

Phylogenetic pipeline: phylogenetic analysis of chosen species from *Armillaria* genus

Jakub J. Guzek

University of Warsaw, Faculty of Mathematics, Informatics and Mechanics, Genomika porównawcza (1000-719GP2),
Student ID: 456616

Fungi from basidiomycete genus *Armillaria* are responsible for root rot in woody plants in various geographic regions. They are characterized by wide host range, low host specificity and are present on all continents, except Antarctica. They are capable of thriving both in natural forests and on woody crops. *Armillaria* root rot – the disease caused by those fungi, can cause serious economic losses when it attacks industrial plantations or orchards. Fungi from this genus, have been the object of interest of many mycologists and plant pathologists for a long time, because of their pathogenicity but also because individuals of *Armillaria* species are among the oldest and largest discovered organisms on earth. This led to multitude of research papers on *Armillaria* physiology, ecology and phylogeny being published over the years. While relatively mature, fungal phylogeny remains a field, that could benefit from more rigorous and extensive analyses of readily available data. As such, most studies of *Armillaria* phylogeny focus on a handful of molecular markers, thus not fully utilizing annotated genomic assemblies. Here I present an analysis of phylogeny, utilizing full proteomic data from 8 species from *Armillaria* genus.

INTRODUCTION

Saprotrophic and necrotrophic fungi were shown to be the most efficient decomposers of lignocellulose – material complex consisting of cellulose, hemicelluloses, lignin and other structural macromolecules, which constitutes the majority of plant biomass[1]. Those wood-decaying fungi are essential agents of carbon cycling in forest ecosystems[2].

Agaricomycetes are the main group within Basidiomycota, responsible for decay of dead wood[3], with two main modes of dead wood decay that can be distinguished among them: white rot and brown rot. Brown rot fungi have the ability to decompose cellulose and hemicelluloses, but cannot fully degrade lignin. White rot fungi, on the other hand, can decompose cellulose, hemicelluloses and also fully degrade lignin.[4]

Wood decaying fungi in Agaricomycetes encompass species with very different lifestyles, from mycorrhizal symbiosis, through saprotrophy and necrotrophy to parasitism[3]. Many fungi are not easily characterized by single lifestyle, and can be mycorrhizal in some circumstances and pathogenic in others. Most of plants on Earth depend on symbiosis with fungi[5] and some plants are even able to be parasites of fungi[6–8].

Although it's important to underline the complexity of fungi-plant relationships, which are often impossible to put into simple, discrete categories, this report will use this mentioned classification for simplicity.

While enzymes encoded in genomes of white rot and brown rot fungi allow for decomposition of wood regardless of plant species, it was shown that white rot fungi are more often specializing in decaying wood of angiosperm plants, while brown rot fungi tend to be generalists, able to decay wood from wide range of angiosperm and gymnosperm species.[3]

Armillaria genus

Fungi from genus *Armillaria* are plant pathogens that cause root rot. They are present on all continents be-

side Antarctica[9] and are generalists able to attack many species of woody plants. Species from this genus are also extraordinarily efficient at colonizing new regions, due to their ability to produce rhizomorphs – specialized root-like structured composed from vegetative mycelium, which allow the fungus to search for new hosts and for clonal dispersal. Their ability to colonize new areas can be also attributed to the fact, that they are not obligatory pathogens, rather they often act opportunistically and can survive as saprotrophic/necrotrophic white rot fungi in right conditions[10–12]. Some species of *Armillaria* were even shown to be symbiotic with plants or be hosts of mycoheterotrophic plants[13, 14].

Armillaria fungi can cause serious economic losses if they attack industrial plantations or orchards. Because of that *Armillaria* have been extensively studied both by academic mycologists and United States Department of Agriculture[15]. Some species of *Armillaria* exhibit bioluminescence, which was additional reason for studying them.[16, 17]

Individuals of some *Armillaria* species can reach extraordinary sizes. One of such individuals, aptly named "humongous fungus" is one of the largest and oldest known terrestrial organisms on Earth. Lower estimate of its size suggest that it spans 15 hectares and weights at least 10 tons. Its age was estimated to be at least 1500 years.[18]. Newer estimates by Anderson et al. suggest that it might be 2500 years old, and weight at least 40 tons[19]. Individuals of those species might be able to reach such age due to extremely slow mutation rates (magnitudes lower than those of other filamentous fungi[20].)

Motivation

There is a huge number of published studies on *Armillaria* and closely related genera's physiology, morphology, ecology and phylogeny. Usually motivation for those studies is importance of those fungi as vectors of *Armillaria* root rot disease. Sometimes they are motivated by special interesting quirks of mushrooms from this genus, such as

Table 1: Trophic strategies and geographic range of chosen *Armillaria* species. Adapted from Koch et al. 2017

Species	Facultative Necrotroph	Rizomorphs in nature	Known range	References
<i>Armillaria borealis</i>	Yes	Yes	Europe	[22–24]
<i>Armillaria fumosa</i>	Unknown	Yes ^a	Australia	[11, 25, 26]
<i>Armillaria gallica</i>	Yes	Yes	Europe, North America, Asia	[23, 27]
<i>Armillaria luteobubalina</i>	Yes	Yes	Australia, Tasmania	[28–30]
<i>Armillaria mellea</i>	Yes	Yes	Europe, North America, Asia	[23, 31]
<i>Armillaria nabsnona</i>	Yes ^b	Yes	North America, Hawaii	[32–34]
<i>Armillaria ostoyae</i> ^c	Yes	Yes	Europe, North America, Japan, China	[23, 25, 35–37]
<i>Armillaria novae-zelandiae</i>	Yes	Yes	Argentina, Australia, Chile New Zeland	[38–41]
<i>Armillaria solidipes</i> ^c	Yes	Yes	Europe, North America, Japan, China	[23, 25, 35, 36, 42]

