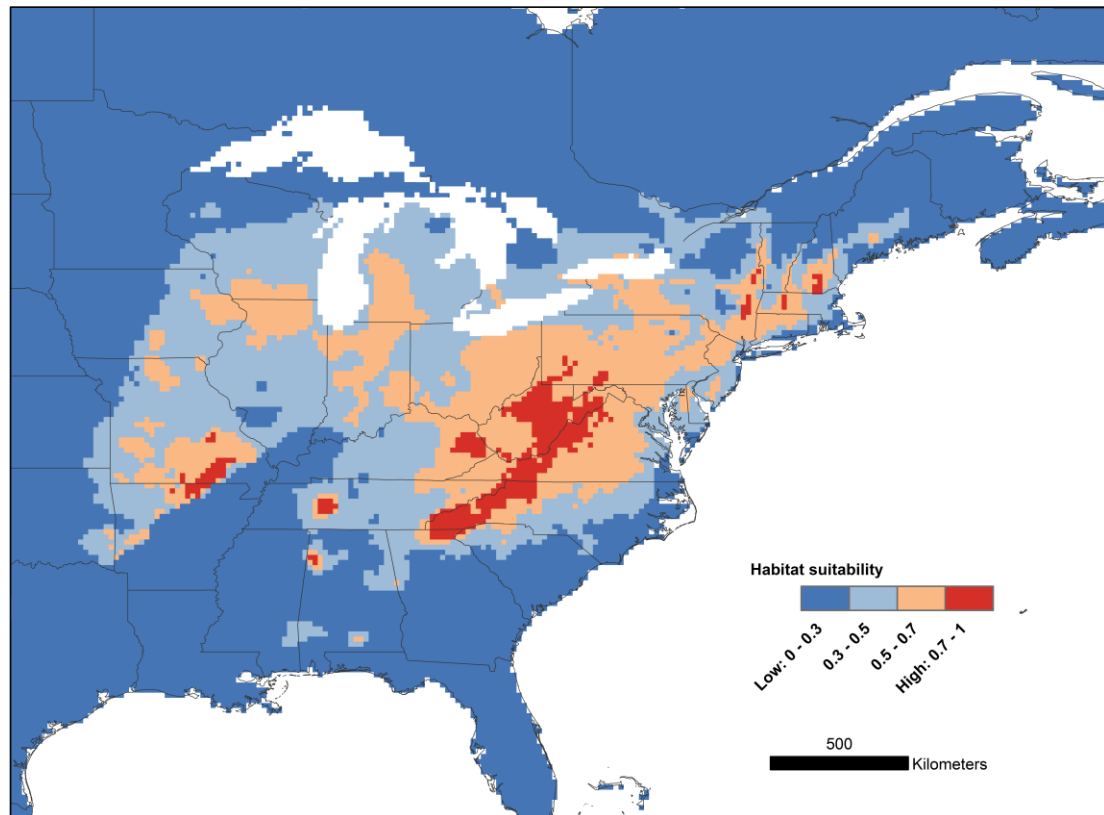


Figure 5. The distribution and potential suitable habitat of *Armillaria mellea* in eastern North America. Predicted probabilities of occurrence are divided into four categories.

The importance of environmental variables in model generation was assessed through two main approaches: i) analyses during model training, and; ii) jackknife estimations. During model training, annual mean temperature (63%) and annual mean moisture index (18%) had the highest percent contributions (Fig. 7). These variables also produced the highest permutation importance (51% and 27% respectively) accounting for 78% of variation in the model (Table 1).

Jackknife analysis of variables showed that the annual mean temperature contributed the most information that was not captured by other environmental predictors (Fig. 6). Mean temperature of the driest quarter and precipitation seasonality also had a significant contribution; however, the model performance was lower for models of training gain, testing gain, and AUC when these two variables were excluded (Fig. 6). Annual mean temperature, annual mean moisture index, and mean diurnal temperature range had the greatest importance after their removal (Fig. 6).

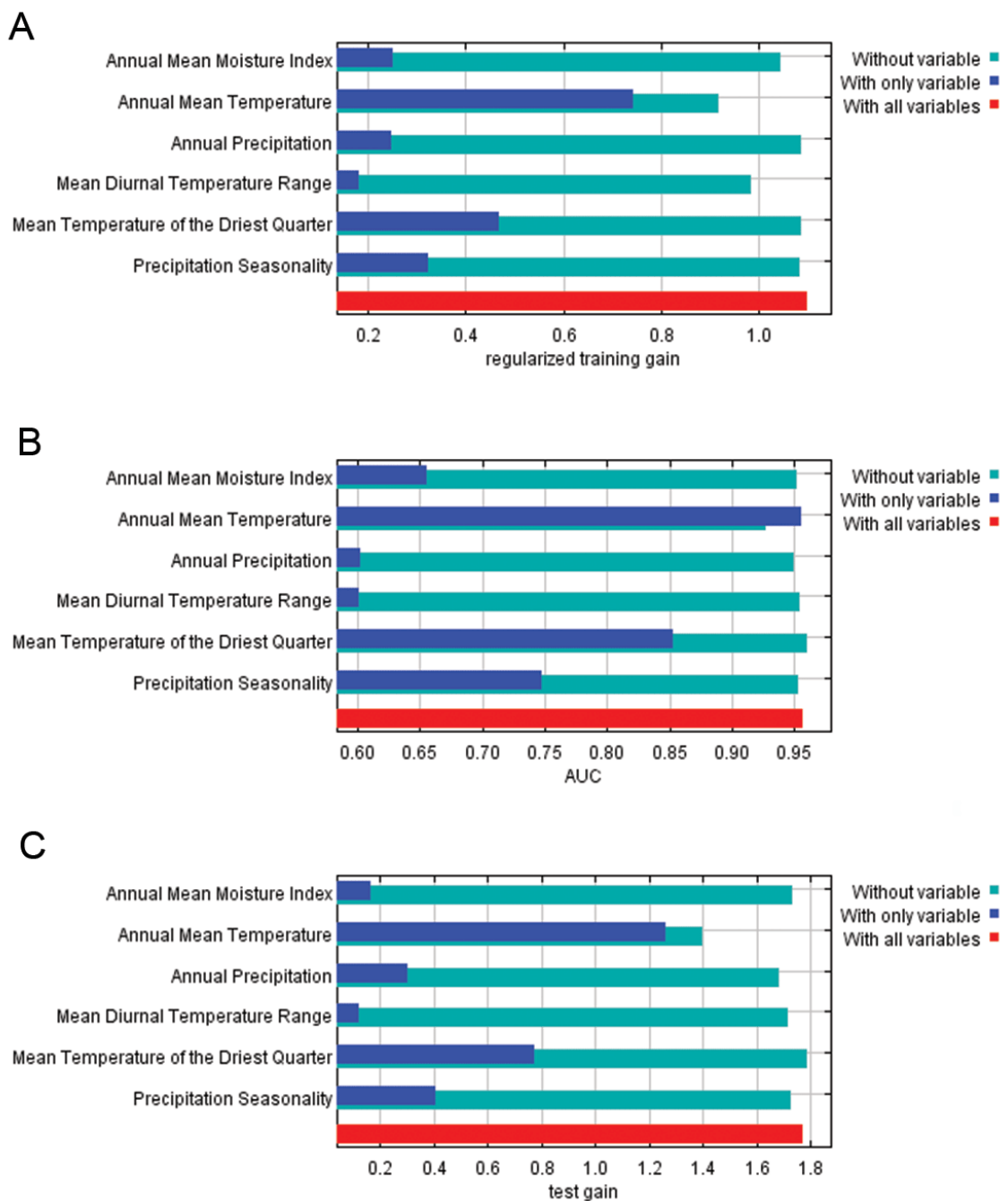


Figure 6. Jackknife estimates of the variable importance for the model predicting the distribution and potential suitable habitat of *Armillaria mellea* in eastern North America: A) regularized training gain; B) Area Under the Curve for training gain and; C) testing gain.

