Genomic Insights into the Phylogenetics, Ecology, and Reproductive Biology of Lorchel Mushrooms (*Gyromitra* spp.) and the Biosynthesis and Evolution of the Gyromitrin Mycotoxin

by

Alden C. Dirks

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in the University of Michigan
2024

Doctoral Committee:

Professor Timothy James, Chair
Professor Catherine Badgley
Professor Regina Baucom
Professor Stephen Smith

Alden C. Dirks

adirks@umich.edu

ORCID iD: 0000-0002-3997-0122

© Alden C. Dirks 2024

Dedication

This thesis is dedicated to the people of Palestine. We are all bearing witness to genocide committed by Israel. Americans are accomplices in the indiscriminate bombings, mass murder, and starvation of Palestinian citizens. Billions of our tax dollars fund an endless stream of weapon transfers that arm the genocide. It must stop.

As academics, we have the tremendous privilege of pursuing our passions in our work. Every human should enjoy the privilege of being able to follow their interests, to feel curiosity, to dream, to wonder. How can somebody do that when they are denied even the basic human right of living without fear of starvation and murder? How many mycologists, ornithologists, entomologists, and evolutionary biologists of Palestine will never know their passions?

Our land acknowledgments are empty tokenism if not supported by action. It is not enough to reiterate the sordid history of University of Michigan—its founding via land theft and genocide of the Anishinaabeg and Wyandot peoples. We need Land Back for these Indigenous Nations. We need to divest from genocide. At every opportunity, we all must be screaming for this to stop, which is why this humble dedication exists.

Acknowledgements

My foray into the topic of lorchels and gyromitrin began with a tentative, unfunded research proposal at the beginning of my PhD that would have me conducting field work in the spring. Mushroom hunting during morel season? “Sign me up!” I said. It didn’t matter much what I was doing if I could be outside serendipitously collecting morels. Little did I know that the lorchels (also called false morels) would become a passion of mine and totally eclipse morels, not least of all for their culinary properties—they are truly delicious! (Please be very careful if you attempt to lorchels, some of which are deadly poisonous). Entering as a self-identified mycologist who simply loved mushrooms, my PhD journey has brought me much deeper than I could have ever imagined into the worlds of bioinformatics, genetics, and chemistry. I am grateful to my advisor, Tim James, for being a steadfast source of advice, guidance, and mentorship along the way. Tim is an exceptional mycologist with broad knowledge of fungi, which allows him to engage in (and advise on) any mycological topic, but he also has tremendous expertise in numerous subjects, which has aided my growth as a scientist. I am also grateful to Tim for entrusting me with a substantial amount of autonomy during my PhD. Shortly after I began, the world shut down because of the COVID-19 pandemic. The ensuing isolation was excruciating and disqualified in-person research activity for a solid year. After two years, I moved to Philadelphia to again be with my partner, Taylor Tai, and other community, and have worked on my PhD research in a hybrid format since. I’ve spent a third of my life with Taylor and I’m so excited for many more years to come. Taylor inspires me every

day with her creativity, moral clarity, and intelligence. We have grown so much together, and I am deeply grateful for her support and love over so many years, especially during these last stages of my PhD. Even though she did not choose the path of the academic, she has a strong scholarly mind and has been a source of intellectual companionship and collaboration in my work. For many delicious shared meals, games of Wingspan, mushroom hunts, and commitment to community, I also want to thank Alex, Amanda, Ami, Ashley, Fianna, Kase, Marlen, Rivka, Ruby, Sasha, Scott, and Thaddeus, among other beloved friends and family, as well as those who so kindly housed and fed me during my visits to Ann Arbor—thank you! My PhD research itself has been accomplished with the help of so many individuals. I'm thankful to members of the James Lab, past and present, for their support, inspiration, and comradery. Elizabeth Feliciano is dearly missed. She was patient with me and my fear of dogs, brought beautiful laughter to this world, and was gone far too soon. I'm grateful to lab manager Michelle Orozco-Quime for being an emotional support during challenging protocols and making scary lab tasks approachable. I'm also grateful to Eva Schwarz who worked diligently on developing a transformation protocol. While not yet finished, I suspect this is not the end of that work. I want to thank my PhD committee members, Catherine Badgley, Regina Baucom, and Stephen Smith for being supportive, encouraging, and nice to me over the years we've gathered to discuss my dissertation progress. Finally, many individuals, both community scientists and fungarium managers, contributed fungal specimens to make this research possible. Sourcing a diversity of specimens to conduct broad biological research is an underappreciated challenge, and many folks working in the background made this possible. They are acknowledged by name throughout the thesis.

Table of Contents

Dedication ii

Acknowledgements iii

List of Tables viii

List of Figures ix

Abstract xiv

Chapter 1 Introduction 1

1.1 Fungal secondary metabolites are diverse and important 1

1.2 Genetic paradigms in secondary metabolite biosynthesis 2

1.3 *Gyromitra esculenta* is infamous for its production of the mycotoxin gyromitrin 3

1.4 Gyromitrin is a small, volatile mycotoxin with a rare N–N bond moiety 4

1.5 Preview of chapters 5

References 7

Chapter 2 Not All Bad: Gyromitrin Has a Limited Distribution in the False Morels as Determined by a New Ultra High-Performance Liquid Chromatography Method 9

Abstract 9

2.1 Introduction 10

2.2 Materials and Methods 14

2.2.1 Specimen acquisition 14

2.2.2 Molecular data 15

2.2.3 Phylogenetic analyses 16

2.2.4 Gyromitrin analytical method development 17

2.3 Results 20
2.4 Discussion 22
2.4.1 On the edibility of lorchels 24
2.4.2 Conclusions and open questions 26
Acknowledgements 28
Funding 28
Figures 29
Tables 35
References 38
Supplement 44
Chapter 3 Taxonomic Revisions and Insights into the Ecology and Mating Systems of Lorchels (*Discinaceae*, *Pezizales*) Using Whole Genome Analyses 54
Abstract 54
3.1 Introduction 55
3.2 Materials and Methods 59
3.2.1 Specimen acquisition and DNA sequencing 59
3.2.2 Genome assembly and annotation 59
3.2.3 Phylogenomics and time divergence estimation 60
3.2.4 CAZyme analysis, mating systems, and ancestral state reconstruction 61
3.3 Results 62
3.4 Discussion 65
3.4.1 Phylogenomics and taxonomy 65
3.4.2 Ecology 67
3.4.3 Mating systems 68
3.4.4 Gyromitrin ancestral state reconstruction 69
3.4.5 Closing remarks 70

Acknowledgements 71

Figures 73

References 77

Supplement 87

Chapter 4 Bioinformatics Identification of a Putative Gyromitrin Biosynthesis Gene Cluster in Lorchel Mushrooms (*Gyromitra* spp., *Pezizales*) 98

Abstract 98

4.1 Introduction 99

4.2 Materials and Methods 102

4.2.1 Homology search 103

4.2.2 Phylogenetic incongruence search 103

4.3 Results 104

4.3.1 Homology search 104

4.3.2 Phylogenetic analysis 105

4.4 Discussion 109

4.4.1 Putative gyromitrin BGC 1 likely produces an iron-chelating siderophore 109

4.4.2 Putative gyromitrin BGC 2 is our leading hypothesis for gyromitrin biosynthesis 109

4.4.3 Putative gyromitrin BGC 3 likely encodes a RiPP cyclic peptide mycotoxin 112

4.4.4 Conclusions 115

Acknowledgements 116

Figures 117

Tables 124

References 125

Supplement 137

Chapter 5 Conclusion 150

References 153

List of Tables

Table 2.1. Specimens tested for gyromitrin and their associated metadata. 35

Table 2.2. Summary of available evidence for the toxicity and presence of gyromitrin in select *Discinaceae* species based on primary literature and tests conducted for this study. Bolded species are taxa that have been shown to contain gyromitrin. Evidence of toxicity does not necessitate that gyromitrin is the etiological factor, and no evidence of toxicity does not mean that gyromitrin is absent from that species. 37

Supplemental Table 2.1. Accession numbers of GenBank sequences used in the phylogenetic analyses. 50

Supplemental Table 2.2. Peak area and recovery percentage of **13** in the pre-column derivatization (18 hours post incubation at 40 C) and UHPLC-DAD analysis (370 nm) of gyromitrin standard at various concentrations. 51

Supplemental Table 2.3. Metadata and results of gyromitrin tests conducted on samples for this study. 52

Supplemental Table 3.1. Metadata for samples whose genomes were sequenced. 92

Table 4.1. Nitrogen-nitrogen bond-containing natural product biosynthesis gene clusters and nitrogen primary metabolism enzymes that were used as queries in a BLASTP homology search against a custom *Discinaceae* proteomes database. 124

Supplemental Table 4.1. Orthogroups whose phylogenies had gyromitrin-producing taxa forming a monophyletic clade and contained *Gyromitra esculenta* CBS101906. The location refers to the position in the Joint Genome Institute MycoCosm genome browser for *G. esculenta*, and the protein ID is a URL to the MycoCosm page for that protein..... 148

List of Figures

Figure 2.1. Chemical structures of gyromitrin (**1**) and its homologs (**2–9**). 29

Figure 2.2. Acid hydrolysis of gyromitrin and homologs (**1–9**) yields *N*-methyl-*N*-formylhydrazine (**10**) and monomethylhydrazine (**11**). Derivatization of **10** and **11** with 2,4-dinitrobenzaldehyde (2,4-DNB) yields Schiff bases **12** and **13**..... 30

Figure 2.3. UHPLC-DAD (370 nm) chromatograms of derivatization progress of gyromitrin in 10 mg of dried *Gyromitra venenata* ascocarp (MICH 352032). a. Ascocarp extract without 2,4-DNB/TFA treatment (negative control). b–e. Ascocarp extract treated with 2,4-DNB/TFA..... 31

Figure 2.4. UHPLC-DAD (370 nm) chromatograms of ascocarp extracts. a. *Gyromitra leucoxantha* (MICH 352087). The peak corresponding to an unidentified Schiff base derivative is shaded red. b. *Gyromitra splendida* (MICH 352089). Photograph of *G. splendida* by Eric and Jen Chandler. Photograph of *G. leucoxantha* by Garrett Taylor..... 32

Figure 2.5. UHPLC-DAD (370 nm) chromatograms of 100 mg of *Gyromitra venenata* (a–g, i–l) or *Gyromitra splendida* (h) dried ascocarps collected in different years. a. MICH 25555 (1883). b. MICH 25541 (1945). c. MICH 25528 (1966). d. MICH 1304 (1972). e. MICH 25573 (1976). f. MICH 25543 (1979). g. MICH 25578 (1982). h. MICH 25554 (1984). i. MICH 39037 (1995). j. MICH 39039 (1998). k. MICH 68345 (2001). l. MICH 68493 (2003).. 33

Figure 2.6. *Discinaceae* maximum likelihood phylogenetic tree based on concatenated aligned sequences of the ITS barcode and 28S rDNA showing the presence and absence of gyromitrin. Specimens tested for gyromitrin are highlighted: blue means no gyromitrin was detected (negative); red means gyromitrin was detected (positive). Bootstrap values ≥ 0.70 are shown at the nodes, and Bayesian posterior probability scores above 0.95 are indicated by thickened branches. Taxon names for sequences derived from taxonomic type material are bolded..... 34

Supplemental Figure 2.1. UHPLC-DAD (370 nm) trace and UV-vis spectrum of synthetic monomethylhydrazine Schiff base (**13**). 44

Supplemental Figure 2.2. Standard calibration curve of synthetic monomethylhydrazine Schiff base (**13**) at different concentrations (µg/mL) against UV peak area (370 nm). 45

Supplemental Figure 2.3. UHPLC-DAD (370 nm) chromatograms of derivatization progress of gyromitrin standard (20 µg/mL) with 2,4-DNB at different time intervals. a) 0 h, b) 2 h, c) 5 h, d) 8 h, e) 13 h, f) 18 h, and g) 24 h. 46

Supplemental Figure 2.4. Hydrolysis and derivatization progress of different concentrations of gyromitrin standard using 2,4-DNB over time. 47

Supplemental Figure 2.5. Correlation between different concentrations of gyromitrin standard solution and UV peak area (370 nm) of derivatized product (**13**). 48

Supplemental Figure 2.6. Hydrolysis and derivatization progress of gyromitrin in different amounts of *Gyromitra venenata* (MICH 352032) over time using 2,4-DNB..... 49

Figure 3.1. Morphological diversity of *Discinaceae* and related taxa. A–M. Species belonging to *Discinaceae*. N–O. Species belonging to *familia incertae sedis*. Subgenus *Discina*: A) *Gyromitra ancilis* MICH352026; B) *Gyromitra mcknightii* MICH346490. Subgenus *Caroliniana*: C) *Gyromitra brunnea* MICH352018; D) *Gyromitra martinii* MICH346002. Subgenus *Gigas*: E) *Gyromitra korfii* MICH352049; F) *Gyromitra* sp. 2 MICH346448. Subgenus *Leucoxantha*: G) *Gyromitra leucoxantha* MICH352087. Subgenus *Melaleucoides*: H) *Gyromitra melaleucoides* iNat134134066. Subgenus *Hydnotrya*: I) *Gyromitra bailii* MICH345143. Subgenus *Paragyromitra*: J) *Gyromitra infula* MICH352090. Subgenus *Gyromitra*: K) *Gyromitra venenata* MICH352028. Subgenus *Pseudorhizina*: L) *Gyromitra californica* MICH352088. Subgenus *Pseudoverpa*: M) *Gyromitra anthracobia* MICH346478. *Familia incertae sedis*: N) *Marcelleina donadinii* VIT9785; O) *Paradiscina melaleuca* MICH346528, LY−NV2020.05.32. Photo credits: Alden Dirks (A, C, K), Bernd Fellmann (D, I), Brittany Marcotte (M), Drew Henderson (B), Eric and Jen Chandler (J), Garrett Taylor (E, G), Jacques Guinberteau (O), Rubén Martínez Gil (N), Steve Ness (L), Vail Paterson (F), and Yi-Min Wang (H). 73

Figure 3.2. Phylogenomic tree of the lorchel family *Discinaceae* and related *Pezizales* families. The phylogenomic tree was inferred with 1542 concatenated BUSCOs that were analyzed using 1000 ultrafast bootstrap replicates and an edge-linked proportional partition model in IQ-TREE2. Gene concordance factors (gCF) are listed on the branches; values less than 50% are colored red. Families are colored according to the legend. *Discinaceae* subgenera are labelled and the diversity of gross morphology represented in our sampling within each subgenus or outgroup family is represented by icons. Sequences derived from type specimens are in **bold**. 74

Figure 3.3. *Discinaceae* time divergence estimation, modeling of AA9 CAZyme family evolutionary history, gyromitrin status, and genome traits. Branches are colored according to the AA9 CAZyme count inferred through time with CAFE5. The background color of the tips of the tree indicates gyromitrin status (red: tested positive; blue: gyromitrin not detected; no background color: not tested). Genome size, average GC content, total number of genes sized by number of transposable elements, number of CAZyme classes, and number of biosynthesis gene clusters predicted by antiSMASH are shown for each sample. The dotted lines indicate mean values. 75

Figure 3.4. *Discinaceae* mating type (*MAT*) locus gene synteny diagram and scatterplot of *MAT* idiomorph coverage. A) *MAT* locus gene diagrams from representatives *Discinaceae* species. Genes are indicated by arrows and homologous genes have the same fill color. Conserved flanking genes of the *MAT* locus are in **bold** (*APN1* and *SLA2*). *MAT* idiomorphs are surrounded by boxes. Vertical lines show breaks in contigs. B) Scatterplot of *MAT* idiomorph coverage for genomes that contained two *MAT* idiomorphs. The size of dots indicates genome coverage. Darker and lighter circles correspond to species that produce

epigeous or hypogeous ascomata, respectively. The diagonal dotted line shows a *MAT* idiomorph coverage ratio of one. The inset shows the expected distribution of *MAT* idiomorphs within *Discinaceae* ascomata across different macromorphological forms. Teal corresponds to tissue containing one primary *MAT* idiomorph; purple corresponds to tissue containing both idiomorphs in roughly equivalent numbers. 76

Supplemental Figure 3.1. Coalescent phylogenomic tree of *Discinaceae* and related *Pezizales* families. A total of 1645 individual gene trees were inferred with IQ-TREE2 and a coalescent phylogenomic tree was created with ASTRAL-III. Numbers on branches are quartet probabilities..... 87

Supplemental Figure 3.2. Cladogram of *Discinaceae* and related *Pezizales* families with concordance pie charts. Gene concordance and site concordance values are listed above and below the branches, respectively. For each branch, the pie charts show the proportion of decisive gene trees that supported the species tree (dark green), the most likely alternate arrangement of the quarter (red), and the second most likely alternate arrangement of the quarter (light red); grey pie wedges indicate the proportion of decisive gene trees that were polyphyletic for a branch, meaning one or more of the groupings of the species tree quartet at that branch were not recovered as monophyletic..... 88

Supplemental Figure 3.3. Scatterplot of log-transformed number of transposable elements and total gene number. The dotted line is the best fit phylogenetic generalized least squares regression. 89

Supplemental Figure 3.4. Heatmap of *MAT* idiomorph coverage. *Discinaceae* phylogram tip labels are colored according to gyromitrin status. *MAT* idiomorph coverage was calculated by mapping raw reads to the *MAT* gene sequence with BBTools. 90

Supplemental Figure 3.5. Ancestral state reconstruction of gyromitrin production in *Discinaceae*. Circles at the tips show observed gyromitrin production (yellow means present, purple means absent). Ancestral nodes show posterior probabilities of gyromitrin production from stochastic character mapping with phytools 2.1-1. 91

Figure 4.1. Chemical structures of gyromitrin (**1**) and its higher homologs (**2**–**9**). 117

Figure 4.2. *Discinaceae* phylogenomic time divergence species tree and distribution of gyromitrin production. The species tree and taxonomy are from Chapter 2 and the gyromitrin information from Chapter 2 and Lagrange *et al.* (in review). Tip labels in red correspond to specimens that tested positive for gyromitrin; those in blue did not have detectable levels of gyromitrin; and ones without color were not tested. Colors corresponding to "top-clade taxa" and "bottom-clade taxa" (orange and purple, respectively) are used in phylogenetic trees throughout the study. 118

Figure 4.3. Schematic of bioinformatics analyses conducted in this study for the identification of putative gyromitrin biosynthesis gene clusters. *N2NP homology search.* A dataset of known N2NP biosynthesis gene clusters was manually curated. The protein sequences were extracted from NCBI GenBank and queried against a custom *Discinaceae* proteomes database with BLASTP. An E-value cutoff of 1E-5 was used to identify homologs. *Phylogenetic*

incongruence search. 62 Discinaceae proteomes were clustered into homologous with OrthoFinder. Orthogroups with four or more taxa and consisting of at least one representative from both gyromitrin-producing clades (subg. *Gyromitra* and subg. *Leucoxantha*) had their protein phylogenies inferred with IQ-TREE2. Gene trees that showed the gyromitrin-producers as more closely related than expected, particularly when they formed a monophyletic clade, were selected (dashed box). Red branches indicate gyromitrin-producing clades. 119

Figure 4.4. Putative gyromitrin biosynthesis gene clusters (BGC) and their predicted functions. The gene maps were created with Clinker and are sized according to the scale. Genes belonging to the putative gyromitrin BGCs are colored. Homologous genes are linked by the corresponding color. Genes are listed with their predicted functioned according to antiSMASH and the annotated JGI *Gyromitra esculenta* CBS101906 reference genome. Protein IDs correspond to the *G. esculenta* JGI genome and can be accessed at "mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=[protein_id]", where "[protein_id]" is to be replaced with the protein ID number given in parentheses. 120

Figure 4.5. Phylogenetic tree of proteins homologous to genes 2.1, 2.2, 2.3, and 2.5 of putative gyromitrin biosynthesis gene cluster 2. The phylogenetic tree was inferred with IQ-TREE2 and midpoint rooted. Ultrafast bootstrap approximation support values greater than or equal to 95% are shown on the branches. Gyromitrin-producing clades are indicated by red bars. "Top-clade" taxa are colored orange, and "bottom-clade" taxa are colored purple (see Figure 4.2). See Figure 6 for the phylogenetic tree corresponding to gene 2.4. See the main text for the associated functions and *Gyromitra esculenta* CBS101906 MycoCosm protein ID..... 121

Figure 4.6. Phylogenetic tree of proteins homologous to gene 2.4 of putative gyromitrin biosynthesis gene cluster 2 and an alignment of the associated amino acid sequences. The phylogenetic tree was inferred with IQ-TREE2 and midpoint rooted. Branch values show ultrafast bootstrap approximation support. Gyromitrin-producing clades are indicated by red bars. "Top-clade" taxa are colored orange, and "bottom-clade" taxa are colored purple (see Figure 4.2). The bounds of the tyrosyl-tRNA synthetase and GCN5-related *N*-acetyltransferase (GNAT) domains of the didomain protein (gene 2.4) were identified with the NCBI Conserved Domains search and aligned separately with MUSCLE. The amino acid alignments are shaded according to their average percent similarity to the *Gyromitra esculenta* CBS101906 sequence. *Gyromitra venenata* MICH346486 was excluded due to erroneous gene prediction resulting in an artificially fragmented gene. 122

Figure 4.7. Phylogenetic tree of proteins homologous to UstYa-family genes 3.6 and 3.9 from putative gyromitrin biosynthesis gene cluster 3. The phylogenetic tree was inferred with IQ-TREE2 and midpoint rooted. Ultrafast bootstrap approximation support values greater than or equal to 95% are shown on the branches. Gyromitrin-producing clades are indicated by red bars. "Top-clade" taxa are colored orange, and "bottom-clade" taxa are colored purple (see Figure 4.2)..... 123

Supplemental Figure 4.1. Gene synteny map of putative gyromitrin biosynthesis gene cluster 1. Red highlights indicate gyromitrin-producing clades..... 137

Supplemental Figure 4.2. Phylogenetic trees of the orthogroups corresponding to putative gyromitrin biosynthesis gene cluster 1 inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. Homologs of gene 1.3 were absent from *Gyromitra* subgen. *Leucoxantha* and no phylogeny was generated. See the main text for corresponding protein IDs. 138