^a As shown by Morrison et al. (2004) *Armillaria fumosa* is able to produce weak rhizomorphs in vitro

^b Likely not pathogenic or very weakly pathogenic, as per Baumgartner et al. (2001)

^c *Armillaria solidipes* is an older name for the *Armillaria ostoyae*. They are the same biological species, and name *Armillaria solidipes* should generally be used. Here, they are distinguished due to them having two different assembly genomes available in GenBank.

bioluminescence, ability to reach massive sizes and slow mutation rate.

However, a lot of phylogenetic studies of this genus, uses only a handful of molecular markers for reconstruction of phylogenetic trees. In this report I present a resolved phylogeny of 8¹ *Armillaria* species, whose annotated genomic assemblies are available in GenBank. Analysis presented here uses full proteomic data for phylogeny inference.

MATERIALS AND METHODS

Acquisition of proteomes from GenBank

Annotated genome assemblies were acquired directly from NCBI ftp server using custom Python programs. All chosen proteomes are part of GenBank database (Supplementary Table 1). Criteria for choice of given assemblies were: (1) availability of protein sequences, (2) date of submission and (3) whether they were tagged as a reference genome or not. Sometimes assemblies tagged as reference either weren't fully annotated or didn't contain protein sequences; in such cases latest annotated assembly was chosen instead.

Identification of protein families

Families of protein sequences for later inference of protein families phylogenetic trees, were calculated using MMseqs clustering method implemented in MMseqs2 program[43–45]. `cluster` and `easy-cluster` commands were tested. Other than `-c` parameter, responsible for setting the minimal threshold for coverage between sequences, default settings were used for `cluster` command. Three values were tested for `-c` parameter: 0.75, 0.8 and 0.85; 0.8, which incidentally is the default value was arbitrarily chosen as the best, offering the best balance between the number of identified clusters, number of sequences in clusters and

¹9 proteomes are used, however two of them belong to a single biological species *Armillaria solidipes*

taxa representation within clusters. `easy-cluster` command, which provides out-of-the-box easy clustering, returned identical results to `cluster` command.

Filtering the obtained clusters

Clusters were then filtered to get a set of protein families in which every taxon would be represented in each family. Two filtering variants were used: (1) only ortholog sequences (1-1 variant) (2) ortholog and paralog sequences ("with paralogs" variant).

1-1 variant

For 1-1 variant only clusters that contained one sequence per taxon, and included all 9 taxa were used. Out of 53 290 obtained clusters 3140 fitted those filtering criteria, and those were used for further analyses.

'with paralogs' variant

For variant with paralogs, in addition to clusters from 1-1 variant, clusters that included all 9 taxa but had no less than 9 and no more than 29 sequences in total were used. Out of 53 290 obtained clusters 4527 fitted those filtering criteria, 1387 of which were clusters containing paralogs.

Multialignment and inference of family trees

For each cluster from filtered subsets, a multiple sequence alignment was inferred using MAFFT[46], with the following settings: global alignment with 1000 iterations of refinement. Next ML trees with bootstrap supports were inferred for each cluster based on obtained multiple sequence alignments. To do so RAXML Next Generation[47] was run with empirically calculated stationary frequencies, GAMMA8 model for among-site heterogeneity and LG model[48] for substitution matrix. One hundred bootstrap replicated were produced for each tree to calculate support of branches.

LG model was chosen, because its widely used in phylogenetic studies of fungi, and because on average it's better

than most other models, while rarely being worse, as was shown in original paper by Le and Gascuel[48].

Consensus tress

Consensus trees were calculated from obtained protein family trees using RAXML Next Generation with `-consense` option. Consensus trees were calculated in two variants: (1) using majority rule with threshold of 50% and (2) using majority rule with greedy heuristic, which extends majority rule consensus tree by adding the compatible branches with the highest support, until a strictly bifurcating tree is obtained.

Supertree

Supertree was computed using DupTre program[49] with default parameters. DupTree uses gene tree parsimony. It searches for the species supertree that results in the fewest gene duplication events.

Filtering of trees with low support

Protein family trees containing branches with low support (<0.7) were filtered out. Then consensus and supertrees were computed on this filtered subset of trees, using the method described above.

Analysis of protein families with paralogs

All of the analyses described above were carried out for both 1-1 variant and 'with paralogs' variant, except for construction of consensus trees, which by definition cannot be used as a method to compute genome trees, using trees that contain paralogs.

Downstream analysis

Visualizations of tree comparisons was done using TreeTools and TreeDist libraries for R programming language.

RESULTS

Clustering of sequences

As expected most of the clusters contained only a small amount of sequences, with majority containing less than five (Supplementary Figure 1b). However majority of the clusters contained around one sequence per unique taxa within cluster (Supplementary Figure 1a) which meant that a lot of clusters contained families of orthologous sequences or families that contained only a small amounts of paralogs.