Response curves were produced for all variables that were included in the ENM. When predicting the distribution of *A. mellea* in eastern North America, cells containing high annual moisture index values often corresponded with areas of higher predicted occurrence (Fig. 7A). The occurrence of *A. mellea* was highest when mean annual temperature was 10°C (± 1 SD = 6-14°C) (Fig. 8B). The occurrence of *A. mellea* was higher when annual precipitation was < 400 mm per year and exhibited a decrease as precipitation increased to 1200 mm annually (Fig. 7C). Mean diurnal temperatures showed a higher occurrence of *A. mellea* as the difference between minimum and maximum annual temperatures increased (Fig. 7D). The mean temperature of the driest quarter variable was most predictive at -10°C and decreased as temperature increased (Fig. 7E). The occurrence of *A. mellea* was higher as the CV of precipitation decreased (Fig. 7F).

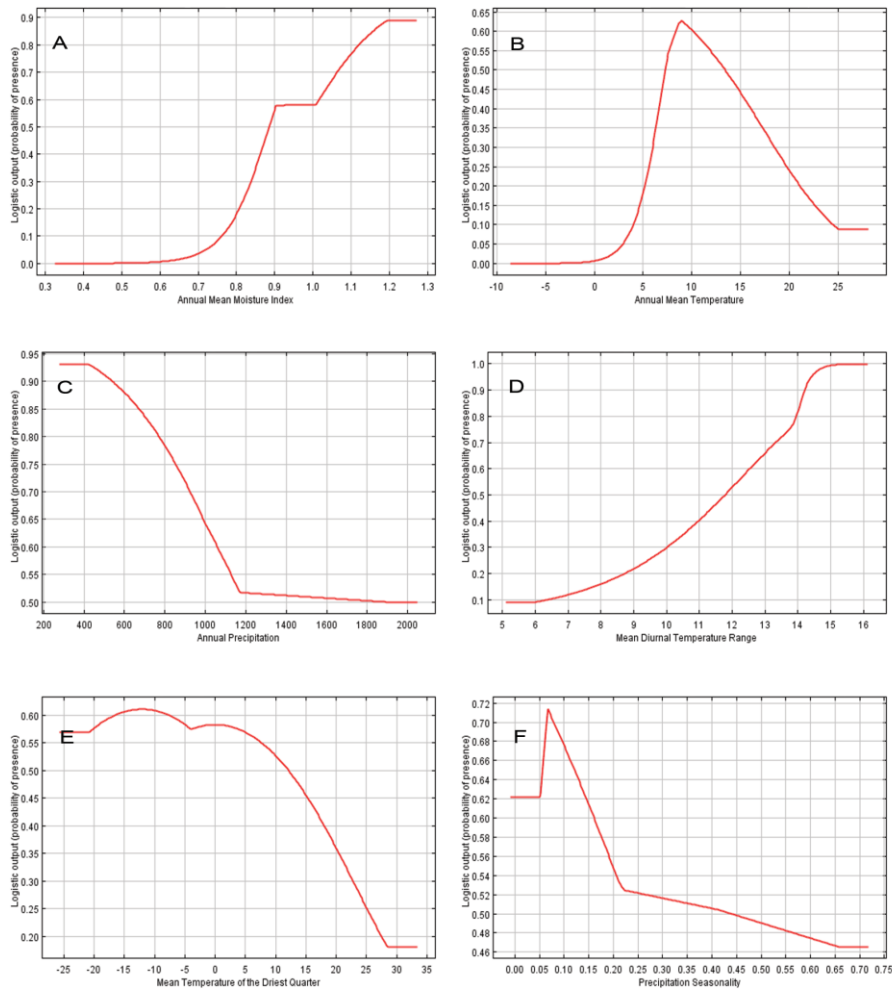


Figure 7. Response curves of each environmental variable in the best fitted Maxent model used to predict the distribution and potential suitable habitat of *Armillaria mellea* in eastern North America. Curves show logistic probability of occurrence of environmental variables used in the model: A) annual mean moisture index; B) annual mean temperature; C) annual precipitation; D) mean diurnal temperature range; E) mean temperature of the driest quarter and; F) precipitation seasonality.

CHAPTER 4. DISCUSSION

Several studies have examined the distribution of *A. mellea* in eastern North America at smaller geographic scales than this study (Banik et al., 1995; Baumgartner et al., 2010b; Bruhn et al., 2000; Harrington and Rizzo, 1993; McLaughlin 2001), and this is the first such study to accurately describe the distribution of the population in eastern North America and the annotation of historical and modern occurrence records to generate accurate distributions. Two of the variables used in the ENM explained > 50% of the ENM's prediction and were related to temperature, suggesting that temperature has significant control on the distribution of *A. mellea*. Five other variables were also important to predicting habitat suitability for *A. mellea* and an ecogeographical distribution based on ENM is provided for the first time.

Prior to this study, the distribution of *A. mellea* in eastern North America was described as predominantly southeastern with a decline in occurrence at northern latitudes and westward to the Great Plains (Burdshall and Volk, 1993). In this study, we observed that there is suitable habitat for *A. mellea* in the southeast, but areas with higher predicted probabilities of occurrences are limited to smaller spatial scales of mountainous habitats at higher elevations (e.g., Appalachian mountains). The northern geographic limit is defined by latitude and the colder temperatures of the northeastern US and Canada prevent the colonization and spread of *A. mellea*.

Armillaria species are broadly distributed in eastern North America, have overlapping ranges, and in most cases, similar macro-morphology (e.g. all species of *Armillaria* are annulate). As a consequence, when using historical data to predict species

distributions, the correct identification of voucher specimens is of critical importance. In this study, a small, but significant percentage of the vouchers obtained were inaccurately identified by the herbaria. This is likely a common scenario for many species, particularly among fungi, when either the lack of experience with a species or inadequate methods for proper identification are employed.