Supplemental Figure 4.3. Phylogenetic trees of orthogroups with all gyromitrin-producing samples present and inferred to form a monophyletic clade with IQ-TREE2 using 1000 ultrafast bootstrap replicates. See Table S1 for corresponding protein IDs. 139

Supplemental Figure 4.4. Phylogenetic trees of orthogroups corresponding to putative gyromitrin biosynthesis gene cluster 2 and flanking genes inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. The phylogenetic trees are organized in the same order as they appear from left to right for *Gyromitra* subg. *Leucoxantha* (*Gyromitra leucoxantha* and *Gyromitra persicula*) in Figure 6. 140

Supplemental Figure 4.5. Gene synteny map of putative gyromitrin biosynthesis gene cluster 2, conserved region A. The putative gyromitrin BGC 2 genes are indicated by a box. Red highlights indicate gyromitrin-producing clades. 141

Supplemental Figure 4.6. Gene synteny map of putative gyromitrin biosynthesis gene cluster 2, conserved region B. The putative gyromitrin BGC 2 genes are indicated by a box. Red highlights indicate gyromitrin-producing clades. 142

Supplemental Figure 4.7. Gene synteny map of putative gyromitrin biosynthesis gene cluster 2, conserved region C. The putative gyromitrin BGC 2 genes are indicated by a box. Red highlights indicate gyromitrin-producing clades. 143

Supplemental Figure 4.8. Gene synteny map of conserved regions A, B, and C of putative gyromitrin biosynthesis gene cluster 2. The putative gyromitrin BGC 2 genes are indicated by boxes. Red highlights indicate gyromitrin-producing clades. 144

Supplemental Figure 4.9. Phylogenetic trees of orthogroups corresponding to *Gyromitra esculenta* CBS101906 biosynthesis gene clusters in which the gyromitrin-producing subgenera were more closely related than expected. The trees were inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. 145

Supplemental Figure 4.10. Gene synteny map of putative gyromitrin biosynthesis gene cluster 3. The UstYa-family genes are inside the boxes (gold and blue genes). Red highlights indicate gyromitrin-producing clades. 146

Supplemental Figure 4.11. Phylogenetic trees of orthogroups corresponding to putative gyromitrin biosynthesis gene cluster 3 (see main text for details) inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. The phylogenies for the core UstYa-family genes 3.6 and 3.9 are displayed in Figure S9 (scaffold 33 006830) and the main text (Figure 4.6). The remaining missing orthogroup phylogenies were not generated since they were absent from *Gyromitra* subg. *Leucoxantha* samples. 147

Abstract

Gyromitrin is a mycotoxin produced by the fungus *Gyromitra esculenta* and related species, which are collectively called false morels or lorchels. Despite the health risks of gyromitrin exposure, *G. esculenta* mushrooms are consumed as a delicacy in some parts of the world after detoxification. While the chemical structure of gyromitrin has been known for decades, the distribution, evolutionary history, and genetics of gyromitrin are poorly studied or completely unknown. My aims in this thesis were to identify the gyromitrin biosynthesis genes and elucidate their evolutionary history while augmenting our knowledge of lorchel biology and taxonomy. To achieve these goals, I first developed a new analytical protocol for detecting gyromitrin in lorchel mushrooms. I then amassed a robust sampling of *Discinaceae* specimens and sequenced rDNA barcodes to build a multi-locus phylogenetic tree for a preliminary investigation of gyromitrin evolution. I discovered that gyromitrin has a more limited distribution in *Discinaceae* than expected based on anecdotal poisoning data. The position of these taxa in the multi-locus phylogenetic tree suggests gyromitrin has a discontinuous distribution, but low support among basal branches makes this conclusion tentative.

Next, I scaled up my investigations and sequenced dozens of whole genomes, including from specimens up to 60 years old. From these I inferred the first-ever phylogenomic tree for the lorchel family, which allowed me to revise genus- and subgenus-level taxonomy in *Discinaceae*. These genomes provided a unique opportunity to explore not only the evolutionary history of *Discinaceae* but also fundamental aspects of lorchel biology, such as their reproduction and

ecology. Through phylogenomic analyses, I resolved the deep-branching relationships in the family and found that gyromitrin does indeed have a discontinuous distribution. This distribution is suggestive of horizontal gene transfer, although I could not rule out numerous gene loss events or widespread transcriptional silencing of the gyromitrin genes. Taxonomic revisions based on the phylogenomic tree resulted in the recognition of a single genus, *Gyromitra*, composed of 10 subgenera. I demonstrated that *Gyromitra* species are predominantly heterothallic (obligately outcrossing) and that all clades possess genomic signatures of saprotrophism except for the truffloid taxa in subg. *Hydnotrya*, which are ectomycorrhizal.

Lastly, I utilized the phenotypic data and genomes from these investigations to search for candidate gyromitrin biosynthesis genes with bioinformatics analyses. I identified three candidate gyromitrin biosynthesis gene clusters. My top candidate gene cluster was one that included four to five genes with predicted functions compatible with gyromitrin production and prominent phylogenetic incongruence. These gene trees showed the gyromitrin-producing species as more closely related than expected, suggesting horizontal gene transfer may have occurred between the ancestors of the gyromitrin-producing clades. The candidate gyromitrin gene clusters identified here serve as exceptional targets for future validation studies using heterologous expression and gene knockouts. This work showcases a comparative genomics approach for the identification of unknown and noncanonical biosynthesis gene clusters that could be applied in the identification of other secondary metabolite gene clusters.

Chapter 1 Introduction

1.1 Fungal secondary metabolites are diverse and important

Fungi are master alchemists. They have evolved the ability to biosynthesize a vast repertoire of specialized bioactive chemicals called secondary metabolites or natural products. "Secondary" refers to their accessory quality and differentiates them from the "primary" metabolites deemed essential for normal cellular physiology. Secondary metabolites include antibiotics, protective toxins, pigments, and agents of pathogenicity, quorum sensing, and nutrient acquisition, making them indispensable chemical toolkits for fungi in natural environments (Rokas et al., 2020). In some regards, fungal secondary metabolites are indispensable to humans, too; we have long capitalized on secondary metabolites in medicine, and the industrialization of fungal natural products such as the antibiotic penicillin and immunosuppressant cyclosporine have shaped human history in no small way (Keller, 2019).

Filamentous fungi are multidimensional organisms—microbial yet expansive—composed of mycelial networks interned in complex substrates like soil, rotting logs, and leaf duff. Their intimate exposure to multi-species microbial assemblages propels the evolution of natural products, which confer a leg up in the struggle for survival, nutrient acquisition, and reproduction in competitive environments. The importance of secondary metabolites in fungal evolution is captured by a few observations: secondary metabolites are ubiquitous in filamentous fungi; fungi expend substantial resources to produce large quantities of secondary metabolites; and fungal secondary metabolites are rapidly evolving, indicated by their variability even among populations of the same species (Rokas et al., 2020). The elucidation of novel natural products,

their adaptive function, genetic underpinnings, biosynthetic pathways, evolutionary history, and pharmaceutical properties are important and active research topics in mycology and chemistry.

1.2 Genetic paradigms in secondary metabolite biosynthesis

Decades of research on fungal secondary metabolites have resulted in the emergence of a few paradigms regarding the genetics of secondary metabolite biosynthesis (Cano-Prieto et al., 2024). First, secondary metabolite biosynthesis is typically encoded by genes that are physically clustered together in the genome. These colocalized genes are called biosynthesis gene clusters (BGCs) and are thought to colocalize due to coordinated expression, epigenetic regulation, or toxicity avoidance (Cano-Prieto et al., 2024; Pfannenstiel and Keller, 2019; Rokas et al., 2020). This colocalization of biosynthesis genes is also thought to facilitate the relatively high rates of horizontal gene transfer observed for secondary metabolites. In the incorporation of foreign DNA, clustered genes are more likely to be inherited together and intact. As a "complete genetic package" that encodes not only the biosynthesis of a natural product but also resistance to its effects, the horizontal gene transfer of a secondary metabolite BGC can suddenly confer a powerful new trait to aid an organism in its growth (Gonçalves et al., 2024; Rokas et al., 2020; Wisecaver and Rokas, 2015).

A second paradigm is that BGCs usually encode a similar set of proteins: a core synthase that makes the backbone of the nascent secondary metabolite; accessory enzymes that tailor the backbone into the final product; transporters to move the secondary metabolite out of the manufacturing cell; transcription factors to regulate biosynthesis; and proteins that confer resistance to the natural product's effects (Rokas et al., 2020). BGCs belong to different classes based on the type of backbone biosynthesis gene, with the predominant classes being

nonribosomal peptide synthases, ribosomally synthesized and post-translationally modified peptides, polyketide synthases, and terpene synthases (Cano-Prieto et al., 2024).

These paradigms have facilitated the development of bioinformatics tools to mine genomes and predict BGCs (Keller, 2019). We now know that it is common for a filamentous fungus species to possess dozens of secondary metabolites gene clusters. Since there are an estimated 2.5 million fungal species, with only about 150,000 species described, the unexplored chemodiversity of fungi is enormous (Keller, 2019; Niskanen et al., 2023). Few of these described species have been chemically characterized, and among the known fungal secondary metabolites, the vast majority have not been linked to their biosynthesis genes.

1.3 *Gyromitra esculenta* is infamous for its production of the mycotoxin gyromitrin

Gyromitrin is one such mycotoxin with a chemical structure characterized over 50 years ago but whose associated genes remain unknown (List and Luft, 1968). Gyromitrin is infamous due to its presence in the controversially edible species *Gyromitra esculenta*, which is at once a delectable and deadly poisonous mushroom. Eaten raw, the mushrooms can destroy organs and induce seizure due to the presence of gyromitrin. However, if prepared properly via a laborious boiling procedure to remove gyromitrin, they are enjoyed as a delicacy, particularly in Scandinavia where their sale and consumption are heavily regulated (Sitta et al., 2021). This has garnered them the nickname “fungul fugu”, an homage to the *fugu* pufferfish dish that is celebrated in Japan after careful preparation to remove tissues containing the deadly tetrodotoxin (Narahashi, 2008).

Outside of Scandinavia, *Gyromitra esculenta* and related species are best known to foragers of the highly sought morel mushrooms (*Morchella* spp.). *Gyromitra* and *Morchella* can be found in similar habitats at the same time of the year and share a vaguely similar morphology.

This has garnered *Gyromitra* species another common name, “false morel”, conveying a trickster nature. Unfortunately, misidentification of gyromitrin-containing *Gyromitra* spp. as *Morchella* spp. can result in their accidental consumption without special preparation, leading to poisoning (Beug, 2014). The toxicity of *Gyromitra esculenta* has cast a shadow on all *Gyromitra* species as potentially deadly mushrooms to be avoided at all costs. Perhaps this disrepute is the cause for the limited degree of scientific investigation into gyromitrin and the “false morels”. To help remedy their maligned status, I refer to *Gyromitra* mushrooms as “lorchels” throughout this thesis, a name that was also favored by the late mycologist Nancy Smith Weber (Weber, 1995).

1.4 Gyromitrin is a small, volatile mycotoxin with a rare N–N bond moiety

Beyond the gastronomical intrigue of *Gyromitra esculenta*, gyromitrin is an interesting secondary metabolite to study because of its structure. Gyromitrin is a small, volatile substance, which explains why *Gyromitra esculenta* mushrooms can be detoxified via boiling. It is also a hydrazone (specifically acetaldehyde *N*-methyl-*N*-formylhydrazone), making it a member of the rare class of secondary metabolites defined by the presence of an N–N bond (Le Goff and Ouazzani, 2014). N–N bond-containing natural products (N2NPs) represent less than 0.1% of the known secondary metabolites but have an outsized role in pharmaceuticals due to their structural diversity and potent bioactivity (He et al., 2022). While the N–N bond can be readily synthesized under harsh conditions in a test tube, bringing these two nucleophiles together in the cytoplasmic milieu is a difficult feat that has required the evolution of fascinating biosynthetic pathways (Katsuyama and Matsuda, 2020). Enzymes that catalyze N–N bonds were only discovered about ten years ago (Waldman et al., 2017). Since then, tremendous progress has been made in understanding N–N bond biosynthesis in bacteria, but nothing is known of how this bond is formed in fungi. The discovery of a single gene cluster associated with N–N bond formation in

fungi could result in the identification of numerous other N2NP-related BGCs across fungi. For example, another infamous fungal N2NP is the mycotoxin agaritine produced by *Agaricus bisporus*, the common button mushroom sold in grocery stores (Roupas et al., 2010). Identifying the gyromitrin genes could facilitate finding the BGC responsible for agaritine production. In both systems, elucidation of the N2NP-associated genes could also permit the development of mycotoxin-free strains via gene deletions, resulting in the production of mushrooms that are safer for human enjoyment (Schuster and Kahmann, 2019).

1.5 Preview of chapters

In this thesis, I explore questions around the distribution, genetics, and evolution of gyromitrin, as well as the taxonomy, reproduction, and ecology of lorchel mushrooms. In Chapter 2, I discuss my work with natural product chemists to develop a new analytical method for gyromitrin detection. I amassed a large collection of lorchel mushrooms from across North America and surveyed them for gyromitrin. Before this study, only a few species had been tested, and our knowledge of gyromitrin's distribution in *Gyromitra* was largely anecdotal. The results of this work have been published (Dirks et al., 2023). In Chapter 3, I share my phylogenomics research. I sequenced over 60 *Gyromitra* genomes, allowing me to infer a robust phylogenomic tree, resolve the early evolutionary history of the group, and revise taxonomy. Using comparative genomics, I explored fundamental aspects of lorchel biology such as their mating systems and trophic modes based on the presence of mating type loci and carbohydrate-active enzymes. With these genomes in hand, I was ready to look for the gyromitrin biosynthesis genes. In Chapter 4, I dive into the bioinformatics analyses that I used to identify putative gyromitrin BGCs. The top candidate cluster has predicted protein functions that align with our expectations for gyromitrin production and notable patterns of phylogenetic incongruence

suggestive of horizontal gene transfer among lorchel mushrooms. Finally, in Chapter 5, I provide some concluding remarks on next steps for the field of gyromitrin research.

Although not presented here, my research uncovered numerous undescribed lorchel species, two of which have been formally described, and ongoing work will continue to describe new species and resolve the taxonomy of various clades (Dirks et al., 2024; Miller et al., 2022). In addition, I invested considerable time and resources to develop a genetic transformation protocol for *Gyromitra* species during my thesis. While not yet fully established, this work will need to continue to enable the validation of the gyromitrin genes via gene knockout studies. Far from being the end of the story, identification of the gyromitrin genes opens exciting new avenues of inquiry into the evolution of gyromitrin and the biosynthesis of fungal N2NPs in general.

References

Beug, M.W., 2014. False morels—age-old questions of edibility: a primer by Michael Beug. FUNGI 7, 28–31.

Cano-Prieto, C., Undabarrena, A., de Carvalho, A.C., Keasling, J.D., Cruz-Morales, P., 2024. Triumphs and challenges of natural product discovery in the postgenomic era. Annu. Rev. Biochem. 93, 20.1-20.35. https://doi.org/10.1146/annurev-biochem-032620-104731

Dirks, A.C., Beug, M.W., Methven, A.S., Miller, A.N., James, T.Y., 2024. *Gyromitra persicula sp. nov.* (*Discinaceae*, *Pezizales*), a novel discoid *Gyromitra* from Northwestern North America. Phytotaxa 644, 49–55. https://doi.org/10.11646/phytotaxa.644.1.7

Dirks, A.C., Mohamed, O.G., Schultz, P.J., Miller, A.N., Tripathi, A., James, T.Y., 2023. Not all bad: gyromitrin has a limited distribution in the false morels as determined by a new ultra high-performance liquid chromatography method. Mycologia 115, 1–15. https://doi.org/10.1080/00275514.2022.2146473

Gonçalves, C., Hittinger, C.T., Rokas, A., 2024. Horizontal gene transfer in fungi and its ecological importance, in: Hsueh, Y.-P., Blackwell, M. (Eds.), Fungal Associations, The Mycota. Springer International Publishing, Cham, pp. 59–81. https://doi.org/10.1007/978-3-031-41648-4_3

He, H.-Y., Niikura, H., Du, Y.-L., Ryan, K.S., 2022. Synthetic and biosynthetic routes to nitrogen–nitrogen bonds. Chem. Soc. Rev. 51, 2991–3046. https://doi.org/10.1039/C7CS00458C

Katsuyama, Y., Matsuda, K., 2020. Recent advance in the biosynthesis of nitrogen–nitrogen bond–containing natural products. Current Opinion in Chemical Biology, Mechanistic Biology ● Energy 59, 62–68. https://doi.org/10.1016/j.cbpa.2020.05.002

Keller, N.P., 2019. Fungal secondary metabolism: regulation, function and drug discovery. Nat. Rev. Microbiol. 17, 167–180. https://doi.org/10.1038/s41579-018-0121-1

Le Goff, G., Ouazzani, J., 2014. Natural hydrazine-containing compounds: biosynthesis, isolation, biological activities and synthesis. Bioorganic & Medicinal Chemistry 22, 6529–6544. https://doi.org/10.1016/j.bmc.2014.10.011

List, P.H., Luft, P., 1968. Gyromitrin, das gift der frühjahrslorchel. 16. Mitt. über pilzinhaltsstoffe. Archiv Der Pharmazie 301, 294–305. https://doi.org/10.1002/ardp.19683010410

Miller, A.N., Dirks, A.C., Filippova, N., Popov, E., Methven, A.S., 2022. Studies in *Gyromitra* II: cryptic speciation in the *Gyromitra gigas* species complex; rediscovery of *G. ussuriensis* and *G. americanigigas* sp. nov. Mycol Progress 21, 81. https://doi.org/10.1007/s11557-022-01832-x

Narahashi, T., 2008. Tetrodotoxin. Proceedings of the Japan Academy, Series B 84, 147–154. https://doi.org/10.2183/pjab.84.147

Niskanen, T., Lücking, R., Dahlberg, A., Gaya, E., Suz, L.M., Mikryukov, V., Liimatainen, K., Druzhinina, I., Westrip, J.R.S., Mueller, G.M., Martins-Cunha, K., Kirk, P., Tedersoo, L., Antonelli, A., 2023. Pushing the frontiers of biodiversity research: unveiling the global diversity, distribution, and conservation of fungi. Annu. Rev. Environ. Resour. 48, 149–176. https://doi.org/10.1146/annurev-environ-112621-090937

Pfannenstiel, B.T., Keller, N.P., 2019. On top of biosynthetic gene clusters: how epigenetic machinery influences secondary metabolism in fungi. Biotechnology Advances, Microbial Engineering Biotechnologies 37, 107345. https://doi.org/10.1016/j.biotechadv.2019.02.001

Rokas, A., Mead, M.E., Steenwyk, J.L., Raja, H.A., Oberlies, N.H., 2020. Biosynthetic gene clusters and the evolution of fungal chemodiversity. Nat. Prod. Rep. 37, 868–878. https://doi.org/10.1039/c9np00045c

Roupas, P., Keogh, J., Noakes, M., Margetts, C., Taylor, P., 2010. Mushrooms and agaritine: a mini-review. Journal of Functional Foods 2, 91–98. https://doi.org/10.1016/j.jff.2010.04.003

Schuster, M., Kahmann, R., 2019. CRISPR-Cas9 genome editing approaches in filamentous fungi and oomycetes. Fungal Genetics and Biology 130, 43–53. https://doi.org/10.1016/j.fgb.2019.04.016

Sitta, N., Davoli, P., Floriani, M., Suriano, E., 2021. Guida ragionata alla commestibilità dei funghi, 1st ed. Regione Piemonte, Italy.

Waldman, A.J., Ng, T.L., Wang, P., Balskus, E.P., 2017. Heteroatom−heteroatom bond formation in natural product biosynthesis. Chem. Rev. 117, 5784–5863. https://doi.org/10.1021/acs.chemrev.6b00621

Weber, N.S., 1995. A Morel Hunter's Companion: A Guide to True and False Morels. Thunder Bay Press, Lansing, Michigan.

Wisecaver, J.H., Rokas, A., 2015. Fungal metabolic gene clusters—caravans traveling across genomes and environments. Front. Microbiol. 6, 161. https://doi.org/10.3389/fmicb.2015.00161

Chapter 2 Not All Bad: Gyromitrin Has a Limited Distribution in the False Morels as Determined by a New Ultra High-Performance Liquid Chromatography Method

Abstract

Gyromitrin (acetaldehyde *N*-methyl-*N*-formylhydrazone) and its homologs are deadly mycotoxins produced most infamously by the lorchel (also known as false morel) *Gyromitra esculenta*, which is paradoxically consumed as a delicacy in some parts of the world. There is much speculation about the presence of gyromitrin in other species of the lorchel family (*Discinaceae*), but no studies have broadly assessed its distribution. Given the history of poisonings associated with the consumption of *G. esculenta* and *G. ambigua*, we hypothesized that gyromitrin evolved in the last common ancestor of these taxa and would be present in their descendants with adaptive loss of function in the nested truffle clade, *Hydnotrya*. To test this hypothesis, we developed a sensitive analytical derivatization method for the detection of gyromitrin using 2,4-dinitrobenzaldehyde as the derivatization reagent. In total, we analyzed 66 specimens for the presence of gyromitrin over 105 tests. Moreover, we sequenced the nuc rDNA internal transcribed spacer region ITS1-5.8S-ITS2 (ITS barcode) and nuc 28S rDNA to assist in species identification and to infer a supporting phylogenetic tree. We detected gyromitrin in all tested specimens from the *G. esculenta* group as well as *G. leucoxantha*. This distribution is consistent with a model of rapid evolution coupled with horizontal transfer, which is typical for secondary metabolites. We clarified that gyromitrin production in *Discinaceae* is both discontinuous and more limited than previously thought. Further research is required to elucidate the gyromitrin biosynthesis gene cluster and its evolutionary history in lorchels.