Genome trees

Both consensus trees (majority rule with 0.5 threshold – MR and majority rule with greedy heuristic – MRE) obtained from variant 1-1, had very similar topology (Fig. 1a-b) with only difference being that majority consensus without greedy heuristic produced a tree with multifurcation in *A. borealis* *A. ostoyae* and *A. solidipes* leaf cluster. This is to be expected as the branch that is introduced there in a variant that used greedy heuristic is present in less than 50% of trees. Species supertree obtained using

DupTree program (Fig. 1c), had identical topology to the MRE consensus tree.

Phylogenetic tree obtained from timetree[51] (Fig. 1d) corroborates the topology of MRE consensus tree and species supertree. All trees were rooted on *A. novae-zelandiae*, *A. fumosa*, *A. luteobubalina* cluster for ease of visual comparison with tree from timetree.org.

To further validate the correctness of resulting genomic trees, MRE consensus tree was compared to trees from literature. MRE consensus tree had identical topology to the abridged *Armillaria* phylogenetic tree obtained by Coetzee et al.[50] (Fig. 1e). *A. ostoyae* was not included in analysis by Coetzee et al., due to it being recognised as synonymous to *A. solidipes*, so it was skipped in this comparison. MRE consensus tree was also compared to abridged phylogenetic tree calculated by Koch et al.[21]. Once again MRE consensus tree had identical topology to that of published tree (Fig. 1f), except for the fact, that tree obtained by Koch et al. did not contain *A. nabsnona* and similarly to Coetzee et al. did not contain *A. ostoyae*.

All obtained genomic trees, except for the MR consensus tree, had identical topology to the tree from timetree.org (Supplementary Figure 2), and to trees adapted from literature, except for missing taxa.

Genomic trees constructed from the subset of family trees, that had only high-supported branches had the same topology as those constructed from unfiltered dataset (Supplementary Figure 3a-c), with a notable exception of majority rule consensus tree, that didn't have a multifurcation, when constructed from filtered dataset (Supplementary Figure 3a). What is also worth noting, is that support values on branches of consensus trees were higher, when those trees were inferred from filtered dataset.

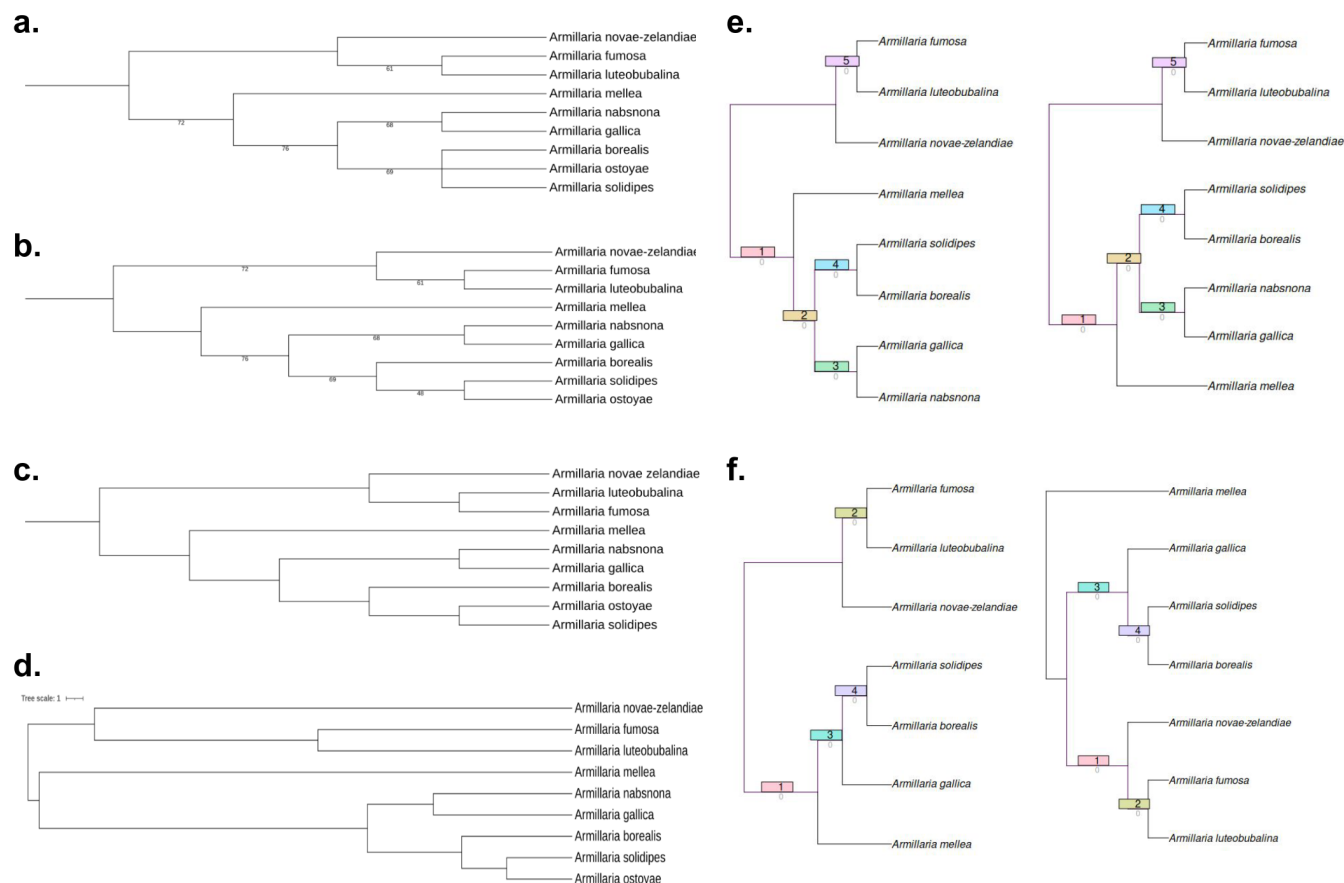
Species supertree obtained from dataset variant with paralogs had the same topology as one obtained from dataset with only orthologous clusters (Supplementary Figure 3d).