This study supports the concept that ENM based on confirmed, georeferenced, occurrence records is a suitable method to obtain accurate occurrence data for *A. mellea*. Collating the data from multiple sources (herbaria, field and samples obtained from culturing symptomatic tissues) was not insignificant, as it required many confirmed and georeferenced records to predict the occurrence of *A. mellea*. *Armillaria mellea* is often adequately represented in repositories (e.g. New York Botanical Garden William and Lynda Steere Herbarium contains 447 vouchers of *A. mellea*); however, in this case, many of the historical collections could not be used for this study since the success of molecular identification depends on age, condition of the specimen when collected, and the curation environment. Ideally, and as in the present case, species identification was confirmed from annotated vouchers with sequence data, which allowed for phylogenetic testing of species concepts in *Armillaria* using two, low-copy nuclear loci to resolve the evolutionary relationships for most of the specimens obtained. In addition, morphological data also served as an important source for occurrence data since a small number of vouchers were carefully examined using microscopic analysis.

Here, ENM were produced to predict the distribution and suitable habitat of *A. mellea* in eastern North America and generate the probability of occurrence areas of

habitat suitability. The ENM performed here extends the occurrence data amassed by surveys by predicting the fundamental niche of *A. mellea* in this region. At local scales, the resulting ENM produced in this study aligns closely with the occurrence of *A. mellea* obtained by extensive surveys of *Armillaria* species (Banik et al., 1995; Bruhn et al., 1997, 2000; Harrington and Rizzo, 1993; McLaughlin 2001).

The decline in the predicted probability of occurrence of *A. mellea* above 46°N, below 33°N, and westward beyond 97°W suggests that the abiotic conditions are unfavorable for establishment and growth for *A. mellea*. Banik et al. (1995) reported a decline in the occurrence of *A. mellea* compared to *A. gallica*, *A. calvescens*, and *A. ostoyae* (Banik et al., 1995; McLaughlin 2001) at higher latitudes in Wisconsin, Minnesota and Michigan. Similarly, McLaughlin (2001) reported a decline in the occurrence for *A. mellea* at the most northern latitudes in Quebec. Taken together, these results suggest that temperature is an important determinant of suitable habitat of *A. mellea* and climate limits the northern distribution.

In the southern extent of the distribution, *A. mellea* is often cited as a causal agent of oak decline (Marçais and Bréda, 2006; Thomas et al., 2002). Bruhn et al. (2000) determined that *A. mellea* was more frequent and exhibited higher levels of virulence on *Quercus* when compared to other hosts of *Armillaria*. Hardwoods dominate the forest covers in the 13 states in the southeastern US, with about two-thirds of the area in upland hardwood, bottomland hardwood or oak-pine mixtures (Haavik et al., 2015; Hodges, 1997; Keča et al., 2009) and many of the hosts in these forest types are susceptible to attack by *A. mellea* (Baumgartner et al., 2011; Raabe, 1962; Schnabel et al., 2005).

Despite suitable hosts in the region, a lower probability of occurrence was predicted for much of the ecogeographic region below 38°N. Higher water saturation and temperatures are found in the lower elevations of the Mississippi flood plain (Bailey, 1995) compared to the surrounding region and this may be an important factor reducing the probability of occurrence of *A. mellea*. Similarly, Mihail (2002) reported that soil saturation reduced the formation of *A. mellea* rhizomorphs. However, the mean temperature of the driest quarter was the predictor that was the highest importance for predicting the occurrence of *A. mellea*. In laboratory studies, the extent of mycelial growth of *A. mellea* was also conditioned by temperature (optimal growth occurs between 20-22°C) (Keča, 2005).

While ENM is a useful tool in this context for understanding the abiotic conditions that govern the distribution of *A. mellea*, it also has its limitations. The data used to predict *A. mellea* occurrence in eastern North America are limited by the potential influence of sampling bias. Spatial bias within the sampling area (sampling bias) may lead to false negatives, i.e. sampling points in which the species occurs but has not been sampled, which makes the geographic boundaries more difficult to discern. However, if effective spatial autocorrelation tests are performed, this helps to reduce this sampling bias by reducing the risk of false pseudo-absence data (Dormann et al., 2012; Lobo et al., 2008). Also, many of the collections used here were originally obtained for use in prior research (Baumgartner et al., 2010b, 2012; Bruhn et al., 1997, 2000; Hughes et al., 2013; McLaughlin 2001; Wargo 1983) and it is possible that over-representation of occurrence reduces the accuracy of the model. This problem was addressed by background manipulation and spatial filtering to reduce autocorrelation and, while sampling bias

introduces uncertainties to the boundaries of the model, it will not likely affect the identification of suitable habitat, if the sample size is sufficient for the scale of the study.

Another consideration that is not captured in estimating the fundamental niche of *A. mellea* by ENM analysis is the importance of biotic interactions in governing the suitable habitat of *Armillaria* species. For example, the prevalence of suitable habitat for *A. mellea* may decline due to interspecific or intraspecific competition with other *Armillaria* species that overlap in range and distribution. Bruhn et al. (2000) observed a decrease in the occurrence of *A. mellea* in the presence to *A. gallica* due to the latter's higher rhizomorph production, restricting much of the *A. mellea* distribution within the study area to where *A. gallica* is less prevalent. Guillaumin et al. (1993) suggested that *A. mellea* in western Europe was a better competitor in suitable habitats at low to mid-elevations when compared to other *Armillaria* species with overlapping geographic ranges. The results of the present study could be combined with similar ENM distributions of other *Armillaria* species to compare where biotic interactions may most influence *A. mellea* habitat suitability.

This study provides an annotated collection of vouchers and cultures for *A. mellea* distributed throughout eastern North America. The ENM applied here predicts ecogeographical distribution of suitable habitat for *A. mellea* in eastern North America and demonstrates how ENM can be used to obtain the realized niche of an important plant pathogen with a broad host distribution. Here, ENM models were consistently produced with a high level of accuracy based on model statistics and demonstrate that habitat suitability maintains the distribution of *A. mellea*. In future research, the

ecogeographical distribution may be refined and utilized for the purposes of understanding the fluctuations in the historical and modern geographic range of *A. mellea* and may be an important tool for assessing the severity of disease under future climate change.