2.1 Introduction

Gyromitrin is a polar, water-soluble, and volatile mycotoxin produced by *Gyromitra esculenta*, a distinctive brain-like mushroom that is consumed as a delicacy, particularly in Scandinavia (Benjamin 2020; Härkönen 1998; Sitta et al. 2021; Svanberg and Lindh 2019). Structurally, gyromitrin refers to the acetaldehyde *N*-methyl-*N*-formylhydrazone (**1**), and it occurs with eight higher aldehyde homologs (**2–9**) present in smaller quantities (List and Luft 1968a; Pyysalo 1975; Pyysalo and Niskanen 1977) (Figure 2.1). Spontaneously at room temperature, upon heating, and especially in acidic environments such as the stomach, gyromitrin hydrolyzes into its aldehyde component and *N*-methyl-*N*-formylhydrazine (**10**), which then loses formaldehyde to yield monomethylhydrazine (**11**), an ingredient in some rocket propellants (List and Luft 1968b; Monteith 2020; Pyysalo et al. 1978) (Figure 2.2). Gyromitrin's volatility, solubility, and reactivity explain how *G. esculenta*, which contains 50–300 mg/kg of gyromitrin per fresh mushroom, can be consumed without ill effect (Michelot and Toth 1991; Pyysalo 1976; Pyysalo and Niskanen 1977). By boiling the mushrooms twice, rinsing them after each boil, and changing the cooking liquid in between, more than 99% of the gyromitrin is released to render the mushroom safe to eat without acute toxicity (Pyysalo 1976; Pyysalo and Niskanen 1977).

Raw or undercooked *G. esculenta* mushrooms are poisonous, primarily due to **10** and **11** formed via acid hydrolysis in the gastric environment. Symptoms usually appear from 5 to 12 h after exposure. Most people poisoned by gyromitrin only experience gastrointestinal issues involving vomiting, abdominal pain, and diarrhea; in more severe cases, a victim suffers cytolytic hepatitis and jaundice (ArłukowiczGrabowska et al. 2019; Mäkinen et al. 1977; Michelot and Toth 1991; Wright et al. 1978). Gyromitrin poisoning can also entail neurological

symptoms such as vertigo, fatigue, and tremor, which can progress to seizure in the worst scenarios (Karlson-Stiber and Persson 2003). This is hypothesized to be a result of the binding of the gyromitrin hydrolytic products to pyridoxine (vitamin B6), inhibiting the enzymes that produce γ-aminobutyric acid (GABA) and serotonin. Death, though rare, is caused by hepatic coma as unstable metabolic intermediates and methyl free radicals destroy the liver (Arłukowicz-Grabowska et al. 2019; Karlson-Stiber and Persson 2003; Michelot and Toth 1991). Gyromitrin may also have long-term toxic effects, including increased rates of cancer (Toth and Patil 1980; Toth and Nagel 1978). The genotoxic potential of gyromitrin informs the emerging hypothesis of its role in the development of neurodegenerative diseases such as sporadic amyotrophic lateral sclerosis (ALS) (Lagrange et al. 2021; Spencer 2020; Spencer and Kisby 2021; Spencer and Palmer 2021).

Beyond *Gyromitra esculenta*, whose toxicity has been well established, there is limited evidence regarding the toxicity of other *Gyromitra* species. At the turn of the 20th century, renowned mycophagist Charles McIlvaine regarded *G. brunnea*, *G. caroliniana*, and "*G. curtipes*" (*G. gigas* group) as "esculent" (McIlvaine and Macadam 1902). In 1976, *Gyromitra* taxonomist Harri Harmaja discussed a poisoning event in northern Sweden involving *G. ambigua* and concluded from the available evidence that this species contained dangerous quantities of gyromitrin, in contrast to the nontoxic *G. infula* (Harmaja 1976). Harmaja also cited a report on four *Gyromitra* poisoning events in the Czech Republic, one of which supposedly involved *G. gigas*. However, this identification was based solely on its purported hefty stature, and the specimen was never examined (Kubička 1966). A pattern emerges in older field guides whereby European *G. gigas* is flagged as toxic (references in Viernstein et al. 1980; Weber 1995), but species in the North American *G. gigas* group—now known to correspond to *G.*

korfii, *G. montana*, and *G. americanigigas* (Miller et al. 2022, 2020)—are listed as edible (references in Weber 1995). One such source, Tylutki's *Mushrooms of Idaho and the Pacific Northwest* (1979), describes *G. montana* (listed as *G. gigas*) as a popular and delicious edible of the Rocky Mountains. From data gathered by the North American Mycological Association, *Gyromitra* spp. were the culprit behind about 4% of North American mushroom poisoning events involving humans between 1973 and 2005 (Beug 2014; Beug et al. 2006). *Gyromitra esculenta* was always involved when organ failure occurred. *Gyromitra brunnea* and *G. montana* resulted in poisonings as well, although symptoms appeared no more severe than those sometimes caused by other widely consumed mushrooms. In 2020, *G. venenata*, a species closely related to *G. esculenta*, was described from China after its consumption caused four people to go to the emergency room (Li et al. 2020). Finally, Lagrange et al. (2021) associated a hot spot of ALS in the French Alps to the consumption of lorchels, specifically *G. gigas*, although this species was perhaps prematurely implicated given that other unidentified *Gyromitra* species were also found stored in the residents' homes.

Many different analytical chemistry techniques have been employed over the years to detect and/or quantify gyromitrin. After purifying and solving the chemical structure of gyromitrin, List and Luft (1968a, 1968b) developed spot tests involving acid hydrolysis of gyromitrin and subsequent derivatization of the hydrolytic products with various chromophores, as well as thin-layer chromatography (TLC) and potassium iodate titration methods. Pyysalo and colleagues developed gas chromatography (GC) methods to quantify gyromitrin hydrazone analogs or benzaldehyde derivatives of **10** and **11** (Pyysalo and Niskanen 1977; Pyysalo et al. 1978). Viernstein et al. (1980) also used GC along with an external standard calibration to quantify gyromitrin in the ether extracts of pressed mushrooms. Andary et al. (1984, 1985)

separated ethanol extracts with TLC and sprayed *p*-dimethylaminocinnamaldehyde to form a red fluorophore quantifiable with spectrofluorometry. Finally, Arshadi et al. (2006) developed a GC–mass spectrometry (MS) analytical method involving the derivatization of acid-hydrolyzed ethanol extracts with pentafluorobenzoyl chloride. Despite all these methods, lacking from the literature is a method that employs high-performance liquid chromatography (HPLC) or ultra high-performance liquid chromatography (UHPLC) coupled with diode array detector (DAD), which are popular chromatographic techniques in modern chemistry laboratories.

Given this variety of published analytical methods, data on the presence of gyromitrin in lorchels are surprisingly scarce. In the first gyromitrin analysis outside of *G. esculenta*, Viernstein et al. (1980) detected small quantities of gyromitrin in two out of three European *G. gigas* specimens (0.05 and 0.74 mg/kg fresh specimen) but not in the single *G. fastigiata* specimen tested. Andary et al. (1985) did not detect gyromitrin in *Gyromitra perlata* or in *Morchellaceae* taxa but claimed to detect it in a curious set of species belonging elsewhere in the *Pezizomycetes* as well as the *Leotiomycetes*. Chemists at a toxicology laboratory in Michigan reported in a clinical toxicology conference abstract that *G. caroliniana* contained "minute" amounts of free **11** (Liang et al. 1998). On the other hand, in her popular morel hunting guide, Weber (1995) cited unpublished spot test data from Dr. Kenneth Cochran indicating the absence of gyromitrin in *G. caroliniana* as well as *G. brunnea* and *G. korfii*.

With the growing popularity of lorchel consumption (as evidenced by online forums such as the Facebook group False Morels Demystified), there is an urgent need to refine our understanding of which taxa contain the gyromitrin mycotoxin. Taking into consideration the available toxicity evidence, we hypothesized that gyromitrin evolved in the last common ancestor of *G. esculenta* and *G. ambigua*, with a loss of function in the closely related *Hydnotrya*

clade due to the negative fitness consequences that mycotoxin production could have on hypogeous fungi relying on mycophagy for spore dispersal. To evaluate the distribution of gyromitrin in *Discinaceae*, we developed a simple, sensitive analytical method involving in situ acid hydrolysis of gyromitrin and chemical derivatization of its hydrolytic products, **10** and **11**, using 2,4-dinitrobenzaldehyde into Schiff bases **12** and **13** (Figure 2.2). This derivatizing agent was chosen due to its proven efficiency in detection of related hydrazine-containing metabolites (Mohamed et al. 2018, 2021). Since the amount of gyromitrin is known to differ within the same species by environment and genotype (Andary et al. 1985; Marjatta and Pyysalo 1978), we aimed for broad taxon sampling and qualitative evaluation of gyromitrin presence rather than attempting to make claims about gyromitrin levels among different specimens within a species.

2.2 Materials and methods

2.2.1 Specimen acquisition

Ascocarp specimens were collected fresh by A.C.D. or donated by mycologists and community scientists from across North America. Some collections were split, with one half preserved at −80 C and the other half dried in an electric dehydrator. All freshly collected dried specimen vouchers were deposited at the University of Michigan fungarium, except for a few that were entirely consumed through the gyromitrin detection method or did not dry properly. A subset of specimens was successfully cultured by sampling a piece of hymenium, placing it in sterile water, agitating the tissue to release ascospores, plating serial dilutions of the ascospore liquid on potato dextrose agar (PDA) and malt extract agar (MEA) with penicillin and streptomycin, and isolating single germinating ascospores onto PDA and MEA without antibiotics. In addition, a set of *Gyromitra* cultures from a recent study on *Pezizales* fungi were provided by Dr. A. Elizabeth Arnold from the living collection at the Robert L. Gilbertson

Mycological Herbarium at the University of Arizona (Healy et al. 2022). Two *Gyromitra* cultures whose genomes have been sequenced were also purchased from the Westerdijk Fungal Biodiversity Institute CBS culture collection. Older dried ascocarp specimen vouchers were acquired as loans from the University of Michigan (MICH), University of Florida (FLAS), and University of Arizona (ARIZ) fungaria. Specimen vouchers and cultures tested for gyromitrin for this study are listed in Table 2.1.

2.2.2 Molecular data

DNA was extracted from dried ascocarps and fresh cultures using the 2× CTAB (cetyltrimethylammonium bromide) method (Porter et al. 2011). Ascocarp tissue (approximately 0.25 cm^2 piece of hymenium) or hyphae (approximately 1 cm^2 scraped from mycelium growing on cellophane) were placed into a 1.5-mL microcentrifuge tube with 500 µL cell lysis solution (2× CTAB extraction buffer: 2% cetyltrimethylammonium bromide, 1.4 M NaCl, 50 mM Tris, 10 mM Na2EDTA, pH 8). Each sample was ground with a Kontes plastic pestle (DWK Life Sciences, Stoke-on-Trent, UK) for up to 2 min or until the tissue appeared homogenized, vortexed, and incubated in a water bath at 65 C for 60 min. Extraction proceeded with the addition of an equal volume of chloroform–isoamyl alcohol (24:1), centrifugation at 13 000 RPM (15 871 RCF) for 12 min, and transfer of the upper aqueous layer to a new 1.5-mL microcentrifuge tube. This extraction procedure was repeated once more, and the final aqueous phase was precipitated with two-thirds volume ice cold isopropyl alcohol, mixed by inversion, and placed in a −20 C freezer overnight. The following day, the DNA was pelleted by centrifugation for 7 min. The alcohol supernatant was discarded, and the pellet rinsed with 1 mL cold 70% ethanol. The DNA was dried in a laminar flow hood and resuspended in 50 µL distilled water.

For each extraction, we attempted to amplify separately the nuc rDNA internal transcribed spacer region ITS1-5.8S-ITS2 (ITS barcode) using the primers ITS1f and ITS4 (Gardes and Bruns 1993; White et al. 1990) and the D1–D2 domains of nuc 28S rDNA using the primers LR0R and LR5 (Moncalvo et al. 2000; White et al. 1990). If ITS amplification failed, we attempted to amplify the nuc rDNA ITS2 region using the primers ITS3 and ITS4 (White et al. 1990). PCR amplifications were carried out using the GoTaq Green Master Mix (Promega, Madison, Wisconsin) in a reaction volume of 12.5 μL, with 5 μL of a 1:20 dilution of the DNA extract as a template. Reactions were completed on an Eppendorf Mastercycler Pro S model 6325 thermal cycler (Hamburg, Germany). For ITS and ITS2, the thermal cycler parameters were as follows: 2 min initial denaturation at 94 C; 35 cycles of 30 s at 94 C, 30 s at 55 C, and 30 s at 72 C; and a 10 min final extension at 72 C. For 28S, they were: 5 min initial denaturation at 95 C; 40 cycles of 30 s at 95 C, 15 s at 52 C, and 1 min at 72 C; and a 10 min final extension at 72 C. Successful PCRs were enzymatically cleaned with ExoSAP-IT (Applied Biosystems, Waltham, Massachusetts) and submitted to Azenta Life Sciences for Sanger sequencing (Azenta Life Sciences, South Plainfield, New Jersey) using the PCR primers. Forward and reverse Sanger sequences were assembled with Geneious Prime (Dotmatics, Boston, Massachusetts) and submitted to GenBank (ON693581–ON693677) (Table 2.1).

2.2.3 Phylogenetic analyses

Published *Discinaceae* ITS and 28S sequences were downloaded from GenBank with Entrez Direct (Kans 2020) (Supplementary Table 2.1). *Discinaceae* ITS and 28S sequences were aligned separately with MUSCLE as implemented in SeaView 5.0.4 (Edgar 2004; Gouy et al. 2010). Ambiguous regions of the alignments were trimmed with Gblocks 0.91b using less stringent parameters (Castresana 2000; Talavera and Castresana 2007) and concatenated with

catfasta2phyml 1.1.0 (github.com/nylander/catfasta2phyml). A maximum likelihood (ML) analysis was conducted on the concatenated alignment with RAxML 8.2.11 using the GTR+Gamma+I model of substitute evolution and 1000 bootstrap replicates (Abadi et al. 2019; Stamatakis 2014). Clades with bootstrap values ≥ 70% were considered significant and strongly supported (Hillis and Bull 1993). Bayesian analyses were performed with the same alignment under the above model using MrBayes 3.2.7 on the CIPRES Science Gateway 3.3 portal (Miller et al. 2011; Ronquist et al. 2012). The Bayesian analyses lasted until the average standard deviation of split frequencies was below 0.01, with trees saved every 1000 generations and burn-in set at 25%. Bayesian posterior probabilities (BPPs) were determined from a consensus tree using Geneious Prime. Clades with BPP ≥ 95% were considered significant and strongly supported (Alfaro et al. 2003; Larget and Simon 1999). The best ML tree was visualized and rooted with *Morchella esculenta* (*Morchellaceae*) as the outgroup in FigTree 1.4.4 (github.com/rambaut/figtree). The alignments (https://doi.org/10.6084/m9.figshare. 21313596.v1) and phylogenetic tree (https://doi.org/10. 6084/m9.figshare.21313599.v1) were deposited in Figshare.

2.2.4 Gyromitrin analytical method development

2.2.4.1 General experimental details

2,4-Dinitrobenzaldehyde (2,4-DNB) and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich (D193607 and 302031; St. Louis, Missouri). Authentic gyromitrin standard was purchased from Toronto Research Chemicals (G931900; Toronto, Canada). HPLC-grade solvents were used in the gyromitrin extractions. Solvents used for preparative HPLC and analytical UHPLC-MS were of HPLC-grade and Optima LC-MS grade, respectively, and supplied by Fisher Chemical (Waltham, Massachusetts). Prior to use, the solvents were

filtered/degassed through a 0.45μm polytetrafluoroethylene membrane. Preparative HPLC was performed using Shimadzu (Kyoto, Japan) LC-20AT HPLC instruments with corresponding detectors, fraction collectors, and software (Phenomenex [Torrance, California] Luna Phenyl-Hexyl column, 21.2 mm × 25 cm, 5 μm, 20 mL/min, isocratic elution with 90% H2O/MeCN for 2 min followed by gradient elution from 90% H2O/MeCN to 100% MeCN over 30 min then isocratic elution with 100% MeCN for 5 min). UHPLC coupled with diode array detection and electrospray ionization mass spectrometry (UHPLC-DAD-ESIMS) was performed using the Shimadzu LC-20AD Separations Module equipped with a Shimadzu LCMS-2020 Series mass detector in both positive and negative ion modes (Phenomenex Kinetex C8 1.7 μm 100 Å column, 50 mm × 2.1 mm, eluting with 0.4 mL/min of isocratic 90% H2O/MeCN for 1 min followed by gradient elution to 100% MeCN with isocratic 0.1% formic acid modifier over 6 min, at 210, 254, 280, and 370 nm).

2.2.4.2 Synthesis of 2,4-dinitrobenzaldehyde Schiff base (13) reference compound

An aliquot of 2,4-DNB (94 mg, 0.48 mmol) dissolved in 4 mL ethanol was treated with **11** (120 mg, 2.67 mmol) and stirred at room temperature for 2 h. Afterward, the reaction mixture was quenched with water (20 mL), extracted with 15 mL dichloromethane (CH2Cl2) two times, and the combined organic layer dried over anhydrous sodium sulfate (Na2SO4), concentrated in vacuo, and purified by preparative reverse-phase HPLC to yield pure **13** (9 mg). Purity was checked by UHPLC-DADESIMS (Supplementary Figure 2.1). A seven-point calibration curve with solutions of reference compound **13** (0.39–50 μg/mL) in 50% H2O/MeCN was established by UHPLC-DAD analysis of an aliquot (10 μL) of each concentration in duplicate (Supplementary Figure 2.2). It is noteworthy that higher concentrations (100 μg/mL) of **13** did not fit into the calibration curve trendline due to saturation of the DAD detector.

2.2.4.3 Detection of gyromitrin hydrazine hydrolytic products using 2,4-DNB

A series of gyromitrin standard solutions (5, 1, 0.5, 0.25, 0.1, 0.05, and 0.01 mg/mL) were prepared in 50% H2O/MeCN. An aliquot (10 µL) of each gyromitrin solution was transferred to a glass vial containing 50% H2O/ MeCN (400 µL) and treated with an aliquot (40 µL) of freshly prepared stock solution (5 mg/mL in MeCN) of 2,4-DNB and an aliquot (50 µL) of freshly prepared stock solution of 10% aqueous TFA. The reaction mixture was incubated at 40 C. Aliquots (10 µL) from each reaction mixture were analyzed by UHPLC-DAD at regular time intervals (0, 2, 5, 8, 13, 18, and 24 h) to detect gyromitrin hydrazine hydrolytic product derivatives with 2,4-DNB (Supplementary Figure 2.3).

2.2.4.4 Detection of gyromitrin in Gyromitra venenata

Aliquots (50, 10, and 1 mg) of powdered *Gyromitra venenata* (MICH 352032) were transferred to glass vials containing 50% H2O/MeCN (820 µL) and treated with an aliquot (80 µL) of freshly prepared stock solution (5 mg/mL in MeCN) of 2,4-DNB and an aliquot (100 µL) of freshly prepared stock solution of 10% aqueous TFA. The reaction mixture was sonicated for 20 s and incubated at 40 C. Aliquots (10 µL) from each reaction mixture were analyzed by UHPLC-DAD at regular time intervals (0, 6, 13, and 24 h) to detect gyromitrin hydrazine hydrolytic product derivatives with 2,4-DNB. The reactions were performed in triplicate. Samples prepared by extracting the same amount of *G. venenata* with 50% H2O/MeCN (1000 µL) without the addition of 2,4-DNB and TFA were used as negative controls.

2.2.4.5 Detection of gyromitrin in different samples

Gyromitrin analysis of samples proceeded as for *G. venenata* except for the following differences. Samples consisted of an aliquot of powdered ascocarp hymenium (ranging from 5 to 100 mg, depending on available starting material) or a piece of mycelium scraped from a plate

with cellophane (half a plate). An aliquot (not in triplicate) of the reaction mixture (10 μL) was analyzed after 13–18 h of incubation at 40 C. An equivalent amount of tissue was extracted with 50% H2O/MeCN (1000 μL) without the addition of 2,4-DNB and TFA as a negative control.

2.3 Results

UHPLC-DAD analysis of different concentrations of gyromitrin standard treated with 2,4-DNB and TFA with varying incubation times at 40 C revealed the need for 13–18 h of incubation to complete the gyromitrin derivatization reaction to the monomethylhydrazine Schiff base (**13**). Our methodology was able to detect even trace amounts of gyromitrin to the extent of 10 ng in samples (Supplementary Figure 2.4). With a standard calibration curve of synthetic **13** (Supplementary Figure 2.2), the reaction recovery percentage was calculated based on peak areas of **13** obtained from derivatization of different gyromitrin standard concentrations (Supplementary Table 2.2). The peak area of the monomethylhydrazine Schiff base (**13**) was plotted against the gyromitrin standard concentration, showing a linear relationship (Supplementary Figure 2.5). This correlation could be used for the quantification of gyromitrin in fungal samples. Equipped with a sensitive gyromitrin analytical method, aliquots of dried *Gyromitra venenata* (MICH 352032) tissue were processed with different incubation times to determine the hydrolysis of gyromitrin present in ascocarp samples (Figure 2.3). The peak area of Schiff base **13** remained relatively constant between 13 and 24 h of extraction regardless of the amount of starting material (Supplementary Figure 2.6). A large amount of gyromitrin (2598 mg/kg dried ascocarp) was detected in these aliquots of *G. venenata*.

We analyzed 66 specimens as dried ascocarps, freshly frozen ascocarps, living cultures, or a combination thereof, resulting in 105 individual tests (Supplementary Table 2.3). All taxa outside the lorchel family (eight specimens) tested negative for gyromitrin. The majority of

Discinaceae taxa also tested negative for gyromitrin. Both *G. leucoxantha* specimens tested positive, with MICH 352087 containing 55 mg/kg of gyromitrin per dried ascocarp (MICH 25407 was not quantified). The chromatograms also included a strong peak between **12** and **13**, which was absent from the negative control (extracts without the addition of 2,4-DNB and TFA) and likely corresponds to an unidentified Schiff base derivative produced by *G. leucoxantha* (Figure 2.4). All four taxa in the *G. esculenta* group (*G. antarctica*, *G. esculenta*, *G. splendida*, and *G. venenata*) tested positive for gyromitrin, with a few caveats. First, gyromitrin levels generally decreased with age of the dried ascocarp, which could be attributed to the volatile property of gyromitrin. Still, even the oldest voucher tested (*G. venenata* [MICH 25555] from 1883) had trace levels of gyromitrin in the chromatogram (Figure 2.5). Second, cultures growing on PDA tested positive, but gyromitrin was not detected in cultures growing on MEA (Supplementary Table 2.3). Third, the two *G. esculenta* group ascocarps from the western United States contained relatively small amounts of gyromitrin. *Gyromitra splendida* (MICH 352089) from Washington State contained 92 mg/kg of gyromitrin per dried ascocarp (Figure 2.4). *Gyromitra venenata* (MICH 1304) from Idaho also appeared to have less gyromitrin than expected, although this is based solely on the relative height of peaks in the chromatogram (Figure 2.5d). Evidence for the toxicity and gyromitrin content of *Gyromitra* species is summarized in Table 2.2.