DISCUSSION

Trees obtained in analyses included in this report are rather high quality, and were able to be corroborated by multiple sources. This can be seen as an evidence supporting usage of full proteomic data for phylogenetic inferences.

Results from herein described analyses are consistent with important papers on *Armillaria* phylogeny, such as the ones by Coetzee and Koch[9, 21, 50]. It's worth noting here that work by Coetzee and Koch included computationally intensive and conceptually difficult Bayesian methods to compute phylogeny, and my analyses were able to reproduce their results using only ML methods. However it needs to be stressed that aims of mentioned studies were broader and thus justified the use of method chosen by authors.

Usage of full proteomic data is powerful method for resolving phylogeny, but it still has its drawbacks. It is dependent on access to annotated genomic data, which is often not available and it can be computationally intensive if either number of obtained clusters is high, or number of species chosen for the analysis is high. This analysis only includes 8 species and 9 proteomes. To make



One additional aspect, on which studies by Coetzee and Koch focused, was biogeography and radiation of *Armillaria* species. While methods used here do not offer any insight into those topics, resulting tree is consistent with the results from those studies (Supplementary Figure 4).

CONCLUSIONS

I managed to provide a reproduction of literature results regarding *Armillaria* genus phylogeny, using full proteomic data, simple sequence clustering and ML methods.

Data availability

All data used for the analyses is non-proprietary and freely available in GenBank under IDs provided within Supplementary Table 1.

Code availability

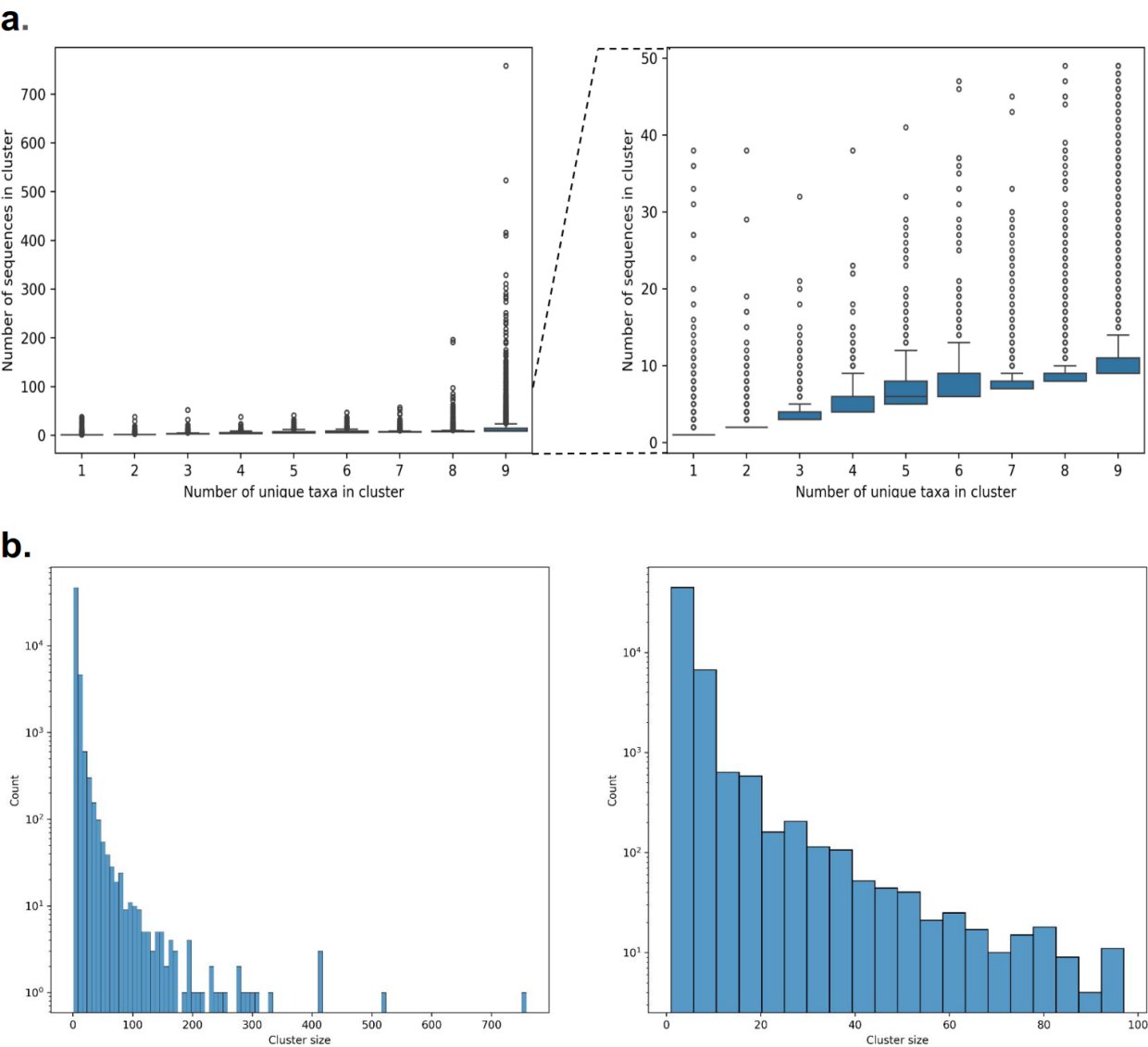
All code used for major processing and analysis steps is available in public GitHub repository under following url <https://github.com/jakubguzek/Armillaria-phylogeny>. All tools and scripts used are open source and free.