REFERENCES

- Alvarado-Serrano, D.F., Knowles, L.L., 2014. Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources* **14**: 233–248.
- Bailey, R.G., 1995. *Description of the ecoregions of the United States*. US Department of Agriculture.
- Baldwin, R.A., 2009. Use of maximum entropy modeling in wildlife research. *Entropy* **11**: 854–866.
- Banik, M., Paul, J.A., Burdsall, H.H., 1995. Identification of *Armillaria* species from Wisconsin and adjacent areas. *Mycologia* **87**: 707–712.
- Baumgartner, K., Rizzo, D.M., 2001a. Ecology of *Armillaria* spp. in mixed-hardwood forests of California. *Plant disease* **85**: 947–951.
- Baumgartner, K., Rizzo, D.M., 2001b. Distribution of *Armillaria* Species in California. *Mycologia* **93**: 821–830.
- Baumgartner, K., Bhat, R., Fujiyoshi, P., 2010a. A rapid infection assay for *Armillaria* and real-time PCR quantitation of the fungal biomass in planta. *Fungal Biology* **114**: 107–119.
- Baumgartner, K., Travadon, R., Bruhn, J., Bergemann, S.E., 2010b. Contrasting patterns of genetic diversity and population structure of *Armillaria mellea* sensu stricto in the eastern and western United States. *Phytopathology* **100**: 708–718.
- Baumgartner, K., Coetzee, M.P.A., Hoffmeister, D., 2011. Secrets of the subterranean pathosystem of *Armillaria*. *Molecular Plant Pathology* **12**: 515–534.
- Baumgartner, K., Baker, B.R., Korhonen, K., Zhao, J., Hughes, K.W., Bruhn, J., Bowman, T.S., Bergemann, S.E., 2012. Evidence of natural hybridization among homothallic members of the basidiomycete *Armillaria mellea* sensu stricto. *Fungal Biology* **116**: 677–691.
- Bruhn, J.N., Brenneman, J.A., James Jr, J., Mihail, J.D. and Leininger, T.D., 1997. Spatial patterns of *Armillaria* populations in the walker branch watershed throughfall displacement experiment, Tennessee, USA. In *In: Pallardy, Stephen G.; Cecich, Robert A.; Garrett, H. Gene; Johnson, Paul S., eds. Proceedings of the 11th Central Hardwood Forest Conference; Gen. Tech. Rep. NC-188. St. Paul, MN: US*

Department of Agriculture, Forest Service, North Central Forest Experiment Station: 49-57 (Vol. 188).

- Bruhn, J.N., Wetteroff, J., Mihail, J.D., Kabrick, J., Pickens, J.B., 2000. Distribution of *Armillaria* species in upland Ozark Mountain forests with respect to site, overstory species composition and oak decline. *Forest Pathology* **30**: 43–60.
- Burdsall, H.H., Volk, T.J., 1993. The state of taxonomy of the genus *Armillaria*. *McIlvainea* **11**: 4–12.
- Canhos, V.P., Souza, S. De, Giovanni, R. De, Canhos, D.A.L., 2004. Global Biodiversity Informatics: setting the scene for a “new world” of ecological forecasting. *Biodiversity Informatics* **1**: 1–13.
- Chillali, M., Idder-Ighili, H., Guillaumin, J.J., Mohammed, C., Escarmant, B.L., Botton, B., 1998. Variation in the ITS and IGS regions of ribosomal DNA among the biological species of European *Armillaria*. *Mycological Research* **102**: 530–540.
- Coetzee, M.P.A., Bloomer, P., Wingfield, M.J., Wingfield, B.D., 2011. Paleogene radiation of a plant pathogenic mushroom. *PLoS ONE* **6**: e28545.
- Dormann, C.F., McPherson, J. M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**: 609–628.
- Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., Schröder, B., Singer, A., 2012. Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography* **39**: 2119–2131.
- Duchesne, L.C., Anderson, J.B., 1990. Location and direction of transcription of the 5S rRNA gene in *Armillaria*. *Mycological Research* **94**: 266–269.
- Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.
- Elías-Román, R.D., Medel-Ortiz, R., Alvarado-Rosales, D., Hanna, J.W., Ross-Davis, A.L., Kim, M.-S., Klopfenstein, N.B., 2018. *Armillaria mexicana*, a newly described species from Mexico. *Mycologia* **110**: 347–360.

- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**: 43–57.
- Ferguson, B.A., Dreisbach, T.A., Parks, C.G., Filip, G.M., Schmitt, C.L., 2003. Coarse-scale population structure of pathogenic *Armillaria* species in a mixed-conifer forest in the Blue Mountains of northeast Oregon. *Canadian Journal of Forest Research* **33**: 612–623.
- Guillaumin, J.J., Mohammed, C., Anselmi, N., Courtecuisse, R., Gregory, S.C., Holdenrieder, O., Intini, M., Lung, B., Marxmüller, H., Morrison, D., Rishbeth, J., Termorshuizen, A.J., Tirrón, A., Dam, B., 1993. Geographical distribution and ecology of the *Armillaria* species in western Europe. *Forest Pathology* **23**: 321–341.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147–186.
- Haavik, L.J., Billings, S.A., Guldin, J.M., Stephen, F.M., 2015. Emergent insects, pathogens and drought shape changing patterns in oak decline in North America and Europe. *Forest Ecology and Management* **354**: 190–205.
- Harrington, T.C., Rizzo, D.M., 1993. Identification of *Armillaria* species from New Hampshire. *Mycologia* **85**: 365–368.
- Harrington, T.C., Wingfield, B.D., 1995. A PCR-based identification method for species of *Armillaria*. *Mycologia* **87**: 280–288.
- Hodges, J.D., 1997. Development and ecology of bottomland hardwood sites. *Forest Ecology and Management* **90**: 117–125.
- Hughes, K.W., Petersen, R.H., Lodge, D.J., Bergemann, S.E., Baumgartner, K., Tulloss, R.E., Lickey, E., Cifuentes, J., 2013. Evolutionary consequences of putative intra- and interspecific hybridization in agaric fungi. *Mycologia* **105**: 1577–1594.
- Jiménez-Valverde, A., Lobo, J.M., 2006. The ghost of unbalanced species distribution data in geographical model predictions. *Diversity and Distributions* **12**: 521–524.
- Keča, N., 2005. Characteristics of *Armillaria* species development and their growth at different temperatures. *Glasnik Sumarskog fakulteta* **91**: 149–162.
- Kluting, K.L., Baroni, T.J., Bergemann, S.E., 2014. Toward a stable classification of genera within the Entolomataceae : a phylogenetic re-evaluation of the Rhodocybe-Clitopilus clade. *Mycologia* **106**: 1127–1142.

- Koch, R.A., Wilson, A.W., Séné, O., Henkel, T.W. and Aime, M.C., 2017. Resolved phylogeny and biogeography of the root pathogen *Armillaria* and its gasteroid relative, *Guyanagaster*. *BMC evolutionary biology*, **17**: 1-16.
- Korhonen, K., 1980. The origin of clamped and clampless basidia in *Armillariella ostoyae*. *Karstenia* **20**: 23–27.
- Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J., Scott, J.K., 2012. CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution* **3**: 53–64.
- Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17**: 145–151.
- Loiselle, B.A., Jørgensen, P.M., Consiglio, T., Jiménez, I., Blake, J.G., Lohmann, L.G. and Montiel, O.M., 2008. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *Journal of Biogeography* **35**: 105–116.
- Maddison, W.P., Maddison, D.R., 2018. Mesquite: a modular system for evolutionary analysis. Version 3.40.
- Maphosa, L., Wingfield, B.D., Coetzee, M.P.A., Mwenje, E., Wingfield, M.J., 2006. Phylogenetic relationships among *Armillaria* species inferred from partial elongation factor 1-alpha DNA sequence data. *Australasian Plant Pathology* **35**: 513–520.
- Marçais, B., Bréda, N., 2006. Role of an opportunistic pathogen in the decline of stressed oak trees. *Journal of Ecology* **94**: 1214–1223.
- McLaughlin, J.A., 2001. Distribution, hosts, and site relationships of *Armillaria* spp. in central and southern Ontario. *Canadian Journal of Forest Research* **31**: 1481-1490.
- Mihail, J., 2002. The effects of moisture and oxygen availability on rhizomorph generation by *Armillaria tabescens* in comparison with *A. gallica* and *A. mellea*. *Mycological Research* **106**: 697–704.
- Miller, M.A., Pfeiffer, W. and Schwartz, T., 2010, November. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE), 2010* (pp. 1-8).