We had high success in sequencing rDNA from *Discinaceae* specimens, including from old fungarium vouchers—56/58 specimens (97%) had either the entire ITS or ITS2 region successfully sequenced and 54/58 specimens (93%) had the 28S rDNA barcode successfully sequenced. *Hydnotrya* specimens proved to be the most challenging to sequence, perhaps because of the contamination associated with their subterranean habitat. The concatenated ITS

and 28S rDNA alignment contained 86 sequences with 1402 characters and 571 distinct patterns, 387 of which were parsimony informative. The maximum likelihood phylogenetic tree for the concatenated data set allowed us to validate species determinations and visualize gyromitrin content in a phylogenetic context (Figure 2.6). Across *Discinaceae*, a general absence of gyromitrin was punctuated by its discontinuous presence in two separate groups, the *G. esculenta* and *G. leucoxantha* clades. Due to the use of just two loci, bootstrap support for basal nodes was generally low and the evolutionary relationships between well-defined clades remain ambiguous.

2.4 Discussion

Gyromitrin determination using our newly developed UHPLC analytical method indicated that gyromitrin production is concentrated in the *G. esculenta* group, represented here by *G. antarctica*, *G. esculenta*, *G. splendida*, and *G. venenata* (although taxonomic names will likely change when the group is treated in a modern systematic revision). Contrary to expectations, we failed to detect gyromitrin in *G. ambigua* or *G. infula*, but to our surprise we detected gyromitrin in *G. leucoxantha*. Accepting at face value earlier reports of gyromitrin production by *G. gigas* sensu stricto (Viernstein et al. 1980), six loss events are required to explain this pattern according to our phylogeny (*G. melaleucoides* clade, last common ancestor of *G. caroliniana* and *G. ancilis* clades, and various losses in the *G. gigas* clade). A similar number of loss events, or one horizontal transfer and four loss events, would be required to explain this distribution according to previously published phylogenetic trees (Miller et al. 2022, 2020; Wang and Zhuang 2019). On the other hand, only two horizontal transfer events could result in the observed distribution, making this the most parsimonious hypothesis. However, repeated occurrences of genetic loss of function have been documented elsewhere (Morris et al.

2012; Patron et al. 2007), so we are unable to infer a definitive history of the evolution of gyromitrin in *Discinaceae*.

The distribution of gyromitrin observed here is consistent with a model of rapid evolution coupled with horizontal transfer, which is typical for secondary metabolites such as gyromitrin (Rokas et al. 2020). In fungi, secondary metabolites—small, bioactive molecules not required for growth but important for interspecific interactions—are usually produced by enzymes encoded by genes physically clustered in the genome as a biosynthesis gene cluster (BGC). BGCs can experience strong selective pressure due to their importance in ecological interactions that often results in a narrow taxonomic distribution as well as intraspecific variation in the ability to produce a given secondary metabolite (Rokas et al. 2020; Yancey et al. 2022). Horizontal gene transfer is also a powerful evolutionary force that has led to the broad and phylogenetically disjunct expression of mycotoxins such as amanitin, psilocybin, and epipolythiodioxopiperazines (Luo et al. 2022; Patron et al. 2007; Rokas et al. 2020; Van Court et al. 2022). Therefore, a disjointed distribution of gyromitrin is not unprecedented.

The UHPLC-DAD analytical method presented in this paper is an accessible method that could facilitate further investigations into the gyromitrin mycotoxin, including determinations in human body fluids. Although other methods such as spot tests and TLC are simpler, they have low detection sensitivity, accuracy, and resolution. Thus, earlier reported conclusions of Andary et al. (1985) that taxonomically disparate ascomycete fungi contain gyromitrin should be reinterpreted. These taxa likely contain hydrazines or natural products that hydrolyze to form hydrazines, but not necessarily gyromitrin. Hydrazines belong to a class of chemicals defined by their nitrogen-nitrogen bond (N–N bond), of which only about 200 have been characterized to date, or less than 0.1% of the total known natural products (Blair and Sperry 2013; Le Goff and

Ouazzani 2014). The frequencies at which hydrazine derivatives were detected by Andary et al. (1985) suggest that they might be more prevalent in fungi than expected and indicate that *Pezizomycetes* and *Leotiomycetes* fungi should be broadly screened for secondary metabolites containing an N–N bond.

2.4.1 On the edibility of lorchels

Given that many people are interested in consuming lorchels and that our research will undoubtedly influence the conversation around their edibility, we believe that it is pertinent that we weigh in on the matter. Although the methods to eat gyromitrin-containing lorchels without acute poisoning exist, the dangers of improperly preparing them and their potential link to neurodegenerative disease lead us to conclude that the *Gyromitra esculenta* group should never be consumed. The same is true for *G. leucoxantha*, which we demonstrated to have the capacity to produce gyromitrin, albeit in relatively small quantities. As far as we are aware, there is no culture of intentionally eating this species, and it would be best for it to stay that way (Tylutki 1979; Weber 1995). As a potential producer of gyromitrin, we recommend that individuals avoid European *G. gigas* as well, at least until its status is clarified with contemporary analysis (Viernstein et al. 1980). We acknowledge that people are likely to continue ingesting some of these species, in which case it is crucial to cook them according to the safety guidelines established for their commercial sale in Europe, such as that of the Finnish Food Safety Authority (Sitta et al. 2021).

Outside of these taxa, we believe that one's decision to consume lorchels is a matter of personal preference based on an informed consideration of the risks involved. In favor of some North American lorchels, there is already an established and widespread culture of consuming *G. brunnea*, *G. caroliniana*, *G. korfii*, *G. montana*, and presumably *G. americanigigas*, and no

evidence of extraordinary toxicity or gyromitrin production—except for a curious, unpublished report of "minute" quantities of free **11** in *G. caroliniana* that our data did not corroborate (Beug 2014; Beug et al. 2006; Liang et al. 1998; Weber 1995). With these lorchels one should adhere to the central tenants of all wild mushroom foraging: positively identify the species, only consume fresh mushrooms (not mushrooms with signs of insect damage, moldy growth, or putrescence), thoroughly cook the mushrooms, try a little before trying a whole meal to evaluate idiosyncratic responses such as an allergic reaction, and be careful with overindulgence.

In opposition to their consumption, we likely have not established the full distribution pattern of gyromitrin in *Discinaceae*. Intraspecific variation in the amount of gyromitrin produced by *G. gigas* resulted in one specimen having no detectable levels (Viernstein et al. 1980), meaning that even some species that tested negative in this study might have the capacity to produce gyromitrin. Furthermore, *Gyromitra* species have up to a dozen biosynthesis gene clusters whose corresponding secondary metabolites are entirely uncharacterized (MycoCosm; Grigoriev et al. 2014). Given that each species could produce a unique set of secondary metabolites, there are likely hundreds of undescribed chemicals with potentially powerful bioactive properties and implications for human health across the lorchel family. For example, the characterization of 1-(2-hydroxyacetyl)pyrazol in *G. fastigiata*, a chemical with structural similarity to gyromitrin, might give pause to consumers of *G. fastigiata* as well as its sister species *G. brunnea* (Jurenitsch et al. 1988). In rebuttal, one could point out that this is the case for almost every mushroom that is regarded as edible. For example, choice edible morels (*Morchella* spp.) can occasionally cause illness, including gastrointestinal distress, vomiting, diarrhea, and sometimes even neurological effects, yet the nature of the toxin is a complete mystery (Benjamin 2015). Even when a species contains a potentially toxic substance, such as

agaritine in the common button mushroom (*Agaricus bisporus*), it may still be consumed because the quantities present are generally recognized as safe enough for humans (Khovpachev et al. 2021; Roupas et al. 2010). Regardless, due to the uncharacterized chemodiversity in *Discinaceae*, unchartered species should be treated with extreme caution, including when a species is closely related to others that are consumed without issue.

2.4.2 Conclusions and open questions

Our study is the most extensive sampling to date of lorchels for their capacity to produce gyromitrin. Further sampling will only improve our resolution of the evolution of gyromitrin production in *Discinaceae*. Along with *G. gigas*, it remains an open question whether *G. anthracobia* produces gyromitrin. Described in 2018 from Cyprus, *G. anthracobia* has the closest phylogenetic affinity to the *Gyromitra esculenta* group and thus is considered suspect (Crous et al. 2018; Li et al. 2020). Beyond gyromitrin, explorations of lorchel chemodiversity would not only reveal novel and potentially useful natural products but could also shed light on the etiology of poisonings caused by gyromitrin-free taxa, such as *G. ambigua*. The relationship, if any, between 1-(2-hydroxyacetyl)pyrazol in *G. fastigiata* and gyromitrin is a critical question. The analytical method reported here has the potential to identify previously unreported hydrazine-containing toxins as shown in the strong, uncharacterized peak in the chromatogram of *G. leucoxantha* (Figure 2.4). This finding deserves dedicated follow-up studies to structurally characterize the unknown entity.

The lower levels of gyromitrin produced by western *G. esculenta* group ascocarps are also an avenue for future study. Tylutki (1979) mentions that "*G. esculenta*" has been eaten by many people in the western United States without ill effect, hypothesizing that mushrooms in this region contain less gyromitrin than those in the east (although severe poisonings are not

unheard of [Leathem and Dorran 2007]). *Gyromitra splendida* (MICH 352089) from Washington had low levels of gyromitrin (28× less than *G. venenata* [MICH 352032] from Michigan), providing some support for this hypothesis. Gyromitrin levels seem to decrease with altitude (Andary et al. 1985), which could explain this pattern, but the full extent of the effect of gene-by-environment interactions on the gyromitrin phenotype is an open question. Future studies should systematically explore the differential production of gyromitrin on different media. Our finding that gyromitrin is produced on PDA but not MEA encourages the use of transcriptomics to identify genes whose expression is correlated with mycotoxin production, facilitating the identification of the gyromitrin biosynthesis gene cluster. Identification of the gyromitrin genes will ultimately allow for the elucidation of the evolution of gyromitrin biosynthesis, discovery of homologous gene clusters in *Discinaceae* and other fungal clades, and characterization of novel secondary metabolites with related chemical structures.

Acknowledgements

We extend our gratitude to all the people who collected and donated specimens, which greatly augmented this project: Leah Bendlin, Eric Chandler, Jen Chandler, Jonathan Frank, Igor Khomenko, Kitty Lundeen-Ness, Bee Marcotte, Steve Ness, Dave Ramos, Jon Shaffer, Huafang Su, Garrett Taylor, Jud Vanwyk, and Jeff Volpert. We also thank the fungarium curators, collection managers, and assistants who supported this research by loaning specimen vouchers and cultures: Dr. A. Elizabeth Arnold, Joseph Myers, and Ming-Min Lee (ARIZ); Dr. Matthew Smith, Dr. Rosanne Healy, and Benjamin Lemmond (FLAS); and Patricia Rogers and Dr. Alison Harrington (MICH). Michelle Orozco-Quime's work in trialing gyromitrin spot tests is much appreciated. Vavřinec Klener provided a helpful translation of a section of Kubička (1966). Taylor M. Tai provided thoughtful editorial advice on earlier drafts of the manuscript.

Funding

The strains provided by Dr. A. Elizabeth Arnold were collected with support from the National Science Foundation under grants DEB-1045766 and DEB-1541496, with additional support from the University of Arizona's College of Agriculture and Life Sciences and U.S. Department of Agriculture National Institute of Food and Agriculture (USDA NIFA) awards ARZT-1361340-H25-242 and ARZT-1259370-S25 -200. This research was supported by the Michigan Translational Research and Commercialization (MTRAC) Innovation Hub for AgBio and the Canadian Institute for Advanced Research (CIFAR). A.T. received funding from the University of Michigan Biological Science Initiative. T.Y.J. is a fellow of the CIFAR research program "Fungal Kingdom: Threats & Opportunities." The authors report there are no competing interests to declare.

Figures

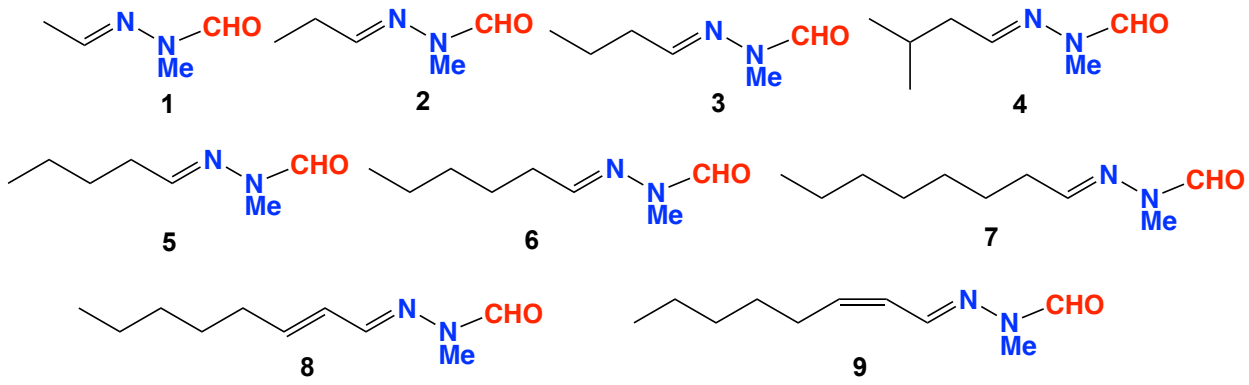


Figure 2.1. Chemical structures of gyromitrin (**1**) and its homologs (**2–9**).

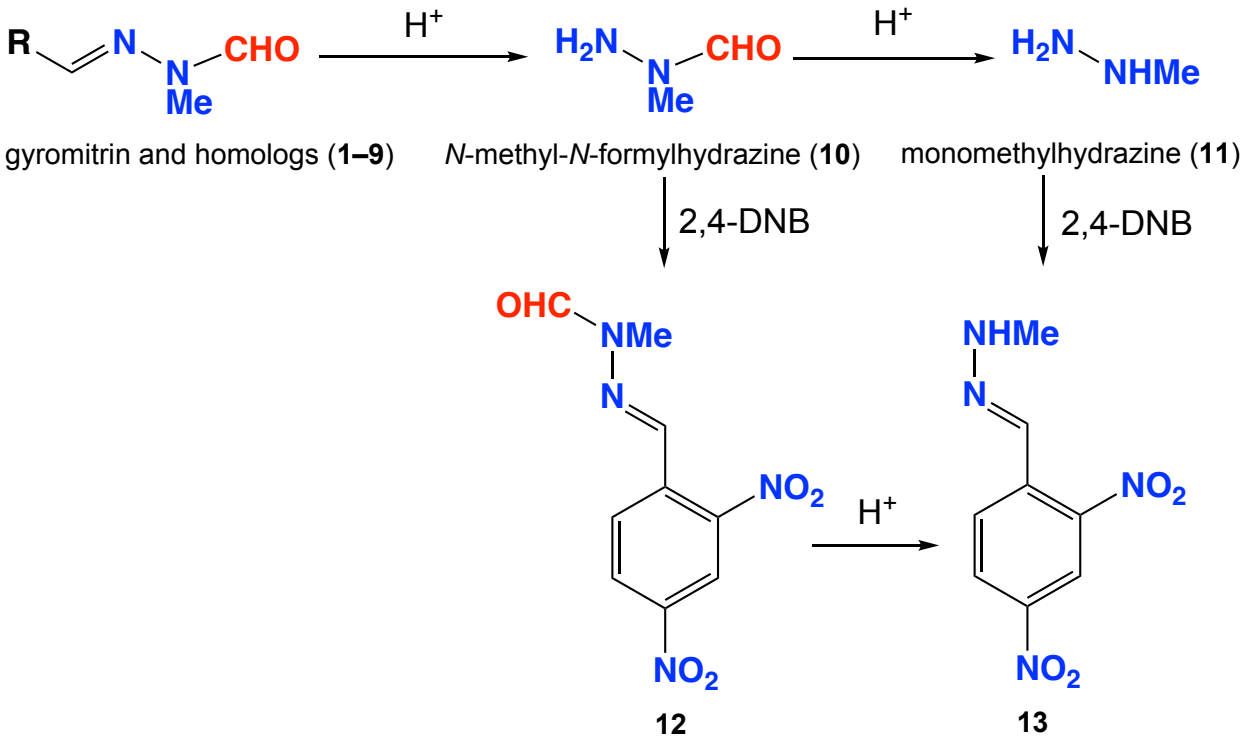


Figure 2.2. Acid hydrolysis of gyromitrin and homologs (**1–9**) yields *N*-methyl-*N*-formylhydrazine (**10**) and monomethylhydrazine (**11**). Derivatization of **10** and **11** with 2,4-dinitrobenzaldehyde (2,4-DNB) yields Schiff bases **12** and **13**.

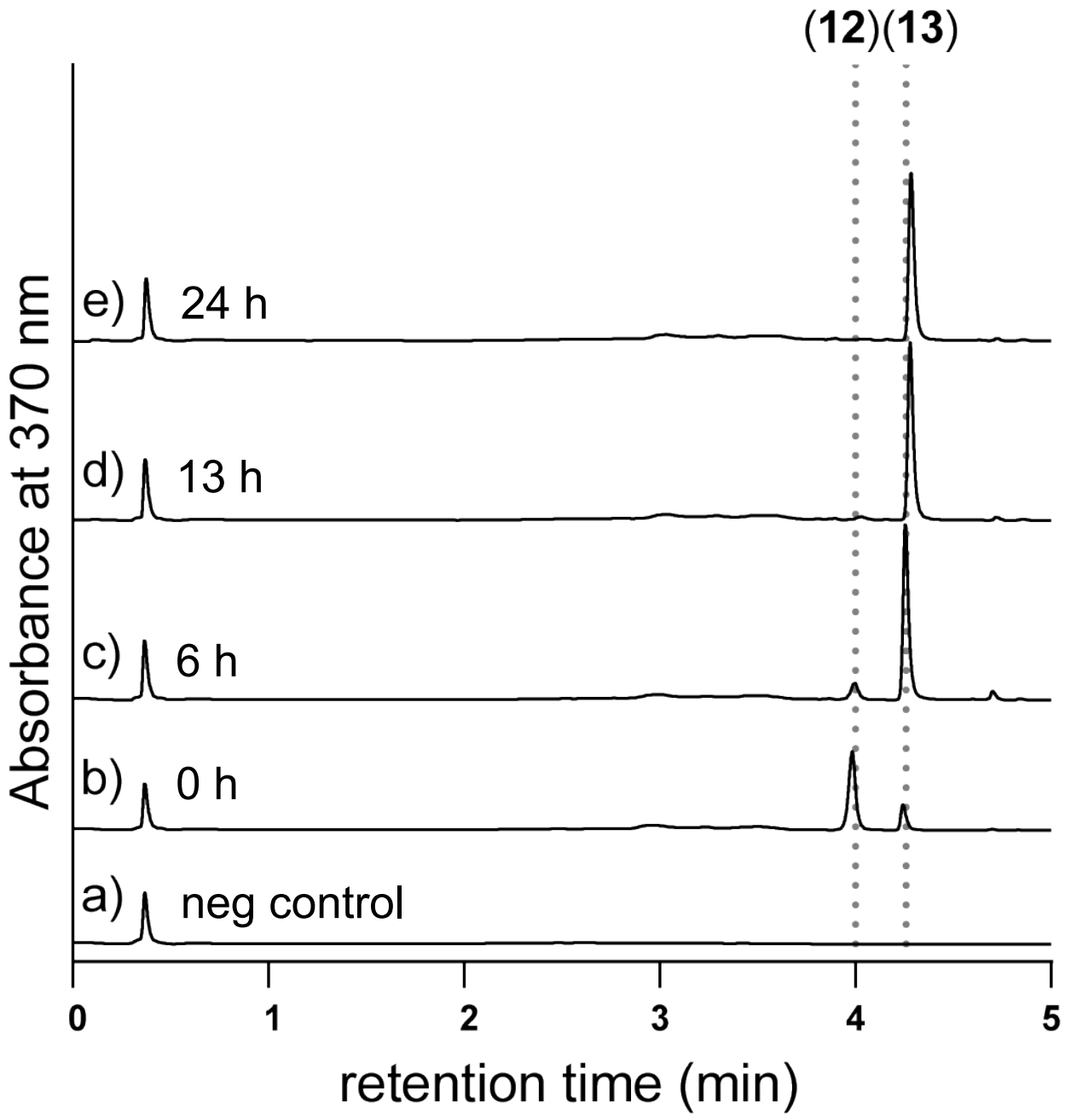


Figure 2.3. UHPLC-DAD (370 nm) chromatograms of derivatization progress of gyromitrin in 10 mg of dried *Gyromitra venenata* ascocarp (MICH 352032). a. Ascocarp extract without 2,4-DNB/TFA treatment (negative control). b–e. Ascocarp extract treated with 2,4-DNB/TFA.