REFERENCES

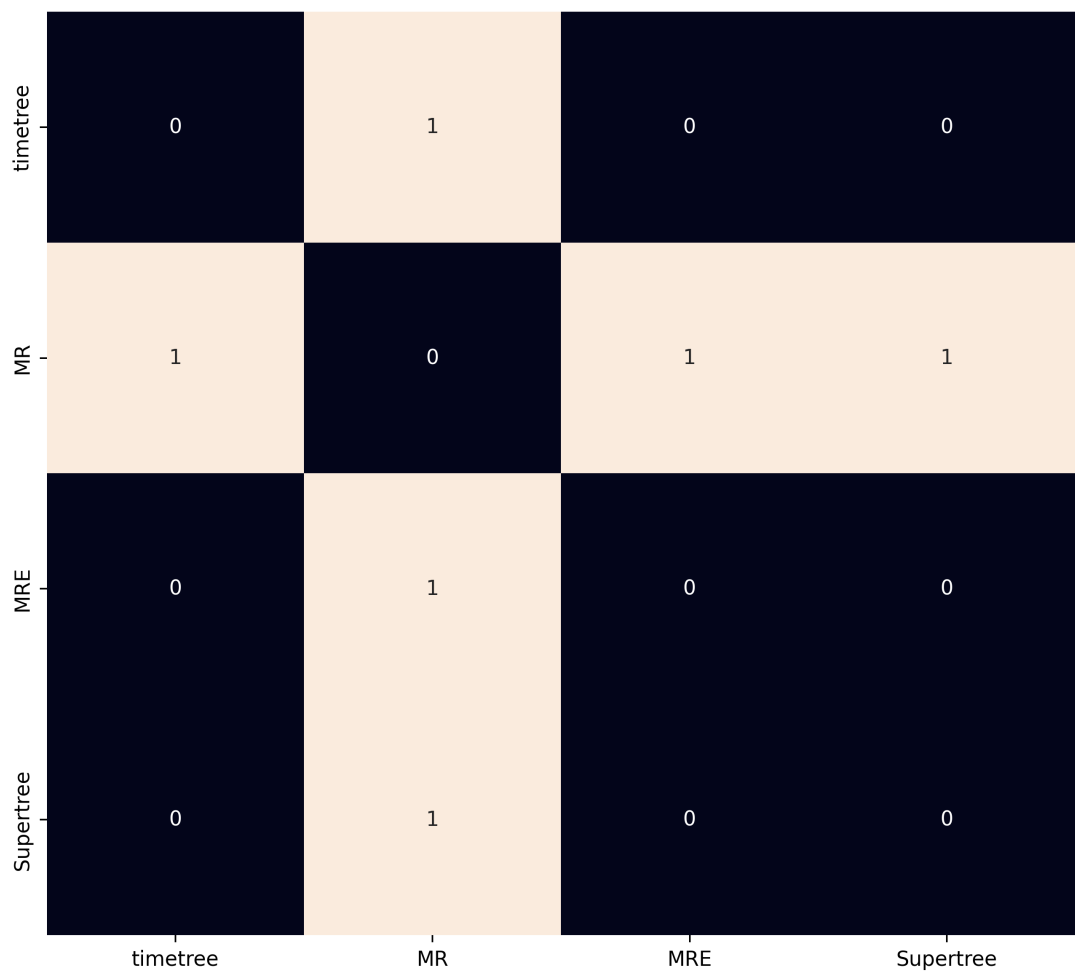
1. Harley, J. Fungi in ecosystems. *Journal of Ecology* **59**, 653–668 (1971).
2. Heimann, M. & Reichstein, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* **451**, 289–292 (2008).
3. Krah, F.-S. *et al.* Evolutionary dynamics of host specialization in wood-decay fungi. *BMC evolutionary biology* **18**, 1–13 (2018).
4. Worrall, J. J., Anagnost, S. E. & Zabel, R. A. Comparison of wood decay among diverse lignicolous fungi. *Mycologia* **89**, 199–219 (1997).
5. Brundrett, M. C. & Tedersoo, L. Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytologist* **220**, 1108–1115 (2018).
6. Leake, J. R. The biology of myco-heterotrophic ('saprophytic') plants. *New Phytologist* **127**, 171–216 (1994).
7. Bidartondo, M. I. *et al.* Epiparasitic plants specialized on arbuscular mycorrhizal fungi. *Nature* **419**, 389–392 (2002).
8. Merckx, V., Bidartondo, M. I. & Hynson, N. A. Myco-heterotrophy: when fungi host plants. *Annals of Botany* **104**, 1255–1261 (2009).
9. Coetzee, M. P., Wingfield, B. D. & Wingfield, M. J. Armillaria root-rot pathogens: species boundaries and global distribution. *Pathogens* **7**, 83 (2018).
10. Gregory, S. The use of potato tubers in pathogenicity studies of Armillaria isolates. *Plant pathology* **34**, 41–48 (1985).
11. Morrison, D. Rhizomorph growth habit, saprophytic ability and virulence of 15 Armillaria species. *Forest Pathology* **34**, 15–26 (2004).
12. Baumgartner, K., Coetzee, M. P. & Hoffmeister, D. Secrets of the subterranean pathosystem of Armillaria. *Molecular Plant Pathology* **12**, 515–534 (2011).
13. Zhan, M. *et al.* Draft genomic sequence of Armillaria gallica 012m: insights into its symbiotic relationship with Gastrodia elata. *Brazilian Journal of Microbiology* **51**, 1539–1552 (2020).
14. Liu, J.-J. *et al.* The role of symbiotic fungi in the life cycle of Gastrodia elata Blume (Orchidaceae): a comprehensive review. *Frontiers in Plant Science* **14** (2023).
15. Shaw, C. G. & Kile, G. A. *Armillaria root disease* **691** (Forest Service, US Department of Agriculture, 1991).
16. Ke, H.-M. *et al.* Mycena genomes resolve the evolution of fungal bioluminescence. *Proceedings of the National Academy of Sciences* **117**, 31267–31277 (2020).
17. Oliveira, A. G. *et al.* Circadian control sheds light on fungal bioluminescence. *Current Biology* **25**, 964–968 (2015).
18. Smith, M. L., Bruhn, J. N. & Anderson, J. B. The fungus Armillaria bulbosa is among the largest and oldest living organisms. *Nature* **356**, 428–431 (1992).
19. Anderson, J. B. *et al.* Clonal evolution and genome stability in a 2500-year-old fungal individual. *Proceedings of the Royal Society B* **285**, 20182233 (2018).
20. Anderson, J. B. & Catona, S. Genomewide mutation dynamic within a long-lived individual of Armillaria gallica. *Mycologia* **106**, 642–648 (2014).
21. Koch, R. A. *et al.* Resolved phylogeny and biogeography of the root pathogen Armillaria and its gasteroid relative, Guyanagaster. *BMC Evolutionary Biology* **17**, 1–16 (2017).
22. Gregory, S. & Watling, R. Occurrence of Armillaria borealis in Britain. *Transactions of the British Mycological Society* **84**, 47–55 (1985).
23. Guillaumin, J.-J. *et al.* Geographical distribution and ecology of the Armillaria species in western Europe. *European Journal of Forest Pathology* **23**, 321–341 (1993).
24. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.bdqqyv>.
25. Mihail, J. D. & Bruhn, J. N. Foraging behaviour of Armillaria rhizomorph systems. *Mycological research* **109**, 1195–1207 (2005).
26. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.6d7rce>.
27. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.edbsbb>.
28. Coetzee, M. P. *et al.* Molecular identification and phylogeny of Armillaria isolates from South America and Indo-Malaysia. *Mycologia* **95**, 285–293 (2003).
29. Pareek, M., Louise, C. & Ashford, A. E. Variations in structure of aerial and submerged rhizomorphs of Armillaria luteobubalina indicate that they may be organs of absorption. *Mycological research* **105**, 1377–1387 (2001).
30. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.wxcsqw>.
31. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.q2udke>.
32. Volk, T. J., Burdsall Jr, H. H. & Banik, M. T. Armillaria nabsona, a new species from western North America. *Mycologia* **88**, 484–491 (1996).
33. Baumgartner, K. & Rizzo, D. M. Distribution of Armillaria species in California. *Mycologia* **93**, 821–830 (2001).
34. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.hgxge9>.
35. Hasegawa, E. *et al.* Ecology of Armillaria species on conifers in Japan. *Forest Pathology* **41**, 429–437 (2011).
36. Burdsall, H. & Volk, T. J. Armillaria solidipes, an older name for the fungus called Armillaria ostoyae. *North American Fungi* **3**, 261–267 (2008).
37. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.5t25gh>.
38. Shaw III, C., Calderon, S., *et al.* Impact of Armillaria root rot in plantations of Pinus radiata established on sites converted from indigenous forest. *New Zealand Journal of Forestry Science* **7**, 359–373 (1977).
39. III, C. S. *et al.* Cultural characteristics and pathogenicity to Pinus radiata of Armillaria novae-zelandiae and A. limonea. *New Zealand journal of forestry science* **11**, 65 (1981).
40. Pildain, M. B. *et al.* Taxonomy of Armillaria in the Patagonian forests of Argentina. *Mycologia* **102**, 392–403 (2010).
41. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.y4ybbx>.
42. GBIF.org. *GBIF Occurrence Download* <https://doi.org/10.15468/dl.yfn5rv>.
43. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology* **35**, 1026–1028 (2017).
44. Steinegger, M. & Söding, J. Clustering huge protein sequence sets in linear time. *Nature communications* **9**, 2542 (2018).
45. Mirdita, M., Steinegger, M. & Söding, J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics* **35**, 2856–2858 (2019).
46. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **30**, 772–780 (2013).
47. Kozlov, A. M. *et al.* RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
48. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Molecular biology and evolution* **25**, 1307–1320 (2008).
49. Wehe, A. *et al.* DupTree: a program for large-scale phylogenetic analyses using gene tree parsimony. *Bioinformatics* **24**, 1540–1541 (2008).
50. Coetzee, M. P. *et al.* Paleogene radiation of a plant pathogenic mushroom. *PLoS One* **6**, e28545 (2011).
51. Kumar, S. *et al.* TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular biology and evolution* **34**, 1812–1819 (2017).