- Motta, J., Korhonen, K., 1986. A note on *Armillaria mellea* and *Armillaria bulbosa* from the middle Atlantic states. *Mycologia* **78**: 471–474.
- Pearce, J.L., Boyce, M.S., 2006. Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology* **43**: 405–412.
- Pearson, R.G., Raxworthy, C.J., Nakamura, M., Townsend Peterson, A., 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* **34**: 102–117.
- Peterson, A.T., 2001. Predicting species' geographic distributions based on ecological niche modeling. *The Condor* **103**: 599–605.
- Peterson, A.T., Nakazawa, Y., 2008. Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography* **17**: 135–144.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species distributions. *Ecological Modelling* **190**: 231–259.
- Phillips, S.J., Dudík, M., Schapire, R.E. 2017. Maxent software for modeling species niches and distributions (Version 3.4.1).
- Pineda, E., Lobo, J., 2009. Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *Journal of Animal Ecology* **78**: 182–190.
- Ponce, R.A., Águeda, B., Ágreda, T., Modrego, M.P., Aldea, J., Fernández-toirán, L.M., Martínez-peña, F., 2011. Rockroses and *Boletus edulis* ectomycorrhizal association: realized niche and climatic suitability in Spain. *Fungal Ecology* **4**: 224–232.
- Prospero, S., Holdenrieder, O., Rigling, D., 2006. Rhizomorph production and stump colonization by co-occurring *Armillaria cepistipes* and *Armillaria ostoyae*: an experimental study. *Forest Pathology* **36**: 21–31.
- Raabe, R.D., 1962. Host list of the root rot fungus *Armillaria mellea*. *Hilgardia* **33**: 24–89.
- Rambaut, A., 2012. FigTree v1.4.0. <http://tree.bio.ed.ac.uk/software/figtree/>
- Schnabel, G., Ash, J.S., Bryson, P.K., 2005. Identification and characterization of *Armillaria tabescens* from the southeastern United States. *Mycological Research* **109**: 1208–1222.

- Segurado, P., Araújo, M.B., Kunin, W.E., 2006. Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology* **43**: 433–444.
- Sato, H., Tsujino, R., Kurita, K., Yokoyama, K., Agata, K., 2012. Modelling the global distribution of fungal species: new insights into microbial cosmopolitanism. *Molecular Ecology* **21**: 5599–5612.
- Smith, B.E., Johnston, M.K., Lücking, R., 2016. From GenBank to GBIF: Phylogeny-based predictive niche modeling tests accuracy of taxonomic identifications in large occurrence data repositories. *PLoS ONE* **11**: e0151232.
- Smith, D.C., Simon, M., Alldredge, A.L., Azam, F., 1992. The fungus *Armillaria bulbosa* is among the largest and oldest living organisms. *Nature* **359**: 139–141.
- Soberon, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* **2**: 1–10.
- Stamatakis, A., 2006. RAxML-VI-HPC : maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.
- Stanke, M., Steinkamp, R., Waack, S., Morgenstern, B., 2004. AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Research* **32**: 309–312.
- Thomas, F.M., Blank, R., Hartmann, G., 2002. Abiotic and biotic factors and their interactions as causes of oak decline in Central Europe. *Forest Pathology* **32**: 277–307.
- Travadon, R., Smith, M.E., Fujiyoshi, P., Douhan, G.W., Rizzo, D.M., Baumgartner, K., 2012. Inferring dispersal patterns of the generalist root fungus *Armillaria mellea*. *New Phytologist* **193**: 959–969.
- Tsykun, T., Rigling, D., Prospero, S., 2013. A new multilocus approach for a reliable DNA-based identification of *Armillaria* species. *Mycologia* **105**: 1059–1076.
- Ullrich, R.C., Anderson, J.B., 1978. Sex and diploidy in *Armillaria mellea*. *Experimental Mycology* **2**: 119–129.
- van Gils, H., Westinga, E., Carafa, M., Antonucci, A., Ciaschetti, G., 2014. Where the bears roam in Majella National Park, Italy. *Journal for Nature Conservation* **22**: 23–34.
- Veldman, G.M., Klootwijk, J., de Regt, V.C., Planta, R.J., Branlant, C., Krol, A., Ebel, J.P., 1981. The primary and secondary structure of yeast 26S rRNA. *Nucleic Acids Research* **9**: 6935–6952.

- Wargo, P.M., 1983. *Armillaria mellea* and mortality of beech affected by beech bark disease. In *In: Proceedings, IUFRO Beech Bark Disease Working Party Conference; 1982 September 26-October 8; Hamden, CT. Sponsored by the USDA Forest Service, Northeastern Forest Experiment Station. Gen. Tech. Rep. WO-37.[Washington, DC]: US Department of Agriculture, Forest Service: 81-88. (37: 81-88).*
- Warren, D., Seifert, S.N., 2011. Ecological niche modeling in Maxent : the importance of model complexity and the performance of model selection criteria. *Ecological Applications* **21**: 335–342.
- Watling, R., Kile, G.A., Gregory, N.M., 1982. The genus *Armillaria*-nomenclature, typification, the identity of *Armillaria mellea* and species differentiation. *Transactions of the British Mycological Society* **78**: 271–285.
- White, A.E.E., Dubetz, C.P., Cruickshank, M.G., Morrison, D.J., Feb, N.J., 1998. DNA diagnostic for *Armillaria* species in British Columbia : within and between species variation in the IGS-1 and IGS-2 regions. *Mycologia* **90**: 125–131.
- Wolfe, B.E., Richard, F., Cross, H.B., Pringle, A., 2010. Distribution and abundance of the introduced ectomycorrhizal fungus *Amanita phalloides* in North America. *New Phytologist* **185**: 803–816.
- Wollan, A.K., Bakkestuen, V., Kausrud, H., Gulden, G., Halvorsen, R., 2008. Modelling and predicting fungal distribution patterns using herbarium data. *Journal of Biogeography* **35**: 2298–2310.

APPENDIX

Appendix. Specimens used to model the distribution and potential suitable habitat of *Armillaria mellea*.