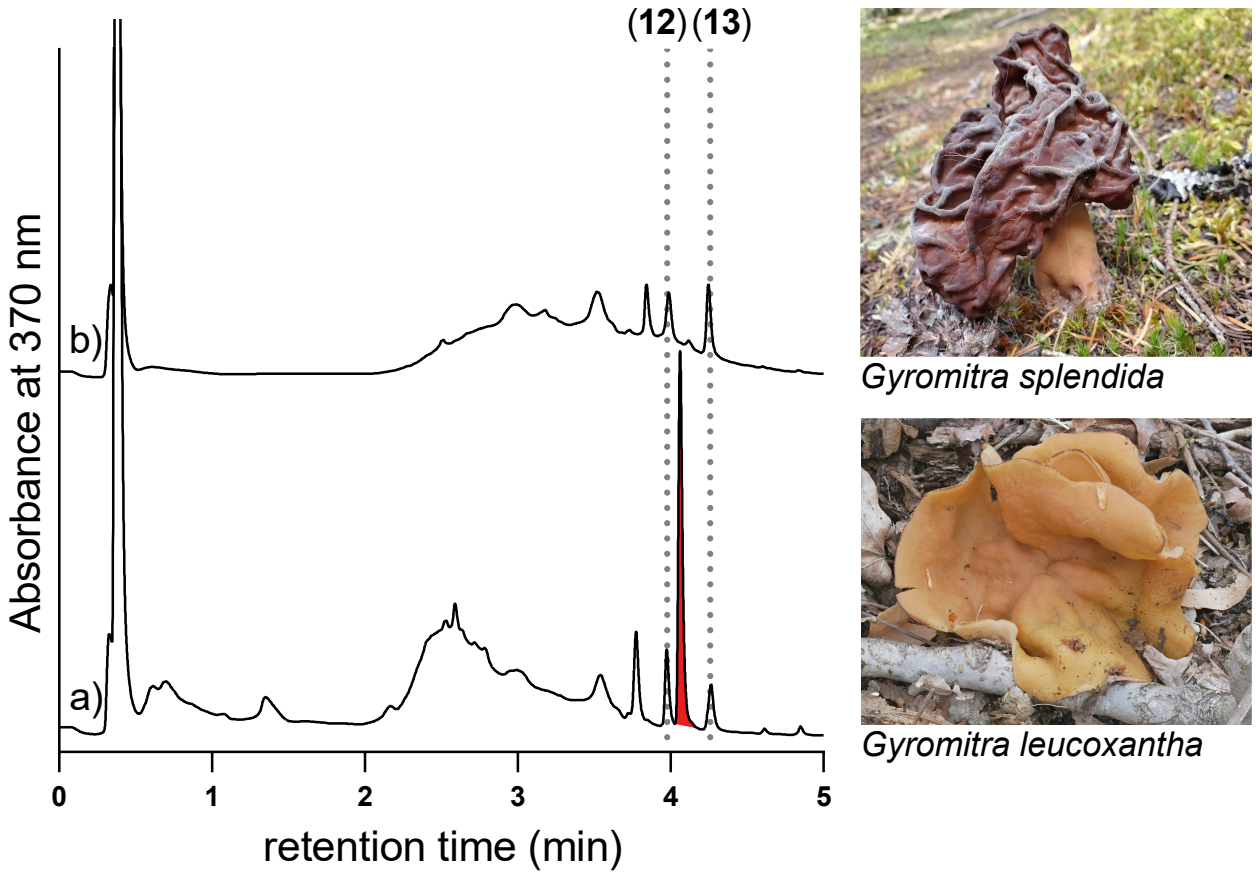


Figure 2.4. UHPLC-DAD (370 nm) chromatograms of ascocarp extracts. a. *Gyromitra leucoxantha* (MICH 352087). The peak corresponding to an unidentified Schiff base derivative is shaded red. b. *Gyromitra splendida* (MICH 352089). Photograph of *G. splendida* by Eric and Jen Chandler. Photograph of *G. leucoxantha* by Garrett Taylor.

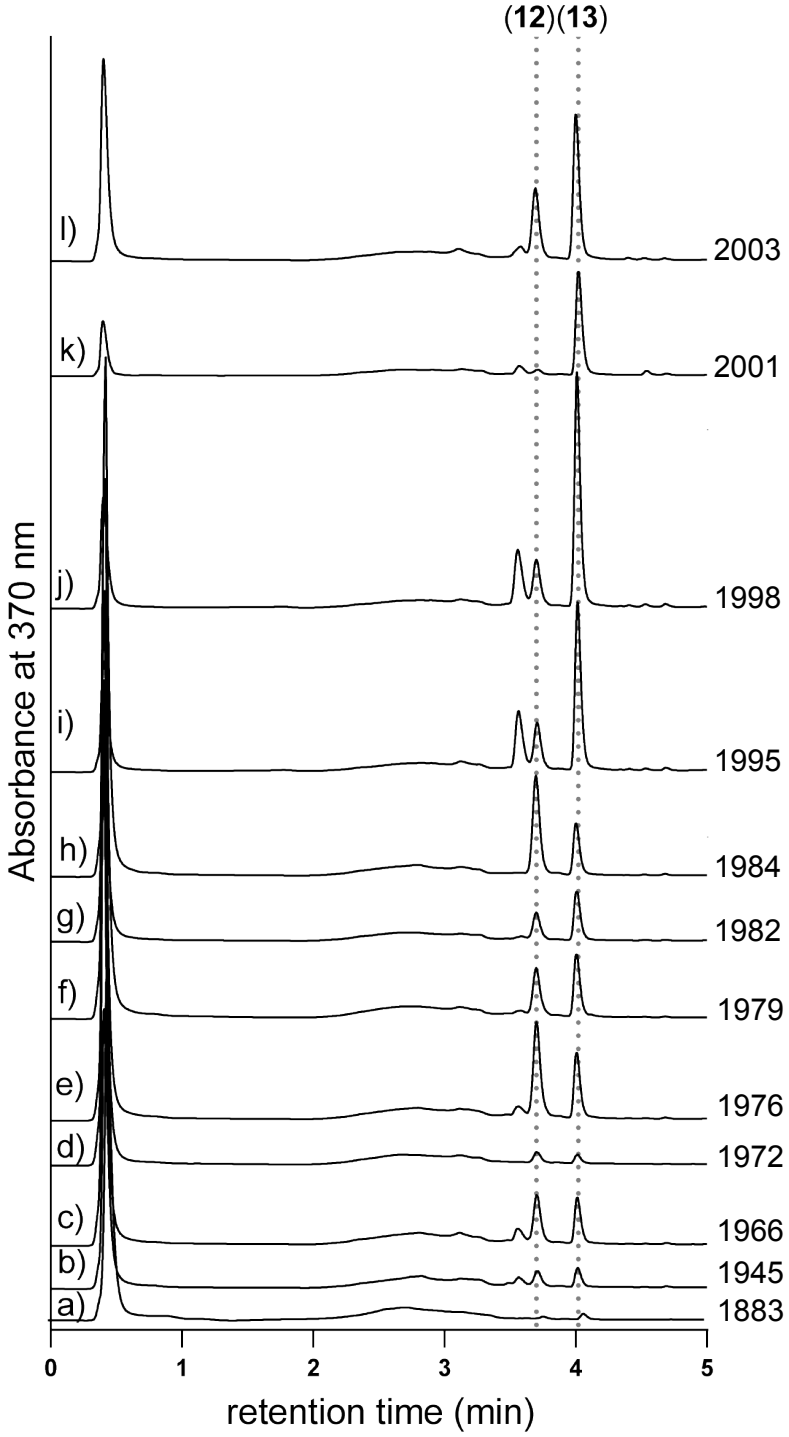


Figure 2.5. UHPLC-DAD (370 nm) chromatograms of 100 mg of *Gyromitra venenata* (a–g, i–l) or *Gyromitra splendida* (h) dried ascocarps collected in different years. a. MICH 25555 (1883). b. MICH 25541 (1945). c. MICH 25528 (1966). d. MICH 1304 (1972). e. MICH 25573 (1976). f. MICH 25543 (1979). g. MICH 25578 (1982). h. MICH 25554 (1984). i. MICH 39037 (1995). j. MICH 39039 (1998). k. MICH 68345 (2001). l. MICH 68493 (2003).

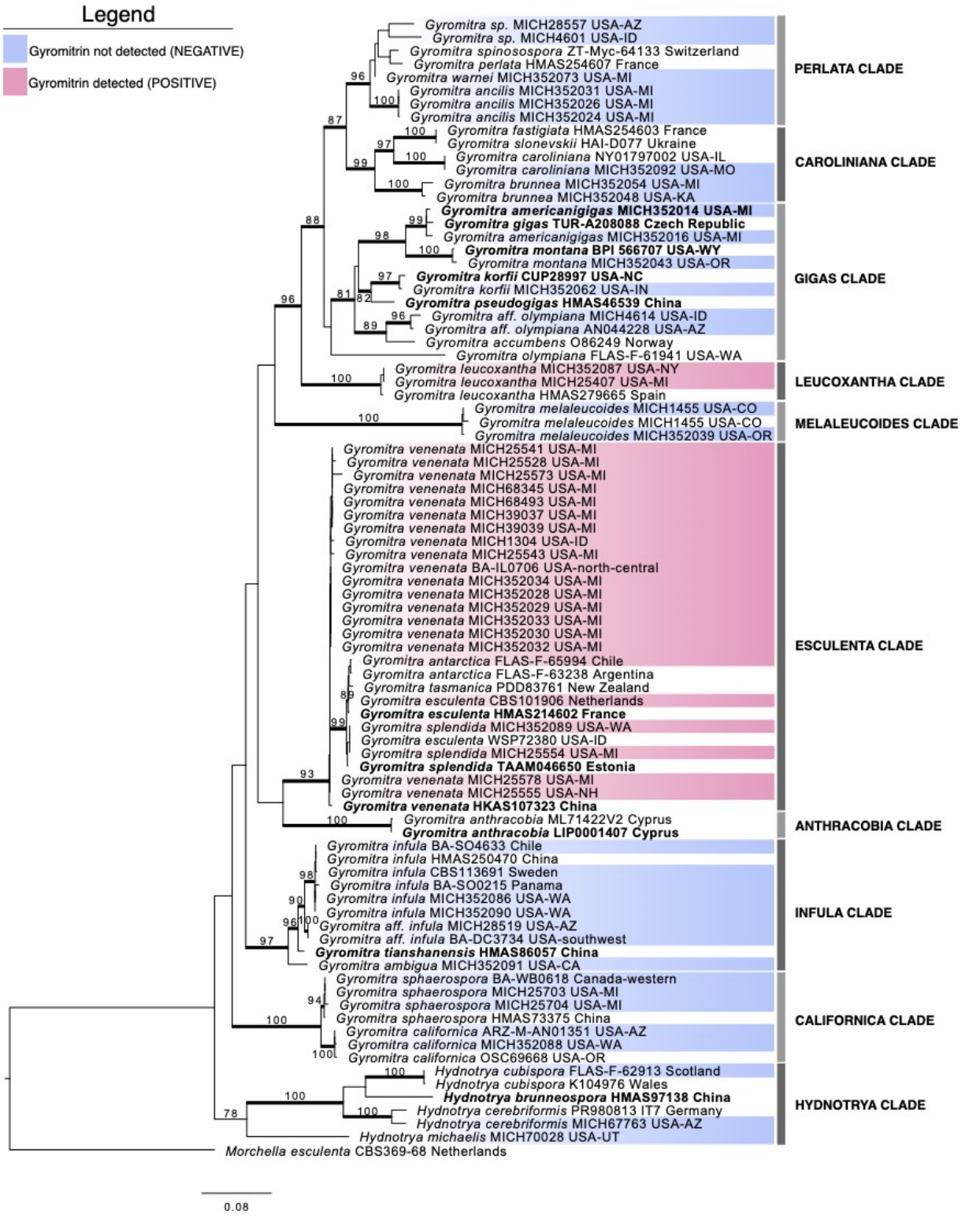


Figure 2.6. *Discinaceae* maximum likelihood phylogenetic tree based on concatenated aligned sequences of the ITS barcode and 28S rDNA showing the presence and absence of gyromitrin. Specimens tested for gyromitrin are highlighted: blue means no gyromitrin was detected (negative); red means gyromitrin was detected (positive). Bootstrap values ≥ 0.70 are shown at the nodes, and Bayesian posterior probability scores above 0.95 are indicated by thickened branches. Taxon names for sequences derived from taxonomic type material are bolded.

Tables

Table 2.1. Specimens tested for gyromitrin and their associated metadata.

Species	Accession Number	Origin	Year	ITS Barcode GenBank No.	28S rDNA GenBank No.
Cudonia grisea	iNat #51411067	USA: Oregon	2020	NA	NA
Disciotis cf. *venosa*	MICH 352035	USA: Michigan	2020	MZ919223	NA
Gyromitra ambigua	MICH 352091	USA: California	2021	ON693648	ON693600
Gyromitra americanigigas	MICH 352014	USA: Michigan	2020	ON527894	ON532830
Gyromitra americanigigas	MICH 352016	USA: Michigan	2020	ON527896	ON532832
Gyromitra ancilis	MICH 352024	USA: Michigan	2020	ON693629	ON693581
Gyromitra ancilis	MICH 352026	USA: Michigan	2020	ON693630	ON693582
Gyromitra ancilis	MICH 352031	USA: Michigan	2020	ON693634	ON693586
Gyromitra antarctica	FLAS-F-65994	Chile	2019	ON693656	ON693611
Gyromitra brunnea	MICH 352048	USA: Kansas	2020	ON693640	ON693592
Gyromitra brunnea	MICH 352054	USA: Michigan	2020	ON693642	ON693594
Gyromitra californica	ARZ-M-AN 01351	USA: Montana	1977	ON693651	ON693604
Gyromitra californica	MICH 352088	USA: Washington	2020	ON693644	ON693596
Gyromitra caroliniana	MICH 352092	USA: Missouri	2021	ON693649	ON693601
Gyromitra esculenta	CBS 101906	Netherlands	culture	ON693653	ON693608
Gyromitra aff. *infula*	DC3764[a]	southwestern USA	culture	MZ091684	MZ018955
Gyromitra aff. *infula*	MICH 28519	USA: Arizona	1990	ON693669	ON693619
Gyromitra infula	SO0215[a]	Panama	culture	MZ091671	ON693606
Gyromitra infula	SO4633[a]	Chile	culture	ON693652	ON693607
Gyromitra infula	CBS 113691	Sweden	culture	ON693654	ON693609
Gyromitra infula	MICH 352086	USA: Washington	2020	ON693643	ON693595
Gyromitra infula	MICH 352090	USA: Washington	2020	ON693646	ON693598
Gyromitra korfii	MICH 352062	USA: Indiana	2021	ON693647	ON693599
Gyromitra leucoxantha	MICH 25407	USA: Michigan	1984	ON693659	ON693613
Gyromitra leucoxantha	MICH 352087	USA: New York	2020	ON693641	ON693593
Gyromitra melaleucoides	MICH 1455	USA: Colorado	1983	ON693658	KC751517
Gyromitra melaleucoides	MICH 352039	USA: Oregon	2020	ON693638	ON693590
Gyromitra montana	MICH 352043	USA: Oregon	2020	ON693639	ON693591
Gyromitra aff. *olympiana*	AN 044228	USA: Arizona	2017	MT483607	ON693603
Gyromitra aff. *olympiana*	MICH 4614	USA: Idaho	1972	ON693674	ON693624
Gyromitra sp.	MICH 28557	USA: Arizona	1983	ON693670	ON693620
Gyromitra sp.	MICH 4601	USA: Idaho	1962	ON693673	ON693623
Gyromitra sphaerospora	WB0618[a]	western Canada	culture	MZ091686	MZ018957
Gyromitra sphaerospora	MICH 25703	USA: Michigan	1970	ON693667	KC751527
Gyromitra sphaerospora	MICH 25704	USA: Michigan	1970	ON693668	KC751526
Gyromitra splendida	MICH 25554	USA: Michigan	1984	ON693663	ON693616
Gyromitra splendida	MICH 352089	USA: Washington	2020	ON693645	ON693597
Gyromitra venenata	IL0706[a]	north-central USA	culture	MZ091661	ON693605
Gyromitra venenata	MICH 1304	USA: Idaho	1972	ON693657	ON693612
Gyromitra venenata	MICH 25528	USA: Michigan	1966	ON693660	ON693614
Gyromitra venenata	MICH 25541	USA: Michigan	1945	ON693661	NA
Gyromitra venenata	MICH 25543	USA: Michigan	1979	ON693662	ON693615
Gyromitra venenata	MICH 25555	USA: New Hamp.	1883	ON693664	ON693617
Gyromitra venenata	MICH 25573	USA: Michigan	1976	ON693665	NA
Gyromitra venenata	MICH 25578	USA: Michigan	1982	ON693666	ON693618
Gyromitra venenata	MICH 352028	USA: Michigan	2020	ON693631	ON693583
Gyromitra venenata	MICH 352029	USA: Michigan	2020	ON693632	ON693584
Gyromitra venenata	MICH 352030	USA: Michigan	2020	ON693633	ON693585
Gyromitra venenata	MICH 352032	USA: Michigan	2020	ON693635	ON693587

Gyromitra venenata	MICH 352033	USA: Michigan	2020	ON693636	ON693588
Gyromitra venenata	MICH 352034	USA: Michigan	2020	ON693637	ON693589
Gyromitra venenata	MICH 39037	USA: Michigan	1995	ON693671	ON693621
Gyromitra venenata	MICH 39039	USA: Michigan	1998	ON693672	ON693622
Gyromitra venenata	MICH 68345	USA: Michigan	2001	ON693676	ON693626
Gyromitra venenata	MICH 68493	USA: Michigan	2003	ON693677	ON693627
Gyromitra warnei	MICH 352073	USA: Michigan	2021	ON693650	ON693602
Hydnotrya cerebriformis	MICH 67763	USA: Arizona	1996	ON693675	ON693625
Hydnotrya cf. *cubispora*	AN 043557	USA: Arizona	2015	NA	NA
Hydnotrya cubispora	FLAS-F-62913	Scotland	2018	ON693655	ON693610
Hydnotrya michaelis	MICH 70028	USA: Utah	1995	NA	ON693628
Leotia lubrica	MICH 352041	USA: Michigan	2020	MZ919236	NA
Morchella punctipes	iNat #46602042	USA: Wisconsin	2020	NA	NA
Pachycudonia monticola	MICH 352040	USA: Oregon	2020	MZ919235	NA
Scutellinia sp.	MICH 352036	USA: Michigan	2020	NA	NA
Sphaerosporella brunnea	MICH 352042	Canada: Ontario	2020	MZ919238	NA
Urnula craterium	iNat #42365327	USA: Wisconsin	2020	NA	NA

Table 2.2. Summary of available evidence for the toxicity and presence of gyromitrin in select *Discinaceae* species based on primary literature and tests conducted for this study. Bolded species are taxa that have been shown to contain gyromitrin. Evidence of toxicity does not necessitate that gyromitrin is the etiological factor, and no evidence of toxicity does not mean that gyromitrin is absent from that species.

Species	**Toxicity**	**Gyromitrin**	**References**
Gyromitra ambigua	Toxic	Initially suspected due to poisoning symptoms; absent	Harmaja 1976; this study
G. americanigigas	No poisoning reports (similar to *G. korfii*)	Absent	This study
G. ancilis	No poisoning reports (similar to *G. perlata*)	Absent	This study
G. antarctica	Toxic	Present	This study
G. brunnea	Nontoxic	Absent	McIlvaine and Macadam 1902; Weber 1995; Beug et al. 2006; this study
G. californica	No poisoning reports	Absent	Tylutki 1979, this study
G. caroliniana	Nontoxic	Previously reported to contain minute quantities; absent	McIlvaine and Macadam 1902; Weber 1995; Liang et al. 1998; this study
G. esculenta	Toxic	Present	List and Luft 1968a, 1968b; this study
G. fastigiata	Nontoxic	Absent	Viernstein et al. 1980
G. gigas	Toxic	Present	Kubička 1966; Viernstein et al. 1980; Lagrange et al. 2021
G. infula	Nontoxic	Absent	Harmaja 1976; this study
G. korfii	Nontoxic	Absent	Weber 1995; this study
G. leucoxantha	Toxic	Present	This study
G. melaleucoides	No poisoning reports (similar to *G. perlata*)	Absent	This study
G. montana	Nontoxic	Absent	Tylutki 1979; Beug et al. 2006; this study
G. perlata	Nontoxic	Absent	Tylutki 1979; Andary et al. 1985; Weber 1995
G. sphaerospora	Unknown	Absent	Weber 1995; this study
G. splendida	Toxic	Present	This study
G. venenata	Toxic	Present	Li et al. 2020; this study
G. warnei	No poisoning reports (similar to *G. perlata*)	Absent	This study
Hydnotrya cerebriformis	Unknown	Absent	This study
H. michaelis	Unknown	Absent	This study

References

Abadi S, Azouri D, Pupko T, Mayrose I. 2019. Model selection may not be a mandatory step for phylogeny reconstruction. Nature Communications 10:1–11.

Alfaro ME, Zoller S, Lutzoni F. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Molecular Biology and Evolution 20:255–266.

Andary C, Privat G, Bourrier M-J. 1984. Microdosage spectrofluorimétrique sur couches minces de la monométhylhydrazine chez Gyromitra esculenta. Journal of Chromatography 287:419–424.

Andary C, Privat G, Bourrier M-J. 1985. Variations of monomethylhydrazine content in Gyromitra esculenta. Mycologia 77:259–264.

Arłukowicz-Grabowska M, Wójcicki M, Raszeja-Wyszomirska J, Szydłowska-Jakimiuk M, Piotuch B, Milkiewicz P. 2019. Acute liver injury, acute liver failure and acute on chronic liver failure: A clinical spectrum of poisoning due to Gyromitra esculenta. Annals of Hepatology 18:514–516.

Arshadi M, Nilsson C, Magnusson B. 2006. Gas chromatography-mass spectrometry determination of the pentafluorobenzoyl derivative of methylhydrazine in false morel (Gyromitra esculenta) as a monitor for the content of the toxin gyromitrin. Journal of Chromatography A 1125:229–233.

Benjamin DR. 2015. Neurological effects of Morchella sp. FUNGI 8:24–25.

Benjamin DR. 2020. Gyromitrin poisoning: More questions than answers. FUNGI 13:36–39.

Beug MW. 2014. False morels: Age-old questions of edibility. FUNGI 7:28–31.

Beug MW, Shaw M, Cochran KW. 2006. Thirty-plus years of mushroom poisoning: Summary of the approximately 2000 reports in the NAMA case registry. McIlvainea: Journal of American Amateur Mycology 16:47–68.

Blair LM, Sperry J. 2013. Natural products containing a nitrogen-nitrogen bond. Journal of Natural Products 76:794–812.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17:540–552.