Supplementary Table 1: IDs and names of annotated assemblies from GenBank chosen for this study.

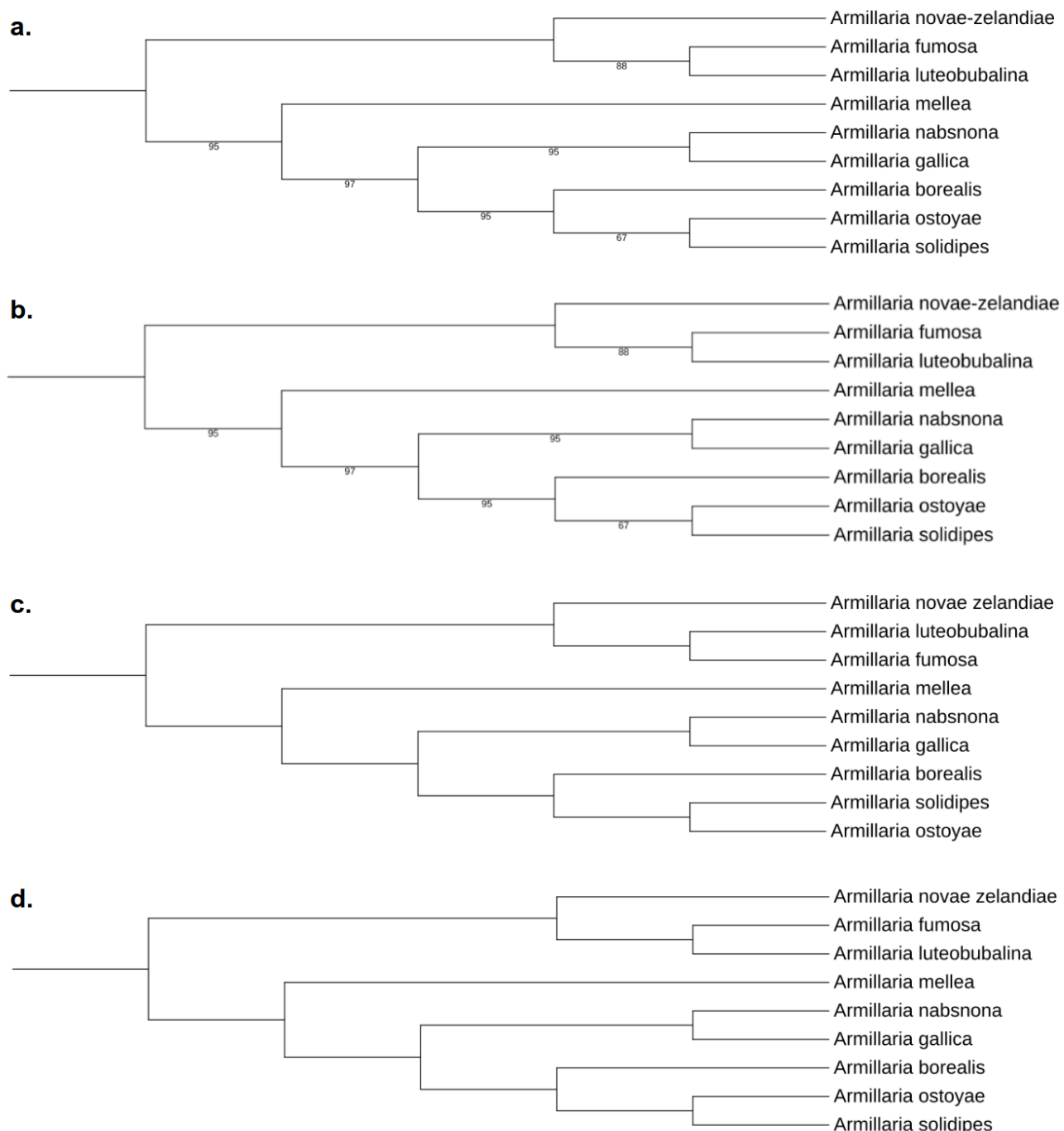
Species	GenBank ID	Assembly name	Level
<i>Armillaria borealis</i>	GCA_030435635.1	Armbor1	Contig
<i>Armillaria fumosa</i>	GCA_030407015.1	Armfum1	Contig
<i>Armillaria gallica</i>	GCA_002307695.1	Armga1	Scaffold
<i>Armillaria luteobubalina</i>	GCA_030435615.1	Armlut1	Contig
<i>Armillaria mellea</i>	GCA_030407055.1	Armmel1	Contig
<i>Armillaria nabsnona</i>	GCA_030407065.1	Armnab1	Contig
<i>Armillaria novae-zelandiae</i>	GCA_030435655.1	Armnov1	Contig
<i>Armillaria ostoyae</i>	GCA_900157425.1	—	Contig
<i>Armillaria solidipes</i>	GCA_002307675.1	Armost1	Scaffold