Species	Collection Identifier/Gen Bank Identifier	Herbarium Accession No.	Herbarium Identifier	Collector(s), Location and Year	Host or substrate	Longitude (°W)	Latitude (°N)	Identification Method
<i>A. calvescens</i>	GP92	C0227792F	F	G. Perlow, Downers Grove, DuPage County, IL, USA, 1993	on ground	88.0153	41.7776	S ^a
<i>A. gallica</i>	JS13	DUKE192990	DUKE	J. Sogani, Craven County, NC, USA, 1990	<i>Quercus</i> sp.	77.0563	34.9143	S ^a
<i>A. mellea</i>	JJ007 ^b	NJJ16_038		J. Justice, St. Genevieve County, MO, USA, 2016		90.2324	37.8347	S ^a
<i>A. mellea</i>	JJ001 ^b			J. Justice, USA, Oconee County, SC, USA, 2016		90.2324	34.9667	RFLP ^c
<i>A. mellea</i>	JJ004 ^b	NJJ16_032		J. Justice, Faulkner County, AK, USA, 2016		92.27948	35.2901	RFLP ^c
<i>A. mellea</i>	JJ005 ^b	NJJ16_033		J. Justice, Faulkner County, AK, USA, 2016		92.27793	35.2906	RFLP ^c
<i>A. mellea</i>	JJ006 ^b	NJJ16_037		J. Justice, St. Genevieve County, MO, USA, 2016		90.2324	37.830	RFLP ^c
<i>A. mellea</i>	JJ008 ^b			J. Justice, Hardin County, TX, USA, 2016		94.515	30.2658	RFLP ^c
<i>A. mellea</i>	115 ^e			D. Mclaughlin, Grattan Township, Pembroke District, Ontario, Canada, 1999	<i>Q. rubra</i>	76.9941	45.7280	S ^a
<i>A. mellea</i>	232 ^e			D. Mclaughlin, Dalhousie Township, Lanark County, Ontario, Canada, 1999	<i>Fagus grandifolia</i>	76.4726	45.0443	S ^a
<i>A. mellea</i>	253 ^e			D. Mclaughlin, Tweed Township, Hastings County, Ontario, Canada, 1999	<i>Betula papyrifera</i>	77.3211	44.2925	S ^a
<i>A. mellea</i>	323 ^e			D. Mclaughlin, Regional Municipality of Niagara, Ontario, Canada, 1999	<i>F. grandifolia</i>	79.3726	43.1414	S ^a

<i>A. mellea</i> ^a	LJ001 ^a			L. Jarnigan, Johnston Shut-ins State Park, Reynolds County, MO, USA, 2016		90.88493	37.5595	S ^a
<i>A. mellea</i>	JJ002 ^a			J. Justice, Oconee Co., SC, USA, 2016		83.12509	34.9667	S ^a
<i>A. mellea</i>	JJ003 ^a	NJJ16_031		J. Justice, Faulkner County, AK, USA, 2016		92.27965	35.2894	S ^a
<i>A. mellea</i>	BAK137 ^a			B. Kerr, Cumberland Mountain State Park, Cumberland County, TN, USA 2016	hardwood	84.9975	35.9014	S ^a
<i>A. mellea</i>	Ovrebo-3230	3330	CFMR	Leflore, OK, USA, 1991	hardwood	94.69619	34.5717	RFLP ^c
<i>A. mellea</i>	HHB-12555	3265	CFMR	HH Burdsall Jr, Sauk County, WI, USA 1988	<i>Q. vetulina</i>	89.6388	43.5555	S ^a
<i>A. mellea</i>	OZ-604 ^f	3229	CFMR	Carter County, MO, USA, 1995	<i>Q. vetulina</i>	81.2362	42.9652	S ^a
<i>A. mellea</i>	OZ-623 ^f		CFMR	Carter County, MO, USA, 1995	<i>Q. vetulina</i>	91.9623	36.9412	S ^a
<i>A. mellea</i>	OZ-1330 ^f		CFMR	Shannon County, MO, USA, 2003	<i>Q. alba</i>	91.1328	37.1542	S ^a
<i>A. mellea</i>	OZ-1592 ^f		CFMR	Shannon County, MO, USA, 2005	<i>Q. stellata</i>	91.0274	37.1706	S ^a
<i>A. mellea</i>	OKM-3704	3324	CFMR	Beltsville, Prince George's County, MD, USA, 1966	<i>Quercus</i> sp.	76.9177	39.0375	S ^a
<i>A. mellea</i>	OKM-4700	3328	CFMR	Bowie, Prince George's County, MD, USA, 1966	hardwood	76.9177	39.0375	S ^a
<i>A. mellea</i>	JMS-339	3295	CFMR	Livonia, Centre County, PA, USA	<i>Q. coccinea</i>	77.17	40.58	S ^a
<i>A. mellea</i>	Wargo-842	3457	CFMR	P. Wargo, Tuscarora State Forest, Carlisle County, PA, USA, 1986	<i>Q. prinus</i>	77.55	40.36	S ^a
<i>A. mellea</i>	OR-138		CFMR	Oak Ridge, Roane County, TN, USA, 1995		84.2192	35.9523	S ^a
<i>A. mellea</i>	DU-1-T	3198	CFMR	Elk Mound, Dunn County, WI, USA, 1990	<i>Acer saccharum</i> and <i>Betula</i> sp.	89.30845	43.4958	S ^a
<i>A. mellea</i>	FP-101989-Sp	3203	CFMR	One mile south of Wyocena, WI, USA, 1984	<i>Quercus</i> sp.	89.30845	43.4958	S ^a