Court RC Van, Wiseman MS, Meyer KW, Ballhorn DJ, Amses KR, Slot JC, Dentinger BTM, Garibay-Orijel R, Uehling JK. 2022. Diversity, biology, and history of psilocybin-containing fungi: Suggestions for research and technological development. Fungal Biology 126:308–319.

Crous PW, Wingfield MM, Burgess TI, Hardy GESJ, Gené J, Guarro J, Baseia IG, García D, Gusmão LFP, Thangavel R, Adamčík S, Barili A, Barnes CW, Bezerra JDP, Bordallo JJ, Cano-Lira JF, Oliveira RJ V. de, Ercole E, Hubka V, Iturrieta-González I, Kubátová A, Martín MP, Moreau P-A, Morte A, Ordoñez ME, Rodríguez A, Stchigel AM, Vizzini A, Abdollahzadeh J, et al. 2018. Fungal Planet description sheets: 716–784. Persoonia 40:240–393.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792–1797.

Gardes M, Bruns TD. 1993. ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. Molecular Ecology 2:113–118.

Goff G Le, Ouazzani J. 2014. Natural hydrazine-containing compounds: Biosynthesis, isolation, biological activities and synthesis. Bioorganic and Medicinal Chemistry 22:6529–6544.

Gouy M, Guindon S, Gascuel O. 2010. Sea view version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution 27:221–224.

Grigoriev I V., Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. 2014. MycoCosm portal: Gearing up for 1000 fungal genomes. Nucleic Acids Research 42:D699–D704.

Härkönen M. 1998. Uses of mushrooms by Finns and Karelians. International Journal of Circumpolar Health 57:40–55.

Harmaja H. 1976. Another poisonous species discovered in the genus Gyromitra: G. ambigua. Karstenia 15:36–37.

Healy RA, Arnold AE, Bonito G, Huang YL, Lemmond B, Pfister DH, Smith ME. 2022. Endophytism and endolichenism in Pezizomycetes: The exception or the rule? New Phytologist 233:1974–1983.

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Systematic Biology 42:182–192.

Jurenitsch J, Aurada E, Kubelka W. 1988. 1-(2-Hydroxyacetyl)-pyrazol, eine neue, mit gyromitrin verwandte verbindung aus Gyromitra fastigiata. Zeitschrift Für Mykologie 54:155–157.

Kans J. 2020. Entrez Direct: E-utilities on the UNIX Command Line. In: Entrez Programming Utilities Help [Internet]. Bethesda, MD: National Center for Biotechnology Information. p. 1–65.

Karlson-Stiber C, Persson H. 2003. Cytotoxic fungi: An overview. Toxicon 42:339–349.

Khovpachev AA, Basharin VA, Chepur S V., Volobuev S V., Yudin MA, Gogolevsky AS, Nikiforov AS, Kalinina LB, Tyunin MA. 2021. Actual concepts of higher fungi's toxins: Simple nitrogen-containing compounds. Biology Bulletin Reviews 11:198–212.

Kubička J. 1966. Čtyři případy otravy ucháčem (Gyromitra). Česká Mykologie 20:178–181.

Lagrange E, Vernoux JPP, Reis J, Palmer V, Camu W, Spencer PSS. 2021. An amyotrophic lateral sclerosis hot spot in the French Alps associated with genotoxic fungi. Journal of the Neurological Sciences 427:1–8.

Larget B, Simon DL. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Molecular Biology and Evolution 16:750–759.

Leathem AM, Dorran TJ. 2007. Poisoning due to raw Gyromitra esculenta (false morels) west of the Rockies. Canadian Journal of Emergency Medicine 9:127–130.

Li H-J, Zuo-Hong C, Cai Q, Zhou M-H, Chen G-J, Cheng-Ye S, Zhang H-S, Zhu-Liang Y. 2020. Gyromitra venenata, a new poisonous species discovered from China. Mycosystema 39:1706–1718.

Liang Y-H, Eisenga BH, Trestrail JTI, Kuslikis B. 1998. Gyromitra mushroom species and their monomethylhydrazine content. In: Platform Session 1: Part 2, Journal of Toxicology: Clinical Toxicology. p. 527.

List PH, Luft P. 1968a. Gyromitrin, das gift der frühjahrslorchel. 16. Mitt. über pilzinhaltsstoffe [Gyromitrin, the poison of Gyromitra esculenta. 16. On the fungi contents]. Arch Pharm Ber Dtsch Pharm Ges 301:294–305.

List PH, Luft P. 1968b. Nachweis und gehaltsbestimmung von gyromitrin in frischen lorcheln. 19. Mitt. über pilzinhaltsstoffe [Detection and content determination of gyromitrin in fresh Gyromitra esculenta. 19. Mushroom contents]. Arch Pharm Ber Dtsch Pharm Ges 302:143–146.

Luo H, Hallen-Adams HE, Lüli Y, Sgambelluri RM, Li X, Smith M, Yang ZL, Martin FM. 2022. Genes and evolutionary fates of the amanitin biosynthesis pathway in poisonous mushrooms. PNAS 119:1–11.

Mäkinen SM, Kreula M, Kauppi M. 1977. Acute oral toxicity of ethylidene gyromitrin in rabbits, rats and chickens. Food and Cosmetics Toxicology 15:575–578.

Marjatta R, Pyysalo H. 1978. Occurrence of N-methyl-N-formylhydrazones in mycelia of Gyromitra esculenta. Zeitschrift Fur Naturforschung 33:472–474.

McIlvaine C, Macadam RK. 1902. One Thousand American Fungi. Indianapolis: Bowen-Merrill Company. 729 p.

Michelot D, Toth B. 1991. Poisoning by Gyromitra esculenta—a review. Journal of Applied Toxicology 11:235–243.

Miller AN, Dirks AC, Filippova N, Popov E, Methven AS. 2022. Studies in Gyromitra II: Cryptic speciation in the Gyromitra gigas species complex; rediscovery of G. ussuriensis and G. americanigigas sp. nov. Mycological Progress 21:1–13.

Miller AN, Yoon A, Gulden G, Stensholt Ø, Vooren N Van, Ohenoja E, Methven AS. 2020. Studies in Gyromitra I: The Gyromitra gigas species complex. Mycological Progress 19:1459–1473.

Miller MA, Pfeiffer W, Schwartz T. 2011. The CIPRES science gateway: A community resource for phylogenetic analyses. In: Proceedings of the TeraGrid 2011 Conference: Extreme Digital Discovery, TG'11. p. 1–8.

Mohamed OG, Khalil ZG, Capon RJ. 2018. Prolinimines: N-amino-L-pro-methyl ester (hydrazine) Schiff bases from a fish gastrointestinal tract-derived fungus, Trichoderma sp. CMB-F563. Organic Letters 20:377–380.

Mohamed OG, Khalil ZG, Capon RJ. 2021. N-amino-L-proline methyl ester from an Australian fish gut-derived fungus: Challenging the distinction between natural product and artifact. Marine Drugs 19:1–14.

Moncalvo JM, Lutzoni FM, Rehner SA, Johnson J, Vilgalys R. 2000. Phylogenetic relationships of agaric fungi based on nuclear large subunit ribosomal DNA sequences. Systematic Biology 49:278–305.

Monteith WR. 2020. Draft Environmental Assessment for SpaceX Falcon Launches at Kennedy Space Center and Cape Canaveral Air Force Station. 110 p.

Morris JJ, Lenski RE, Zinser ER. 2012. The black queen hypothesis: Evolution of dependencies through adaptive gene loss. MBio 3:1–7.

Patron NJ, Waller RF, Cozijnsen AJ, Straney DC, Gardiner DM, Nierman WC, Howlett BJ. 2007. Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evolutionary Biology 7:174–188.

Porter TM, Martin W, James TY, Longcore JE, Gleason FH, Adler PH, Letcher PM, Vilgalys R. 2011. Molecular phylogeny of the Blastocladiomycota (Fungi) based on nuclear ribosomal DNA. Fungal Biology 115:381–392.

Pyysalo H. 1975. Some new toxic compounds in false morels, Gyromitra esculenta. Naturwissenschaften 62:395.

Pyysalo H. 1976. Tests for gyromitrin, a poisonous compound in false morel Gyromitra esculenta. Zeitschrift Für Lebensmittel-Untersuchung Und -Forschung 160:325–330.

Pyysalo H, Niskanen A, Wright A von. 1978. Formation of toxic methylhydrazine during cooking of false morels, Gyromitra esculenta. Journal of Food Safety 1:295–299.

Pyysalo H, Niskanen AA. 1977. On the occurrence of N-methyl-N-formylhydrazones in fresh and processed false morel, Gyromitra esculenta. Journal of Agricultural and Food Chemistry 25:644–647.

Rokas A, Mead ME, Steenwyk JL, Raja HA, Oberlies NH. 2020. Biosynthetic gene clusters and the evolution of fungal chemodiversity. Natural Product Reports 37:868–878.

Ronquist F, Teslenko M, Mark P Van Der, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61:539–542.

Roupas P, Keogh J, Noakes M, Margetts C, Taylor P. 2010. Mushrooms and agaritine: A mini-review. Journal of Functional Foods 2:91–98.

Sitta N, Davoli P, Floriani M, Suriano E. 2021. Guida ragionata alla commestibilità dei funghi. 180 p.

Spencer PS. 2020. Consumption of unregulated food items (false morels) and risk for neurodegenerative disease (amyotrophic lateral sclerosis). Health Risk Analysis 2020:93–98.

Spencer PS, Kisby GE. 2021. Role of hydrazine-related chemicals in cancer and neurodegenerative disease. Chemical Research in Toxicology 34:1953–1969.

Spencer PS, Palmer VS. 2021. Direct and indirect neurotoxic potential of metal/metalloids in plants and fungi used for food, dietary supplements, and herbal medicine. Toxics 9:57–69.

Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Svanberg I, Lindh H. 2019. Mushroom hunting and consumption in twenty-first century post-industrial Sweden. Journal of Ethnobiology and Ethnomedicine 15:1–23.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology 56:564–577.

Toth B, Nagel D. 1978. Tumors induced in mice by N-methyl-N-formylhydrazine of the false morel Gyromitra esculenta. Journal of the National Cancer Institute 60:201–204.

Toth B, Patil K. 1980. Carcinogenesis by a single dose of N-methyl-N-formylhydrazine. Journal of Toxicology and Environmental Health 6:577–584.

Tylutki EE. 1979. Mushrooms of Idaho and the Pacific Northwest: Vol. 1 Discomycetes. Moscow, Idaho: University Press of Idaho. 133 p.

Viernstein H, Jurenitsch J, Kubelka W. 1980. Vergleich des giftgehaltes der lorchelarten Gyromitra gigas, Gyromitra fastigiata und Gyromitra esculenta. Ernahrung/Nutrition 4:392–395.

Wang X-CC, Zhuang W-YY. 2019. A three-locus phylogeny of Gyromitra (Discinaceae, Pezizales) and discovery of two cryptic species. Mycologia 111:69–77.

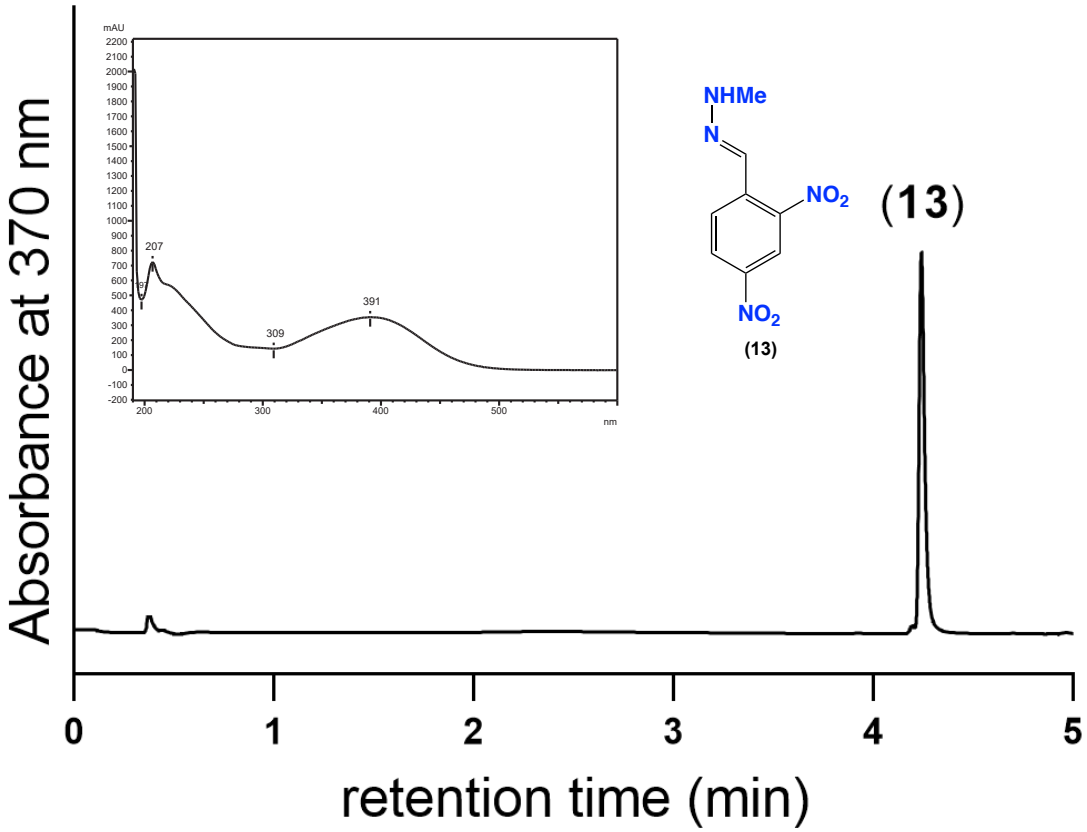
Weber NS. 1995. A Morel Hunter's Companion: A Guide to True and False Morels. Holt, Michigan: Thunder Bay Press. 209 p.

White TJ, Bruns TD, Lee SB, Taylor JW, White TJ, Bruns TD, Lee SB, Taylor JW. 1990. Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. In: PCR Protocols: A Guide to Methods and Applications. Academic Press, Inc. p. 315–322.

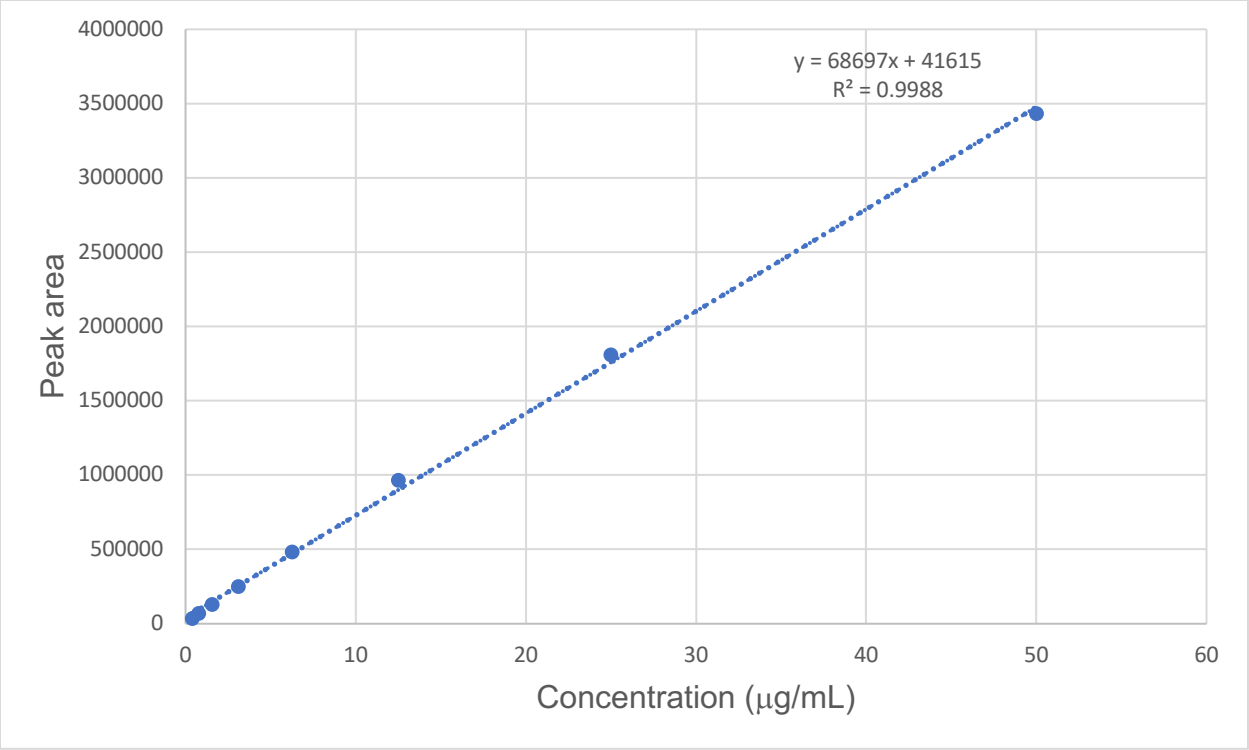
Wright A V., Niskanen A, Pyysalo H, Korpela H. 1978. The toxicity of some N-methyl-N-formylhydrazones from Gyromitra esculenta and related compounds in mouse and microbial tests. Toxicology and Applied Pharmacology 45:429–434.

Yancey CE, Smith DJ, Uyl PA Den, Mohamed OG, Yu F, Ruberg SA, Chaffin JD, Goodwin KD, Tripathi A, Sherman DH, Dick GJ. 2022. Metagenomic and metatranscriptomic insights into population diversity of Microcystis blooms: Spatial and temporal dynamics of mcy genotypes, including a partial operon that can be abundant and expressed. Applied and Environmental Microbiology 88:1–17.

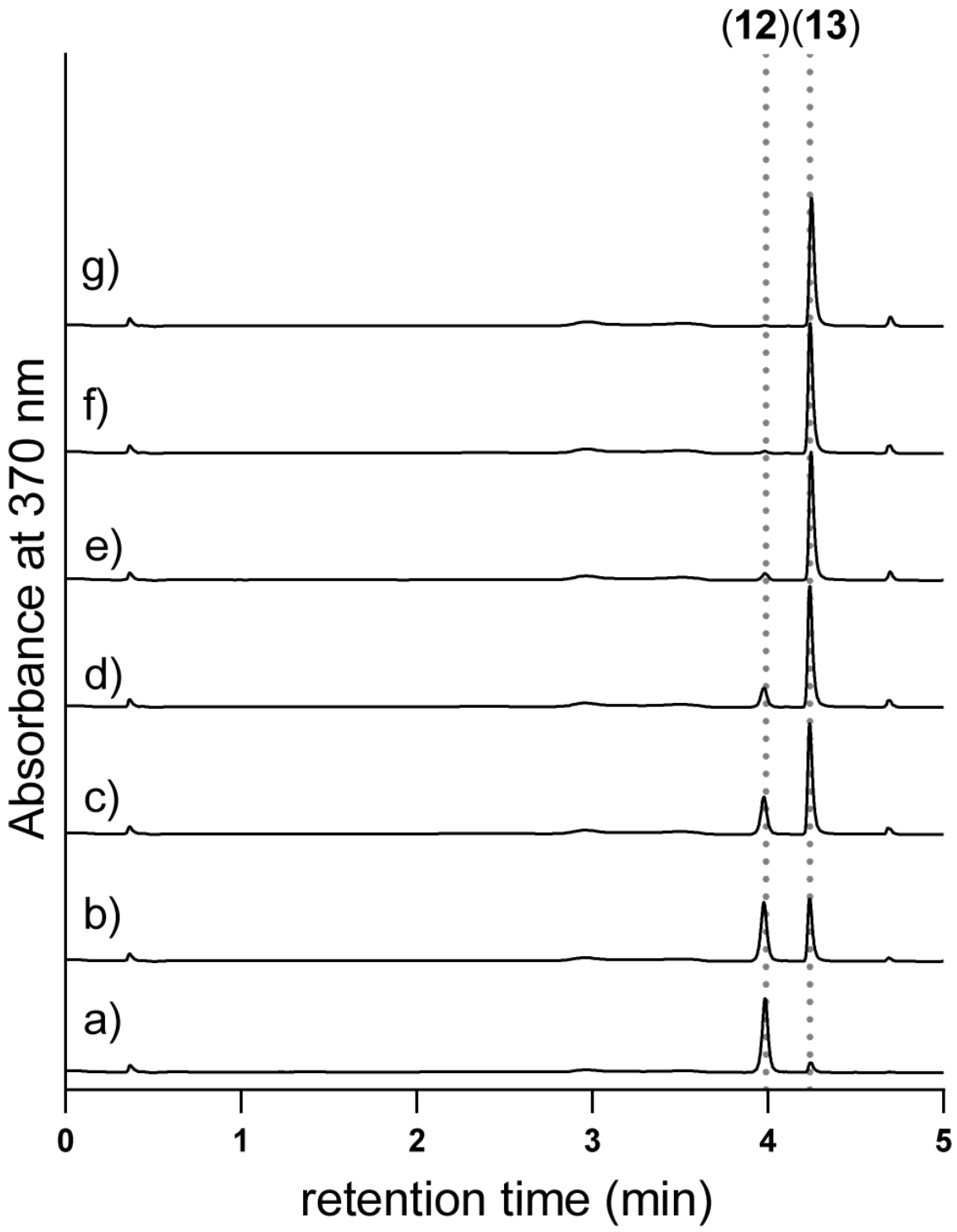
Supplement



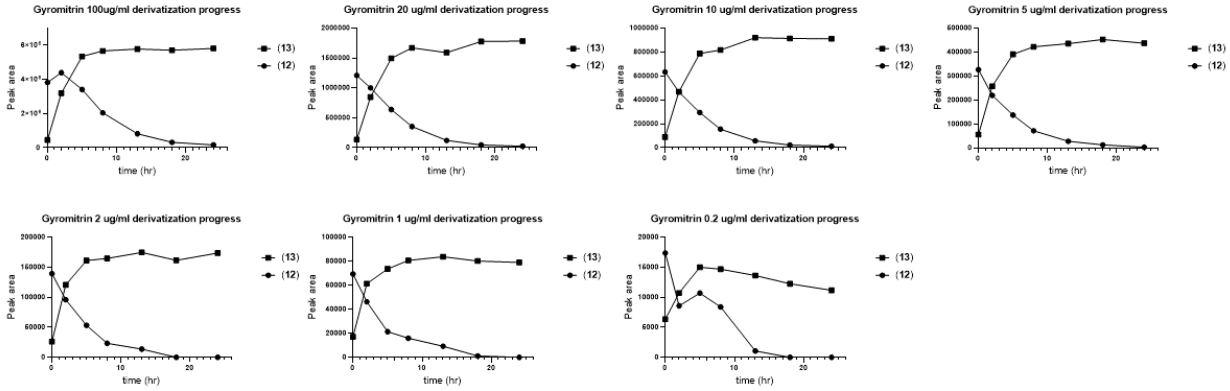
Supplemental Figure 2.1. UHPLC-DAD (370 nm) trace and UV-vis spectrum of synthetic monomethylhydrazine Schiff base (**13**).