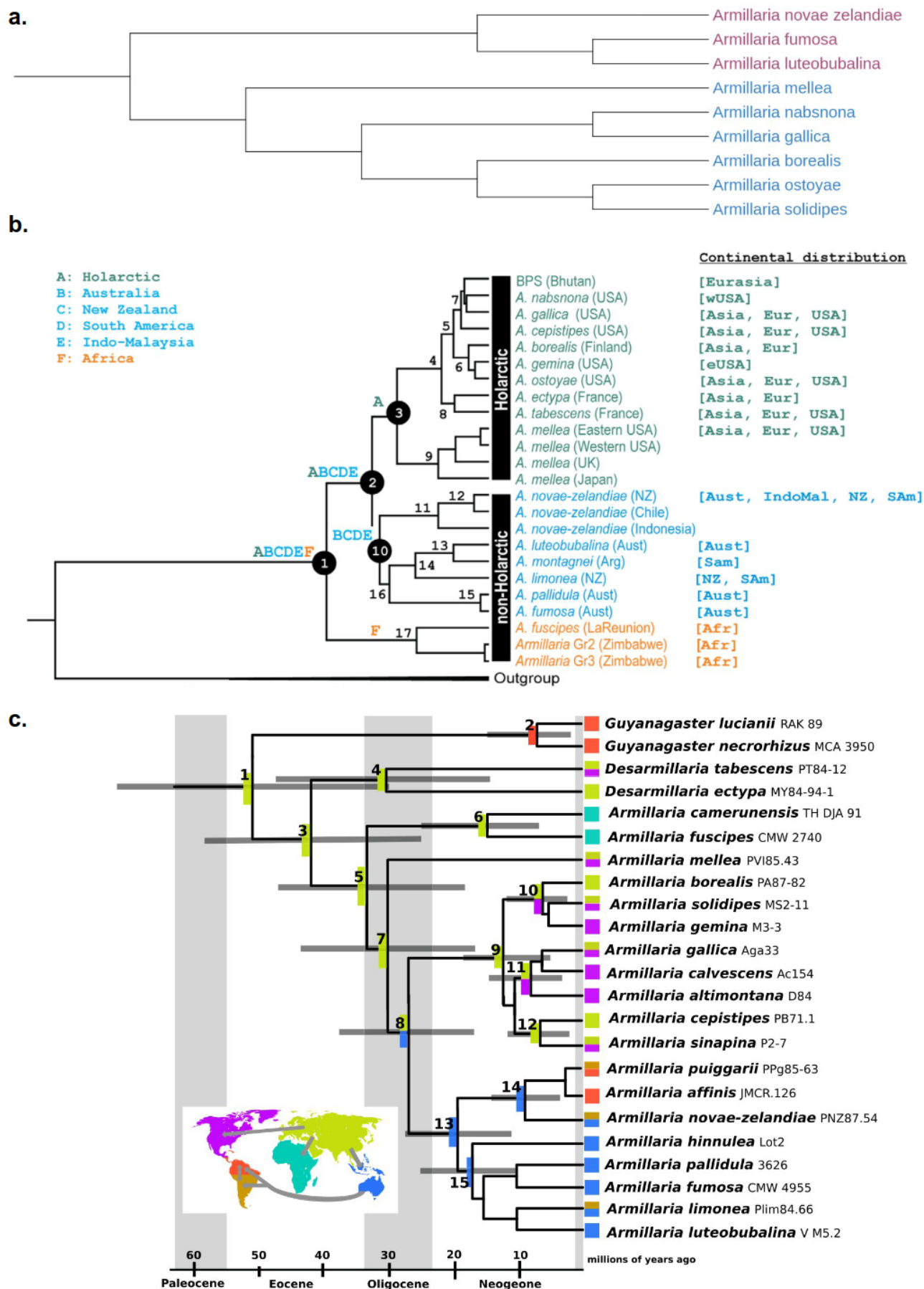
Supplementary Figure 1: **Most clusters contained only one sequence, but majority of those that had 9 unique taxa had around 10 sequences.** **a.** Box-plot showing the spread of numbers of sequences per cluster in clusters with given number of unique taxa. Most clusters contained one sequence per taxon and only some contained more sequences per taxon, and thus contained no paralogs, or just one or two paralogs **b.** Histogram of numbers of sequences per cluster. Most clusters contained small number of sequences, but a small amount of clusters contained a lot.



Supplementary Figure 2: **Robinson-Foulds distance between all obtained trees, except for MR consensus tree, and timetree was 0.** Heatmap showing distances between trees obtained from variant 1-1 clusters.



Supplementary Figure 3: **Filtering trees with low-support branches does not impact the topology of resulting genomic trees. Neither does including the clusters with paralogs** **a.** Rooted majority consensus tree obtained from 1-1 variant after filtering out trees with branches that had low (<0.7) bootstrap support. This is the only tree for which filtering seems to improve the topology as, there is no longer a multifurcation present in *A. ostoyae* *A. solidipes* *A. borealis* cluster. **b.** Rooted greedy majority consensus tree obtained from 1-1 variant after filtering out tree with branches that had low bootstrap support. **c.** Rooted species supertree obtained using DupTree from the subset of trees after filtering. **d.** Species supertree obtained using DupTree from variant with paralogs. As can be observed, the topology of supertree obtained from trees with paralogic sequences is identical to that obtained from tree containing only orthologic sequences.



Supplementary Figure 4: **Species supertree is consistent with findings on paleogene radiation and biogeography of *Armillaria* studied by Coetzee et al. and Koch et al.** a. Species supertree with leaves colored by major geographic regions: blue – Holarctic, purple – Australia and South America b. Ancestral area reconstruction by Coetzee et al. 2011 c. Chronogram with ancestral area reconstruction by Koch et al. 2017