<i>A. mellea</i>	FP-102028-T	3206	CFMR	Hancock Experimental Station, Hancock, WI, USA, 1984	<i>Q. macrocarpa</i>	89.52318	44.1335	S ^a
<i>A. mellea</i>	Mad-703-Sp	3302	CFMR	Wisconsin Rapids, Wood County, WI, USA	<i>Quercus</i> sp.	89.81735	44.3835	S ^a
<i>A. mellea</i>	MI-3(T-1)	3306	CFMR	Baraga County, MI, USA, 1991	<i>Quercus</i> sp.	88.48901	46.7785	S ^a
<i>A. mellea</i>	OH-2-T	3314	CFMR	Barrancas, Lucas County, OH, USA, 1991		83.61269	41.6621	S ^a
<i>A. mellea</i>	SC00i147		CFMR	Spartainburg County, SC USA, 1987	<i>Prunus persica</i>	81.95777	35.1497	S ^a
<i>A. mellea</i>	SCNP04		CFMR	Anderson County, SC, USA, 2004	<i>Quercus</i> sp.	82.83736	34.6834	S ^a
<i>A. mellea</i>	FP-104077-Sp	3210	CFMR	Duval County, FL, USA, 1954	<i>Castanea alnifolia</i>	81.47877	30.4158	S ^a
<i>A. mellea</i>	HHB-138-Sp	3270	CFMR	Laurel County, MD, USA, 1967	<i>Quercus</i> sp.	76.90747	39.0348	S ^a
<i>A. mellea</i>	TJVIL-40-T	2930	CFMR	T. Volk, Jo Daviess County, IL, USA, 1992	<i>Quercu</i> sp.	90.27957	42.2564	S ^a
<i>A. mellea</i>	LC314	2957	CFMR	Lanphear-Cook, Juneau, Necedah, WI, USA, 1989	<i>Q. palustris</i>	90.07402	44.0260	S ^a
<i>A. mellea</i>	LC355	2958	CFMR	Lanphear-Cook, Portage County, WI, USA, 1989	hardwood	89.27345	44.5747	S ^a
<i>A. mellea</i>	2263TJB	CORT0054 94	CORT	TJ Baroni, Hampshire County, MA, USA, 1975	<i>Quercus</i> sp.	72.51278	42.3900	RFLP ^c
<i>A. mellea</i>	2639TJB	CORT0054 92	CORT	TJ Baroni, Conway State Forest, Franklin County, MA, USA, 1976		72.7108	42.4674	S ^a
<i>A. mellea</i>	6483TJB	CORT0054 93	CORT	TJ Baroni, Tompkins Co., Ringwood Preserve, NY, USA, 1990	<i>F. grandifolia</i>	76.1561	42.4501	S ^a
<i>A. mellea</i>	8868MC	CORT0054 79	CORT	Tompkins Co., Ringwood Preserve, NY, USA, 1998	hardwoo	76.1561	42.4501	Sa
<i>A. mellea</i>	6540TJB	CORT0054 81	CORT	T. Baroni, Onondago, Rand Tract, NY, USA, 1991	<i>Q. rubra</i>	76.1561	42.9947	S ^a
<i>A. mellea</i>	DMWV00- 867	DEWV-F- 00550	DEWV	D. Mitchell, Pine Crest Cemetery, Pendleton County, WV, USA, 2000	<i>Quercus</i> sp.	79.3128	38.6648	S ^a

<i>A. mellea</i>	DPWV00-1044	DEWV-F-000601	DEWV	I. Phares, Walton League, Randolph County, WV, USA, 2000	<i>Quercus</i> sp.	79.8338	38.7957	S ^a
<i>A. mellea</i>	WRWV00-735	DEWV-F-000814	DEWV	W. Roody, Blue Bend Recreation Area, Greenbrier County, WV, USA, 2000	mixed hardwoods	80.2724	37.9181	S ^a
<i>A. mellea</i>	DMWV01-202	DEWV-F-001508	DEWV	D. Mitchell, Thorn Creek WMA, Pendleton County, WV, USA, 2001	hardwood	81.1331	37.5052	S ^a
<i>A. mellea</i>	WRWV04-975	DEWV-F-005308	DEWV	W. Roody, Moncove Lake State Park, Monroe, WV, USA, 2004	<i>Quercus</i> sp., <i>F. grandifolia</i>	80.3549	37.6157	S ^a
<i>A. mellea</i>	DMWV01-110	DEWV-F-001464	DEWV	D. Mitchell, Mercer County, WV, USA, 2001	<i>Acer</i> sp., <i>F. grandifolia</i>	82.3848	37.9888	S ^a
<i>A. mellea</i>	WRWV01-1338	DEWV-F-002391	DEWV	W. Roody, Wayne County, WV, USA, 2001	<i>Liquidambar styraciflua</i> , <i>P. strobus</i>	82.3518	37.9742	S ^a
<i>A. mellea</i>	WRWV02-542B	DEWV-F-003343	DEWV	W. Roody, Kanawha County, WV, USA, 2002	<i>Quercus</i> sp., <i>Pinus</i> sp., <i>Carya</i> sp., <i>Carpinus americanus</i>	81.5733	38.3852	S ^a
<i>A. mellea</i>	WRWV02-813	DEWV-F-003535	DEWV	W. Roody, Sleepy Creek, Berkeley County, WV, USA, 2002	mixed <i>Quercus</i> sp., <i>Pinus</i> sp., <i>Carya</i> sp.	78.1678	39.6706	S ^a
<i>A. mellea</i>	WRWV03-1076	DEWV-F-004825	DEWV	W. Roody, Monongalia County, WV, USA, 2003	<i>Quercus</i> sp.	81.2220	37.3678	S ^a
<i>A. mellea</i>	WRWV04-1211	DEWV-F-006504	DEWV	W. Roody, Wood County, WV, USA, 2004	<i>Quercus</i> sp.	81.2931	39.2407	S ^a
<i>A. mellea</i>	CSWV06-903	DEWV-F-008680	DEWV	C. Stihler, Elk River Wildlife Management Area, Braxton County, WV, USA, 2006	hardwood	80.6433	38.6231	S ^a
<i>A. mellea</i>	WRWV06-928	DEWV-F-008705	DEWV	W. Roody, Terra Alt Camp, Preston County, WV, USA, 2006	hardwood	79.5277	39.4545	S ^a

<i>A. mellea</i>	WRWV10-105A	DEWV-F-010306	DEWV	W. Roody, French Creek Game Farm, Upshur County, WV, USA, 2010	<i>Quercus</i> sp.	80.3213	38.8738	S ^a
<i>A. mellea</i>	WRWV02-793	DEWV-F-003695	DEWV	W. Roody, Hardy County, WV, USA, 2002	<i>Quercus</i> sp.	78.92125	38.8950	S ^a
<i>A. mellea</i>	JJ15	DUKE 0042215	DUKE	J. Johnson, Buncombe County, NC, USA, 1993		82.6273	35.4854	S ^a
<i>A. mellea</i>	Guan17	DUKE 0193473	DUKE	Guan, Beaufort County, NC, USA, 2009	hardwood	76.6473	34.7259	S ^a
<i>A. mellea</i>	JDG12	DUKE 0192947	DUKE	J. Green, Buncombe County, NC, USA, 2009	hardwood	76.6473	34.7259	S ^a
<i>A. mellea</i>	RP17	DUKE 0193475	DUKE	R. Palmer, Durham, Eno River State Park, NC, USA 2009	hardwood	79.0069	36.0780	S ^a
<i>A. mellea</i>	CHC17	DUKE 0350840	DUKE	C. Cannon, Raven Rock State Park, Harnett County, NC, USA, 1995	<i>Quercus</i> sp.	78.9048	35.4610	S ^a
<i>A. mellea</i>	CKG17	DUKE 0348589	DUKE	CKG, Duke Forest, Durham County, NC, USA, 1996		78.9830	36.0200	S ^a
<i>A. mellea</i>	MDP362	C0213449F	F	P. County, IN, USA, 1998	<i>Quercus</i> sp.	87.0897	41.6458	RFLP ^c
<i>A. mellea</i>	SAB352	C0213485F	F	S. Berge, Cook County, IL, USA, 1997	<i>Quercus</i> sp.	87.8644	41.6761	S ^a
<i>A. mellea</i>	JPS782	C0213459F	F	J. Schmit, Indiana Dunes National Lakeshore, Porter County, IN, USA, 1994	hardwood	87.0423	41.6657	S ^a
<i>A. mellea</i>	DR140	C0213524F	F	D. Richter, North Canal Entry Park, Houghton County, MI, USA, 1985	<i>Alnus serrulata</i>	88.6338	47.2285	S ^a
<i>A. mellea</i>	WCG2503	C0257792F	F	W. Gaswick, Warren Woods, Berrien County, MI, USA, 2011	<i>Liriodendron tulipifera</i>	86.6306	41.8395	S ^a
<i>A. mellea</i>	IMA9132	C0257793F	F	J. Ware, Volo Bog, Lake County, IL, USA, 1994		88.1911	42.3545	S ^a
<i>A. mellea</i>	NAMA2013-010	C0257784F	F	Ozark National Forest, Baxter County, AK, USA, 2013	hardwood	92.353	36.0298	S ^a
<i>A. mellea</i>	JM2614	C0213518F	F	J.M. Murphy, Manistee County, MI, USA, 1995	<i>Pinus strobus</i>	86.1066	44.2908	S ^a