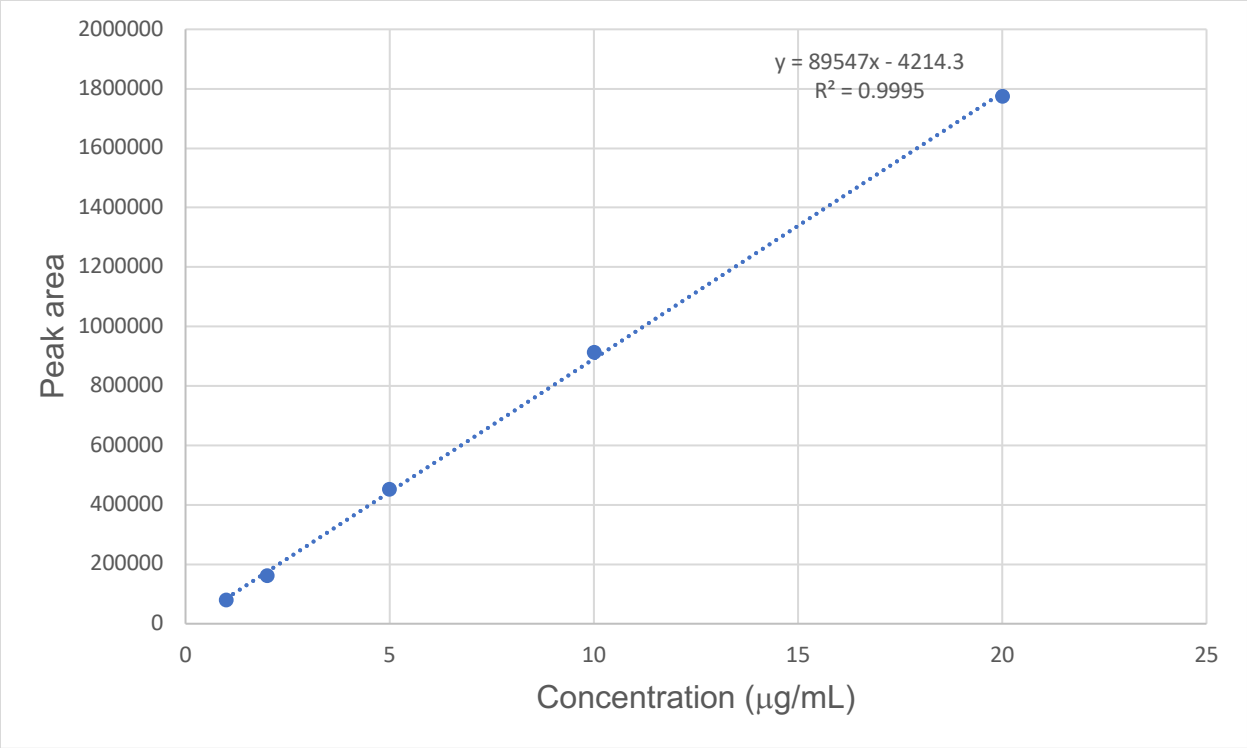
Supplemental Figure 2.2. Standard calibration curve of synthetic monomethylhydrazine Schiff base (**13**) at different concentrations (μg/mL) against UV peak area (370 nm).



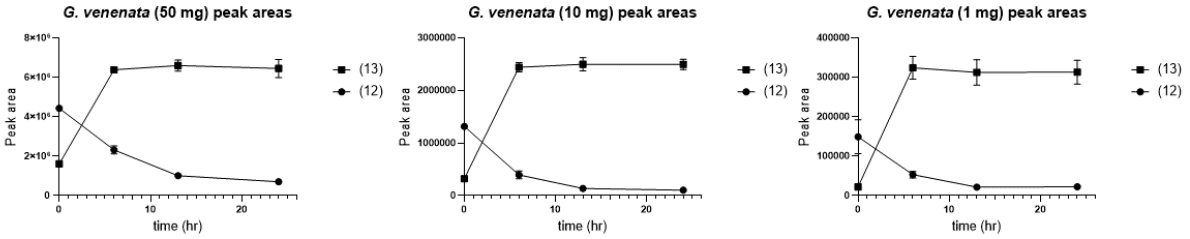
Supplemental Figure 2.3. UHPLC-DAD (370 nm) chromatograms of derivatization progress of gyromitrin standard (20 µg/mL) with 2,4-DNB at different time intervals. a) 0 h, b) 2 h, c) 5 h, d) 8 h, e) 13 h, f) 18 h, and g) 24 h.



Supplemental Figure 2.4. Hydrolysis and derivatization progress of different concentrations of gyromitrin standard using 2,4-DNB over time.



Supplemental Figure 2.5. Correlation between different concentrations of gyromitrin standard solution and UV peak area (370 nm) of derivatized product (**13**).



Supplemental Figure 2.6. Hydrolysis and derivatization progress of gyromitrin in different amounts of *Gyromitra venenata* (MICH 352032) over time using 2,4-DNB.

Supplemental Table 2.1. Accession numbers of GenBank sequences used in the phylogenetic analyses.

Species	Fungarium	ID	Location	ITS	LSU	Type	Reference(s)
Gyromitra accumbens	O86249	NA	Norway		KX008332		Miller unpublished
G. antarctica	FLAS-F-63238	MES-1298	Argentina	MH930384	MZ018861		Mujic and Smith unpublished; Healy et al. 2022
G. anthracobia	LIP0001407	NA	Cyprus	MH014751	MH014750	HOLOTYPE	Crous et al. 2018
G. anthracobia	NA	ML71422V2	Cyprus	MH014752	MH014749		Crous et al. 2018
G. californica	OSC69668	Mike Roantree s.n.	USA-OR	EU837203			Gordon unpublished
G. caroliniana	NY01797002	KUO-04100601	USA-IL		KC751528		Methven et al. 2013
G. esculenta	WSP72380	Ge 8-1	USA-ID	KM204673	KM204699		Carris et al. 2015
G. esculenta	HMAS214602	NV2015.04.01	France	MG846990	MG847002	EPITYPE	Wang and Zhuang 2019
G. fastigiata	HMAS254603	NV2004.05.01	France	MG846992	MG847003		Wang and Zhuang 2019
G. gigas	TUR-A208088	NA	Czech Republic	MH938663	MH938309	EPITYPE	Carbone et al. 2018
G. infula	HMAS250470	NA	China	MG846956	MG847006		Wang and Zhuang 2019
G. korfii	CUP28997	NA	USA-NC	MW075385	MW078425	PARATYPE	Miller et al. 2020
G .leucoxantha	HMAS279665	NV2017.06.10	Spain	MG846991	MG847020		Wang and Zhuang 2019
G. melaleucoides	MICH1455	NSW 4520	USA-CO		KC751517		Methven et al. 2013
G. montana	BPI 566707	NA	USA-WY	MW077452	MW077442	ISOTYPE	Miller et al. 2020
G. olympiana	FLAS-F-61941	MES-3061	USA-WA	MT373908	MT350422		Healy unpublished
G. perlata	HMAS254607	NV2014.05.14	France	MG846993	MG847022		Wang and Zhuang 2019
G. pseudogigas	HMAS46539	NA	China	MG846994	MG847023	HOLOTYPE	Wang and Zhuang 2019
G. slonevskii	HAI-D077	D-077	Ukraine	JQ691488			Barseghyan et al. 2012
G. sphaerospora	HMAS73375	NA	China	MG846997	MG847027		Wang and Zhuang 2019
G. spinosospora	ZT-Myc-64133	NA	Switzerland	MW489531			Senn-Irlet et al. 2021
G. splendida	TAAM046650	KL398	Estonia	KX185090	KX185094	HOLOTYPE	Miller et al. 2015
G. tasmanica	PDD83761	JAC9664	New Zealand	MK432693			Cooper et al. unpublished
G. tianshanensis	HMAS86057	NA	China	MG846963	MG847024	HOLOTYPE	Wang and Zhuang 2019
G. venenata	HKAS107323	NA	China	MT424856	MT421932	PARATYPE	Li et al. 2020
Hydnotrya brunneospora	HMAS97138	NA	China	NR_161073		HOLOTYPE	Xu et al. 2018
H. cerebriformis	NA	PR980813 IT7	Germany	GQ140236			Stielow et al. 2010
H. cubispora	K104976	B.J. Lack s.n.	Wales	EU784273			Brock et al. 2009

Supplemental Table 2.2. Peak area and recovery percentage of **13** in the pre-column derivatization (18 hours post incubation at 40 C) and UHPLC-DAD analysis (370 nm) of gyromitrin standard at various concentrations.

Gyromitrin standard concentration (µg/mL)	Peak area	Recovery percentage
100	5704347	36.80
20	1775369	56.33
10	913033	56.63
5	451864	53.32
2	161334	38.90
1	80118	25.02

Supplemental Table 2.3. Metadata and results of gyromitrin tests conducted on samples for this study.

Species	Accession #	Collection #	Origin	Year	Analyzed	Gyromitrin
Cudonia grisea	NA	iNat #51411067	OR	2020	dried, frozen	0
Disciotis cf. *venosa*	MICH 352035	ACD0290	MI	2020	dried, frozen, culture (PDA)	0
Gyromitra ambigua	MICH 352091	ACD0408	CA	2021	dried ×2, culture (PDA)	0
G. americanigigas	MICH 352014	ACD0256	MI	2020	dried, frozen	0
G. americanigigas	MICH 352016	ACD0258	MI	2020	dried, frozen	0
G. ancilis	MICH 352024	ACD0260	MI	2020	dried, frozen	0
G. ancilis	MICH 352026	ACD0263	MI	2020	dried, frozen	0
G. ancilis	MICH 352031	ACD0285	MI	2020	frozen	0
G. antarctica	FLAS-F-65994	NA	Chile	2019	dried	1
G. brunnea	MICH 352048	ACD0324	KA	2020	dried, frozen	0
G. brunnea	MICH 352054	ACD0332	MI	2020	dried	0
G. californica	ARZ-M-AN 01351	NA	MO	1977	dried	0
G. californica	MICH 352088	ACD0394	WA	2020	dried ×2, culture (PDA)	0
G. caroliniana	MICH 352092	ACD0409	MO	2021	dried ×2	0
G. esculenta	CBS 101906	NA	Netherlands	culture	culture (PDA, MEA ×2)	1 (PDA), 0 (MEA)
G. aff. *infula*	Betsy Arnold	DC3764	USA	culture	culture (PDA)	0
G. aff. *infula*	MICH 28519	NA	AZ	1990	dried ×2	0
G. infula	Betsy Arnold	SO0215	Panama	culture	culture (PDA)	0
G. infula	Betsy Arnold	SO4633	Chile	culture	culture (PDA)	0
G. infula	CBS 113691	NA	Sweden	culture	culture (PDA, MEA)	0
G. infula	MICH 352086	ACD0393	WA	2020	dried	0
G. infula	MICH 352090	ACD0396	WA	2020	dried ×2, culture (PDA)	0
G. korfii	MICH 352062	ACD0399	IN	2021	dried	0
G. leucoxantha	MICH 25407	NA	MI	1984	dried	1
G. leucoxantha	MICH 352087	ACD0326	NY	2020	dried ×3	1
G. melaleucoides	MICH 1455	NA	CO	1983	dried	0
G. melaleucoides	MICH 352039	ACD0308	OR	2020	dried	0
G. montana	MICH 352043	ACD0318	OR	2020	dried, frozen	0
G. aff. *olympiana*	AN 044228	NA	AZ	2017	dried	0
G. aff. *olympiana*	MICH 4614	NA	ID	1972	dried ×2	0
G. sp.	MICH 28557	NA	AZ	1983	dried ×2	0
G. sp.	MICH 4601	NA	ID	1962	dried	0
G. sphaerospora	Betsy Arnold	WB0618	Canada	culture	culture (PDA)	0
G. sphaerospora	MICH 25703	NA	MI	1970	dried ×2	0
G. sphaerospora	MICH 25704	NA	MI	1970	dried	0
G. splendida	MICH 25554	NA	MI	1984	dried	1
G. splendida	MICH 352089	ACD0395	WA	2020	dried ×2, culture (PDA)	1
G. venenata	Betsy Arnold	IL0706	USA	culture	culture (PDA)	1
G. venenata	MICH 1304	NA	ID	1972	dried	1
G. venenata	MICH 25528	NA	MI	1966	dried	1
G. venenata	MICH 25541	NA	MI	1945	dried	1
G. venenata	MICH 25543	NA	MI	1979	dried	1
G. venenata	MICH 25555	NA	NH	1883	dried	1
G. venenata	MICH 25573	NA	MI	1976	dried	1
G. venenata	MICH 25578	NA	MI	1982	dried	1
G. venenata	MICH 352028	ACD0282	MI	2020	dried, frozen	1
G. venenata	MICH 352029	ACD0283	MI	2020	dried, frozen	1
G. venenata	MICH 352030	ACD0284	MI	2020	dried, frozen, culture (PDA)	1
G. venenata	MICH 352032	ACD0286	MI	2020	dried, frozen, culture (PDA)	1
G. venenata	MICH 352033	ACD0287	MI	2020	dried, culture (PDA)	1
G. venenata	MICH 352034	ACD0288	MI	2020	dried	1
G. venenata	MICH 39037	NA	MI	1995	dried	1
G. venenata	MICH 39039	NA	MI	1998	dried	1
G. venenata	MICH 68345	NA	MI	2001	dried	1
G. venenata	MICH 68493	NA	MI	2003	dried	1
G. warnei	MICH 352073	ACD0420	MI	2021	dried	0
Hydnotrya cerebriformis	MICH 67763	NA	AZ	1996	dried ×2	0
H. cf. *cubispora*	AN 043557	NA	AZ	2015	dried	0
H. cubispora	FLAS-F-62913	NA	Scotland	2018	dried	0

H. michaelis	MICH 70028	NA	UT	1995	dried	0
Leotia lubrica	MICH 352041	ACD0314	MI	2020	dried, frozen	0
Morchella punctipes	NA	iNat #46602042	WI	2020	frozen	0
Pachycudonia monticola	MICH 352040	ACD0312	OR	2020	dried, frozen	0
Scutellinia sp.	MICH 352036	ACD0295	MI	2020	dried	0
Sphaerosporella brunnea	MICH 352042	ACD0317	Canada	2020	dried, frozen, culture (PDA)	0
Urnula craterium	NA	iNat #42365327	WI	2020	frozen	0

Chapter 3 Taxonomic Revisions and Insights into the Ecology and Mating Systems of Lorchels (*Discinaceae*, *Pezizales*) Using Whole Genome Analyses

Abstract

Lorchels, also known as false morels (*Gyromitra* spp.), are iconic due to their brain-shaped mushrooms and production of gyromitrin, a deadly mycotoxin. Molecular phylogenetic studies have hitherto failed to resolve deep-branching relationships in the lorchel family, *Discinaceae*, hampering our ability to settle longstanding taxonomic debates. We newly generated 73 draft genomes from cultures and mushroom samples as old as 1960 that covered all major clades in *Discinaceae* and conducted phylogenomic analyses with 1542 BUSCO gene families to infer the early evolutionary history of lorchels. In addition, we used comparative genomics to search for signatures of trophic status and mapped mating-type loci to better understand lorchel reproductive biology. Our phylogenomic tree was supported by high gene tree concordance, facilitating genus-level taxonomic revisions in *Discinaceae*. We recognized a single genus, *Gyromitra*, encompassing 10 subgenera; two subgenera were newly erected and 21 new combinations were formalized. *Paradiscina melaleuca* and *Marcelleina donadinii* formed their own family-level clade sister to *Morchellaceae.* Genome size and CAZyme content agreed with a mycorrhizal lifestyle for the sequestrate species formerly classified as *Hydnotrya*, while the other clades possessed genomic properties of a saprotrophic habit. *Gyromitra* species were found to be predominantly heterothallic—either MAT1-1 or MAT1-2—but a single occurrence of colocalized mating type idiomorphs indicative of homothally was observed in *Gyromitra esculenta* strain CBS101906 and will require additional confirmation follow-up study. Mating-

type gene coverage data indicated that hypogeous and epigeous ascomata differ in the physical association of the two parental strains contributing to the ascocarp. Our genomic dataset will facilitate investigations into the gyromitrin biosynthesis genes and their evolutionary history. Additional sampling of *Geomoriaceae* and *Helvellaceae*, closely related families with no publicly available genomes, could facilitate comprehensive studies on the evolution of hypogeous growth and ecological diversification in a morphologically intriguing and economically important group of *Pezizales* fungi.

3.1 Introduction

Lorchels (*Gyromitra* spp.) are distinctive mushroom-forming fungi with a long taxonomic history and substantial cultural lore (Beug, 2014). They are often called false morels due to their close relationship to the prized true morels (*Morchella* spp.). False morel, however, is a vague name that is used for several *Pezizales* groups including *Verpa* spp. (*Morchellaceae*) and *Helvella* spp. (*Helvellaceae*). Thus, like the late Nancy Smith Weber, we favor the common name "lorchel" for *Gyromitra* species (Weber, 1995). Some lorchels contain the mycotoxin gyromitrin (acetaldehyde *N*-methyl-*N*-formylhydrazone), which can cause severe illness, neurodegenerative disease, and death (Dirks et al., 2023; Lagrange et al., 2022, 2021; List and Luft, 1969, 1968; Michelot and Toth, 1991). In a recent survey of the lorchel family *Discinaceae*, gyromitrin was shown to have a limited and discontinuous distribution suggestive of horizontal gene transfer (Dirks et al., 2023). Further investigations into the genetics and evolution of gyromitrin have been hampered by a scarcity of sequenced *Discinaceae* genomes, poor phylogenetic resolution, and unstable taxonomy—issues that we address in the present study.

Discinaceae is a morphologically diverse family of fungi that produce discoid ascomata; stipitate ascomata with a convoluted apical hymenophore; or hypogeous (underground) to emergent (partially exposed) sequestrate ascomata (Figure 3.1). Based largely on gross morphology, these taxa have been traditionally placed in the genera *Discina* and *Paradiscina* (discoid); *Gyromitra*, *Pseudorhizina*, and *Neogyromitra* (stipitate); or *Hydnotrya* (sequestrate), with several other catch-all or little used names employed throughout the centuries. For a comprehensive synopsis of the early taxonomic history, see Methven *et al.* (2013) and Wang *et al.* (2023). Harmaja recognized a continuum in microscopic traits across the epigeous genera, prompting their synonymization with *Gyromitra* (Harmaja, 1976a, 1976b, 1969). Other mycologists agreed with Harmaja's broad framework and combined additional discoid species with *Gyromitra* (Koch et al., 1996; Pfister, 1980). Consequently, in current times, *Gyromitra* is the single genus name in common usage for the epigeous *Discinaceae* species.

Early molecular phylogenetic studies determined that the "false morels" and "truffles" were polyphyletic. *Discinaceae* was emended to include *Gyromitra sensu lato* and the sequestrate genus *Hydnotrya* (Landvik et al., 1997; O'Donnell et al., 1997). This, however, resulted in *Gyromitra* being paraphyletic, an issue that remained unresolved until Wang *et al.* split the family into eight genera (2023). Alternative taxonomic frameworks could not be adequately explored due to low bootstrap support among some basal branches in their phylogenetic trees. All *Discinaceae* multi-locus phylogenetic studies have encountered these recalcitrant nodes (Dirks et al., 2023; Methven et al., 2013; Miller et al., 2015; Wang et al., 2023; Wang and Zhuang, 2019). For difficult groups such as this, phylogenomics is a powerful tool to infer a robust backbone tree (Hensen et al., 2023; Steenwyk et al., 2019; Tremble et al., 2024; Vandepol et al., 2020; Zhang et al., 2017). Denser sampling of conserved single-copy gene sets

can increase phylogenetic accuracy and allow us to resolve taxonomic relationships in *Discinaceae* (Binder et al., 2013; Rokas and Carroll, 2005; Salgado et al., 2023). Due to the difficulty of culturing certain fungi and the rarity of some mushrooms, fungarium specimens are a crucial resource for the sampling of diverse taxa (Andrew et al., 2018; Hawksworth, 2010). Short-read sequencing technologies and methodologies developed for the analysis of highly fragmented "ancient DNA" further expand the potential sampling space to preserved collections that are decades to centuries old (Bakker et al., 2020; Yoshida et al., 2014).

Beyond phylogenomics, genomes provide opportunities to learn about fundamental aspects of an organism's biology such as trophic mode and reproduction (Díaz-Escandón et al., 2022; Melie et al., 2023; Murat et al., 2018; Zhao et al., 2023). To date, little is known about lorchel ecology. Stable isotope analyses, ectomycorrhizal synthesis studies, and environmental DNA sequencing have identified *Hydnotrya* as ectomycorrhizal with angiosperms, gymnosperms, and orchids, and the epigeous *Discinaceae* as putative soil or wood-decay saprotrophs (Godbout and Fortin, 1985; Healy et al., 2022; Hobbie et al., 2001; Tedersoo et al., 2010; Zuev et al., 2023). Some of these putative saprotrophs have also been found growing within lichen thalli and as endophytes within gymnosperms and bryophytes, an ecological dynamism akin to *Morchella* species (Healy et al., 2022). Many endophytes may be latent saprotrophs, forming facultative associations with plants and switching to saprotrophism for first access to the host tissue upon senescence, a multi- trophic lifestyle deemed "viaphytism" (Chen et al., 2022; Nelson et al., 2020; Promputtha et al., 2007; Szink et al., 2016). With little experimental data available on the ecology of *Discinaceae* species, genomic signatures of ecological niche—such as genome size, carbohydrate-active enzyme (CAZyme) content, and

secondary metabolite repertoire— could provide insights into the diversity and evolution of trophic modes in lorchels (Melie et al., 2023).