<i>A. mellea</i>	PRL10404	C0257786F	F	P. Leacock, Wonder Lake, McHenry County, IL, USA, 2012	<i>Q. alba</i>	88.3699	42.4209	S ^a
<i>A. mellea</i>	PRL8618	C0257791F	F	P. Leacock, Big Rock Visitor Station, DuPage County, IL, USA, 2010	<i>Quercus</i> sp.	88.045	41.8183	S ^a
<i>A. mellea</i>	PRL10432	C0257788F	F	P. Leacock, Glacial Park, McHenry County, IL, USA, 2012	<i>Quercus</i> sp.	88.327	42.42	S ^a
<i>A. mellea</i>	PRL3268	C0213481F	F	P. Leacock, Palos Park, Cook County, IL, USA, 1998	<i>Quercus</i> sp.	87.8641	41.6758	S ^a
<i>A. mellea</i>	PRL10471	C0257790F	F	P. Leacock, Lake County, IL, USA, 2012	<i>Q. alba</i>	88.1946	42.2291	S ^a
<i>A. mellea</i>	JPS1322	C0213504F	F	Cook County, IL, USA, 1996	<i>Quercus</i> sp.	87.8601	41.6681	S ^a
<i>A. mellea</i>	MDP538	C0213469F	F	Cook County, IL, USA, 1998	<i>Quercus</i> sp.	87.8644	41.6761	S ^a
<i>A. mellea</i>	N-5178	C0213507F	F	Singer County, IL, USA 1984	<i>Quercus</i> sp.	88.4333	41.95	S ^a
<i>A. mellea</i>	WBC44787	MU000150	MU	Union County, OH, USA, 1971	hardwood	83.28514	40.2026	M ^d
<i>A. mellea</i>	WBC49618	MU000150	MU	Greene County, OH, USA, 1974		83.85531	39.7901	M ^d
<i>A. mellea</i>	WBC57976	MU000150	MU	Hamilton County, OH, USA, 1979		84.39909	39.0736	M ^d
<i>A. mellea</i>	WBC49553	MU000150	MU	Hocking County, OH, USA, 1974		82.57658	39.4575	M ^d
<i>A. mellea</i>	WBC44982	MU000150	MU	Tuscarawas County, OH, USA, 1971		81.55762	40.6370	M ^d
<i>A. mellea</i>	WBC45112	MU000150	MU	Adams County, OH, USA, 1971	<i>Quercus</i> sp.	83.40699	38.7593	M ^d
<i>A. mellea</i>	WBC51700	MU000150	MU	Pike County, OH, USA, 1975	hardwood	83.21717	39.1599	M ^d
<i>A. mellea</i>	TL024	MU000150	MU	Butler County, OH, USA, 1992	hardwood	84.72619	39.5125	M ^d
<i>A. mellea</i>	MAV6434	MU000150	MU	Columbiana County, OH, USA 1993		80.58598	40.7106	M ^d
<i>A. mellea</i>	MAV3492	MU000150	MU	Lake County, OH, USA, 1989		81.33682	41.6105	M ^d
<i>A. mellea</i>	WBC46015	MU000150	MU	Scioto Trail State Forest, Ross County, OH, USA, 1972	hardwood	83.4063	39.2220	RFLP ^c

<i>A. mellea</i>	WBC41521	MU000150 856	MU	Hueston Woods, Preble County, OH, USA, 1969	<i>A. nigrum</i>	82.93014	39.5698	RFLP ^c
<i>A. mellea</i>	WBC47093	MU000150 834	MU	The Seven Caves, Highland County, OH, USA, 1972	hardwood	83.35516	39.2228	RFLP ^c
<i>A. mellea</i>	13439	TENN-F- 061702	TENN	Cataloochee Cove, Haywood County, TN, USA, 2006		83.1208	35.6163	S ^a
<i>A. mellea</i>	12071	TENN-F- 060319	TENN	Cades Cove, Blount County, TN, USA, 2004		83.8136	35.6022	S ^a
<i>A. mellea</i>	12931	TENN-F- 061396	TENN	Cades Cove, Blount County, TN, USA, 2005		83.7825	35.6105	S ^a
<i>A. mellea</i>	GM909	TENN-F- 044687	TENN	Oconee, TN, USA		82.9930	34.7491	S ^a
<i>A. mellea</i>	OKM-2636	VPI-F- 0000497	VPI	O. Miller, VA, Giles, Mt. Lake Biol. Station, 1995	<i>Quercus</i> sp.	80.5232	37.3752	S ^a
<i>A. mellea</i>	OKM-8160	VPI-F- 0000496	VPI	O. Miller, Prince Georges County, MD, USA, 1969	hardwood	76.8205	38.9876	S ^a
<i>A. mellea</i>	OKM-9281	VPI-F- 0000501	VPI	O. Miller, Craig County, VA, USA	<i>Q. alba</i> , <i>Q. rubra</i>	80.4480	37.3397	S ^a
<i>A. mellea</i>	TFW-12508	VPI-F- 0000488	VPI	T. Wiedboldt, Giles County, VA, USA	hardwood	80.5505	37.4318	S ^a
<i>A. mellea</i>	OKM-17896	VPI-F- 0000494	VPI	Roanoke County, VA, USA,	hardwood	79.9492	37.3886	S ^a
<i>A. ostoyae</i>	#86-T	3161	CFMR	Houghton County, MI, USA, 1985		88.44167	47.17	S ^a

^aSamples were delimited through phylogenetic analyses of the *tef1* and *actin-1* sequences.

^bVouchers collected for the duration of this study and currently at Middle Tennessee State University, Murfreesboro, Tennessee, USA (MTSU).

^cVouchers were delimited with restriction fragment length polymorphisms of the nuclear ribosomal IGS1 region.

^dVouchers were delimited with morphological analyses.

^eRecords georeferenced from Mclaughlin et al., (2001).

^fRecords georeferenced from Bruhn et al., (2000).