Regarding lorchel reproduction, nearly nothing is known, but paradigms in ascomycete sex are expected to apply equally well to lorchels. Mating in ascomycetes is typically governed by a single mating-type (*MAT*) locus and the variable presence of two *MAT* idiomorphs, *MAT1-1* and *MAT1-2* (James, 2023). The *MAT* genes encode transcription factors with conserved DNA-binding domains that regulate the expression of hundreds of other sex-related genes (Du and Yang, 2021). Heterothallic individuals contain one *MAT* idiomorph; they are self-incompatible and must outcross with an individual of the opposite *MAT* idiomorph. In contrast, homothallic strains contain both *MAT* idiomorphs and can reproduce with any individual in the population, including themselves. In *Morchella* and *Verpa (Morchellaceae)*, both heterothallism and homothallism have been documented, as well as pseudohomothallic behavior in which self-compatibility is achieved by packaging nuclei of opposite mating types together in one ascospore (Chai et al., 2022; Du and Yang, 2021; Sun et al., 2023). Mapping *MAT* idiomorphs in *Discinaceae* genomes could shed light on the basic reproductive strategies in this understudied family.

In summary, there is much we do not know about the evolutionary history and basic biology of lorchels. In this study, we generated 73 new genomes with the aim of 1) resolving the deep-branching relationships in *Discinaceae* and revising genus-level taxonomy, 2) elucidating trophic status, and 3) mapping mating systems in lorchels. These genomic resources will not only aid future investigations into gyromitrin but also support comparative genomics studies on the ecological diversification and convergent evolution of the truffloid form in *Pezizales* fungi.

3.2 Materials and methods

3.2.1 Specimen acquisition and DNA sequencing

Two genomes were generated by the U.S. Department of Energy Joint Genome Institute (JGI) according to the methods described in Supplement. The remaining 71 genomes were generated according to the methods described here in sections 3.2.1 and 3.2.2. Ascomata were collected from the field, donated by community scientists, or loaned by fungaria. Fresh specimens were dried in an electric food dehydrator and deposited at the University of Michigan fungarium (MICH). Cultures were derived from single spore isolates according to Dirks et al. (2023). Additional cultures were obtained from the Bonito Lab culture collection at Michigan State University and the Robert L. Gilbertson Mycological Herbarium culture collection at University of Arizona (Healy et al., 2022). Of special note, molecular phylogenetic studies of *Marcelleina* (*Pezizaceae*) revealed an affinity between *Marcelleina donadinii* and *Discinaceae*, prompting the inclusion of this taxon in our analysis (Van Vooren *et al*., unpublished). Total DNA was extracted with the 2× CTAB method (Dirks et al., 2023; Porter et al., 2011). Library preparation and Illumina NovaSeq 151 bp paired-end sequencing were completed by the University of Michigan Advanced Genomics Core. A list of specimens sequenced for this study and their metadata are presented in Supplementary Table 3.1.

3.2.2 Genome assembly and annotation

Raw read quality was checked with FastQC 0.11.9 (github.com/s-andrews/FastQC). Adapter sequences and low-quality regions were trimmed with Trimmomatic 0.36 (Bolger et al., 2014). Paired sequences were assembled with SPAdes 3.15.5 using the “isolate” and “meta” assembly options for cultures and ascocarp samples, respectively (Bankevich et al., 2012).

Ascocarp assemblies were further processed with SCGid 0.9b0 and the Foreign Contamination Screen tool suite of the National Center for Biotechnology Information (NCBI) to remove contaminant contigs (Amses et al., 2020; Astashyn et al., 2023). Assembly quality and statistics were evaluated with QUAST 5.2.0 and BUSCO 5.3.0 (Gurevich et al., 2013; Simão et al., 2015). Genome coverage was calculated by mapping raw reads to the assembled genomes with BBTools 38.96 (sourceforge.net/projects/bbmap), and rDNA sequences were extracted with a modified script (github.com/aldendirks/Extract-ITS-sequences-from-a-fungal-genome) implementing Barrnap 0.9 (github.com/tseemann/barrnap) and ITSx 1.1.3 (Bengtsson- Palme et al., 2013). Gene prediction and annotation were executed using the Funannotate 1.8.14 pipeline (github.com/nextgenusfs/funannotate) employing antiSMASH 7.1.0, eggNOG-mapper 2.1.9, and InterProScan 5.60.92.0, among other software and databases (Blin et al., 2023; Buchfink et al., 2021; Cantalapiedra et al., 2021; Drula et al., 2022; Eddy, 2011; Huerta-Cepas et al., 2019; Jones et al., 2014; Steinegger and Söding, 2017; The UniProt Consortium, 2023). Custom scripts were used to tabulate secondary metabolite gene clusters and CAZyme family counts from the annotations.

3.2.3 Phylogenomics and time divergence estimation

The BUSCO phylogenomics pipeline was used for coalescence and concatenation (supermatrix) phylogenomic analyses on high-quality *Discinaceae* genomes and outgroup genomes (github.com/jamiemcg/BUSCO_phylogenomics). High-quality genomes were defined as ones with greater than 90% of *Ascomycota* BUSCOs present as single copies and more than 15-times average genome coverage. Outgroup genomes were downloaded from the Joint Genome Institute's MycoCosm portal and NCBI (Table S1) (Grigoriev et al., 2014; Sayers et al., 2022). BUSCOs present in at least four samples (coalescence analysis) or 70 samples

(supermatrix analysis) were independently aligned with MUSCLE 3.8.31 and trimmed with trimAl 1.4.rev22 (Capella-Gutiérrez et al., 2009; Edgar, 2004). For each alignment, the best protein substitution model (constrained to WAG, LG, and JTT) was selected with ModelFinder, and a phylogeny was inferred using IQ-TREE2 2.2.6 (Kalyaanamoorthy et al., 2017; Minh et al., 2020b). The coalescence species tree was created with ASTRAL-III 5.15.5 (Zhang et al., 2018). The concatenation analysis proceeded with IQ-TREE2 2.2.6 using 1000 ultrafast bootstrap replicates and an edge-linked proportional partition model (Chernomor et al., 2016; Minh et al., 2020a; Mo et al., 2023). IQ-TREE2 was also used to calculate gene concordance factors (gCF), maximum- likelihood site concordance factors (sCF), and time divergence estimates with the integrated LSD2 least-squares methodology (To et al., 2016). A divergence time of 46.27 Mya for the *Tuberaceae* taxa *Choiromyces venosus* and *Tuber magnatum* was used as a calibration point (Díaz-Escandón et al., 2022).

3.2.4 CAZyme analysis, mating systems, and ancestral state reconstruction

Significant contractions and expansions of CAZymes and secondary metabolite gene clusters across the phylogenomic tree were evaluated with CAFE5 (Mendes et al., 2021). One to five gamma rate categories were tested, and runs were replicated five times to ensure lambda convergence. Additional statistical analyses and visualizations were conducted in R 4.2.2 (R Core Team, 2022). Phylograms and cladograms were plotted with ggtree 3.6.2 (Yu et al., 2017). Correlations between continuous genome traits were evaluated with the phylogenetic generalized least squares method from nlme 3.1-164 (Mundry, 2014; Pinheiro et al., 2023). Mating systems in lorchels were determined by first identifying homologs of *Morchellaceae MAT1-1* and *MAT1-2* idiomorphs in the MycoCosm *Gyromitra esculenta* genome and then using BLASTP to find homologs of the *Gyromitra esculenta MAT* loci in other *Discinaceae* genomes (Altschul et al.,

1990; Camacho et al., 2009; Du and Yang, 2021; Sun et al., 2023). *MAT* idiomorph coverage was calculated as above for genome coverage. Gene synteny maps were plotted with clinker 0.0.28 (Gilchrist and Chooi, 2021). Gyromitrin distribution data were taken from Dirks *et al*. (2023) and Lagrange *et al.* (in press). Gyromitrin ancestral state reconstruction was performed with stochastic character mapping from phytools 2.1-1 (Revell, 2024). One thousand character stochastic maps were sampled from four continuous-time Markov chain models according to the models' relative support, and the posterior probabilities were plotted on the internal nodes of the phylogenomic tree.

3.3 Results

In addition to the two JGI genomes, we generated 73 new draft genomes from 10 cultures and 63 ascomata, 10 of which (16%) did not meet our quality standards of greater than 90% of *Ascomycota* BUSCOs present as single copies and more than 15-times average genome coverage. Ascomata collected prior to the year 2000 were more frequently of low quality (7/17, 41%). The oldest high-quality genome belonged to *Gyromitra* sp. 13, a putative undescribed species in subgenus *Gigas*, collected in Alaska in 1960 by Virginia Wells and Phyllis Kempton (97.2% BUSCO completeness, 125.6 × average coverage) (Kempton and Wells, 1973; Miller et al., 2022, 2020; Wells and Kempton, 1968). However, even low-quality genomes from old specimens, including type specimens, could be successfully mined for rDNA sequences, making them valuable resources even if they were not used for phylogenomics (Supplementary Table 3.1).

Our final dataset consisted of 62 *Discinaceae* genomes and 16 outgroup genomes: seven *Morchellaceae*, six *Tuberaceae*, and three *familia incertae sedis* (*Paradiscina melaleuca* and *Marcelleina donadinii*). Coalescent and concatenation analyses were conducted with 1,695 and

1,542 BUSCOs, respectively, and yielded identical trees (Figs 2, S1). Bootstrap support was 100% across all branches, except for one within the *Gyromitra venenata* clade (70%). This branch was also the only one that had smaller gCF support than alternative arrangements of the quartet. In this case, there appeared to be uncertainty regarding whether the Eastern North American *G. venenata* samples were more closely related to the European sample or the Western North American sample.

Gene concordance was generally greater than 50%, indicating most branches in the inferred species tree were highly supported and accurately reflected the underlying gene phylogenies (Figure 3.2). A few short basal branches—as well as a few branches among closely related species towards the tips—scored lower gCF values, which indicate conflicting signals in the gene trees. The lowest gCF (24%) belonged to the branch leading to the most recent common ancestor of subg. *Pseudoverpa* and the other subgenera in the lower portion of the tree. A total of 363 gene trees supported the quartet shown in the concatenated species tree, while 238 and 186 gene trees supported the two alternative arrangements of the quartet. A plurality of gene trees was polyphyletic, meaning the quartet clades were not retrieved as monophyletic groupings (738 gene trees, 48%). In general, discordant gene trees corresponding to less supported branches were typically highly polyphyletic and provided little consensus for an alternative topology (Supplementary Figure 3.2). Site concordance factors generally provided less support than gCF. The branch leading to the most recent common ancestor of subg. *Hydnotrya* and subg. *Paragyromitra* had a greater gCF than alternative arrangements of that quartet (40% compared to 20% and 12%) but a smaller sCF than the alternative hypothesis in which subg. *Gyromitra* and subg. *Paragyromitra* formed a monophyletic clade (30% compared to 37%).

Discinaceae had an estimated crown age of 81 Mya and showed an early bifurcation leading to two major clades (Figure 3.3). These are informally called the "top clade" (subgenera *Discina*, *Caroliniana*, *Gigas*, *Leucoxantha*, and *Melaleucoides*) and the "bottom clade" (subgenera *Hydnotrya*, *Paragyromitra*, *Gyromitra*, *Pseudorhizina*, and *Pseudoverpa*). These clades had notable differences in GC content. Taxa in the top clade typically had higher GC content ($\bar{x}$ = 49.4%) while those in the bottom clade typically had lower GC content ($\bar{x}$ = 46.5%). *Gyromitra* genomes were an average of 46.14 Mbp in size with a mean of 11,551 genes. Genome size and gene content was conspicuously elevated in subg. *Hydnotrya* (mean of 63.45 Mbp and 12,756 genes). Across *Gyromitra* taxa, gene count was positively correlated with the log-transformed number of transposons (p = .01) (Supplementary Figure 3.3). In the CAFE5 analysis of gene family expansion and contraction, lambda failed to converge for CAZyme classes when allowing for multiple gamma rate categories (k) but converged at k = 3 for individual CAZyme families and secondary metabolite families. No secondary metabolite families or CAZyme classes were determined to have undergone significant changes in *Discinaceae*. On the other hand, the CAZyme families AA7 (glucooligosaccharide oxidases) and AA9 (lytic cellulose monooxygenases, formerly classified as GH61) experienced significant contraction along the branch leading to subg. *Hydnotrya*, the only ectomycorrhizal subgenus (Figure 3.3).

The *MAT* locus consisted of the mating-type idiomorphs *MAT1-1* and *MAT1-2* and was flanked by the highly conserved genes *APN1* and *SLA2,* as well as a DEAD box helicase, *SHY1*, ubiquitin, and two hypothetical protein-encoding genes. The gene synteny was overall largely conserved among the top-clade taxa and bottom-clade taxa (Figure 3.4). *Gyromitra* was found to be overwhelmingly heterothallic—that is, usually only one *MAT* idiomorph was present (Supplementary Figure 3.4). All cultures contained a single *MAT* idiomorph, except for

Gyromitra esculenta CBS101906, which was the only sample with colocalized *MAT* idiomorphs. Ascomata typically contained a single *MAT* idiomorph. When they contained two, one was flanked by conserved genes associated with the *MAT* locus and the other was isolated on a small contig. In comparing the raw read coverage of the two *MAT* idiomorphs in ascoma genomes, the hypogeous taxa had a ratio close to one whereas the epigeous taxa had a ratio typically much greater than one (Figure 3.4). *Gyromitra leucoxantha* and *G. olympiana* were initially found to contain no *MAT* idiomorphs but had a conspicuous lack of predicted genes where the *MAT* locus was expected. For both taxa, a BLASTX search against the *G. esculenta* genome revealed *MAT1-1* idiomorphs in the expected location.

Gyromitrin was detected in all tested specimens from the distantly related subg. *Gyromitra* and subg. *Leucoxantha* clades but not in any other tested group. An irreversible model in which taxa can evolve to produce gyromitrin but do not lose this ability once gained was the best supported model of character transition evolution in our ancestral state reconstruction. The posterior probabilities of gyromitrin presence or absence at each internal node is presented in Supplementary Figure 3.5. The probability of gyromitrin production in the root of *Discinaceae* was estimated at 12.7%.

3.4 Discussion

3.4.1 Phylogenomics and taxonomy

We resolved the deep branching relationships within *Discinaceae*, facilitating comprehensive taxonomic revisions. *Discinaceae* has been reduced to a single genus, *Gyromitra*, comprising 10 subgenera. The use of subgenera facilitates further divisions of speciose groups into sections if deemed practical in the future (Li et al., 2023). A broad and monophyletic *Gyromitra* necessitated the synonymization of *Hydnotrya*, a derived sequestrate clade most

closely related to subg. *Paragyromitra*. Since *Hydnotrya* is older than *Gyromitra*, a conservation proposal has been submitted (Miller *et al*., unpublished). There is strong precedence for subsuming sequestrate taxa into larger, more familiar genera, reflecting high evolvability of gross morphology in mushroom-forming fungi (Vellinga et al., 2015). For example, the *Pezizales* family *Geomoriaceae* consists of a single genus, *Goemorium*, and contains both epigeous and hypogeous taxa, the latter having independently evolved three times (Kraisitudomsook et al., 2020). Our taxonomic approach is a utilitarian one that preserves familiar names while largely upholding traditional morphology-based taxonomic concepts in *Discinaceae* (Harmaja, 1969).

Alternative taxonomic frameworks could prioritize gross morphology or time divergence in defining *Discinaceae* genera. However, preserving the genus *Hydnotrya* simply due to its sequestrate form would result in a cascade of splits across *Discinaceae* (Wang et al., 2023). If one were to delineate genera based strictly on monophyletic clades of shared macromorphology in this way, one would be left with 13 genera (one sequestrate clade, seven discoid clades, and seven stipitate clades), requiring artificial segregation of congeneric taxa and uncomfortable changes to iconic species. Time-divergence classification, which some argue is a more objective and meaningful taxonomic framework, would be a more reasonable alternative (Zhao et al., 2016). One could define the crown age of the node connecting subg. *Paragyromitra* and subg. *Hydnotrya* as a standard cutoff for delimiting genera in *Discinaceae*. Applying this criterion to our time- calibrated phylogenomic tree, generic delimitations would match the seven-genus taxonomy proposed by Wang *et al.* 2023 (excluding *Paradiscina melaleuca*, which does not belong in *Discinaceae*). Unfortunately, a robust time-divergence classification is severely limited by available fossil data, and the choice of calibration points can have dramatic impacts on

estimated divergence times (Taylor and Berbee, 2006). The resultant groupings may also feel artificial due to deviations from familiar morphological classifications.

3.4.2 Ecology

Consistent with their mycorrhizal habit, species in subg. *Hydnotrya* had markedly larger genomes, greater numbers of transposable elements, and decreased CAZyme content. This is a general pattern for mycorrhizal fungi (Chang et al., 2019; Hess et al., 2014; Kohler et al., 2015; Murat et al., 2018). Transposable elements may drive gene family expansion and thus be an agent of adaptation for symbiotic fungi (Lanfranco and Bonfante, 2023; Sánchez-Vallet et al., 2018). While not significant according to CAFE5, perhaps due to a single lambda modeling a wide numerical range, four CAZyme classes (CAZymes with auxiliary activities, carbohydrate esterases, glycoside hydrolases, and polysaccharide lyases) appeared substantially reduced in subg. *Hydnotrya*; notably, polysaccharide lyases were nearly absent in the hypogeous *Gyromitra* species. Considering the significant contractions observed in AA7 and AA9 CAZyme families, subg. *Hydnotrya* has a reduced enzymatic toolkit to degrade lignocellulose compounds compared to other *Gyromitra* clades (Levasseur et al., 2013). All the other *Gyromitra* subgenera were fairly uniform in their CAZyme content and had traits consistent with a saprotrophic lifestyle.

In our attempts to culture *Gyromitra* (Sow et al., 2024), we observed that spores belonging to species from the bottom clade readily germinate on potato dextrose agar and malt extract agar but those belonging to species from the top clade do not. Interestingly, this pattern is consistent with that observed in the culturing of endolichenic and endophytic *Pezizales* fungi (Healy et al., 2022). Subgenera *Paragyromitra*, *Gyromitra*, and *Pseudorhizina* have been isolated from bryophytes, gymnosperms, and lichen thalli; on the other hand, no top-clade taxa have been isolated from plant or lichen tissue. While the epigeous species of the two major

clades appear similar in their CAZyme profiles, these ecological nuances might be detected in other ways, for example in their repertoires of small secreted proteins. The prevalence of endophytes in bottom-clade taxa and the phylogenetic position of the derived ectomycorrhizal clade raises the interesting hypothesis that facultative endophytism facilitates the evolution of ectomycorrhizal associations by priming fungi for an obligate relationship with plants.

3.4.3 Mating systems

Lorchels have typical *MAT* loci and are predominantly heterothallic. All cultures contained a single *MAT* idiomorph apart from *Gyromitra esculenta* strain CBS101906, whose *MAT* locus consisting of colocalized *MAT1-1* and *MAT1-2* idiomorphs is indicative of homothally. While the presence of unlinked *MAT* idiomorphs as observed in some ascoma assemblies is an atypical but possible gene organization in homothallic organisms (Rydholm et al., 2007), the fact that conspecific strains were heterothallic implies this was likely a result of sequencing multiple parent strains. Indeed, one would expect to find both *MAT* idiomorphs in an ascoma, the absence of which could indicate unisexual reproduction. *Discinaceae* ascomata whose genomes contained a single *MAT* idiomorph were found to produce ascospores (data not shown), which is inconsistent with unisexual reproduction in morels in which ascocarps are formed by homokaryons but no ascospores are produced (Du and Yang, 2021). Therefore, the presence of a single *MAT* idiomorph in ascoma genomes is likely an assembly artifact caused by low read coverage rather than a consequence of unisexual reproduction.

When both *MAT* idiomorphs were present in ascoma genomes, read coverage ratio variation suggested that *Gyromitra* species organize their ascomata differently depending on their growth form. Specifically, hypogeous taxa are expected to form type I ascomata (both mating types equally distributed in the ascoma) whereas epigeous taxa are expected to form type

II ascomata (one primary mating type idiomorph forms the sterile tissue of the ascoma but both are present in the hymenium) (Du and Yang, 2021). Future research could use *MAT*-specific primers and qPCR to map the distribution of *MAT* idiomorphs in lorchels both within (stipe compared to hymenophore) and between (relative frequency of heterothallism and homothallism) samples. Increased sampling of *Gyromitra esculenta* is needed to elucidate the prevalence of homothallism in this species. In morels, a strong preference for strains carrying the *MAT1-1* idiomorph to form the bulk of the ascoma has been documented (Du and Yang, 2021). The preponderance of *MAT*-specific sexual roles within the *Pezizales* is an important topic of investigation to better understand fungal reproduction.

3.4.4 Gyromitrin ancestral state reconstruction

Gyromitrin was found to have a small but nonzero chance of being present in the most recent common ancestor of all *Gyromitra* species. Assuming this to be true, and assuming gyromitrin production is constrained to subg. *Gyromitra* and subg. *Leucoxantha*, five loss events would be required to explain its distribution—a seemingly unlikely evolutionary history but not one without precedent (Patron et al., 2007; Reynolds et al., 2017). Alternatively, the ability to produce gyromitrin may have been independently gained; transferred horizontally; or a product of interspecific hybridization (Du and Yang, 2021; Steenwyk et al., 2023). Comprehensive metabolomics of *Gyromitra* species is needed to understand the full chemodiversity of N-N bond-containing natural products (N2NPs) produced by lorchels. One other N2NP, namely 1-(2-hydroxyacetyl)-pyrazol, has been characterized from *Gyromitra grandis*, but this finding deserves contemporary validation (Jurenitsch et al., 1988). Discovery of additional N2NPs could greatly improve genome mining of fungal N2NP biosynthesis genes. With our broad sampling of

lorchel genomes, we are better poised to elucidate the genes responsible for the biosynthesis of gyromitrin and other fungal N2NPs.

3.4.5 Closing remarks

We used phylogenomic analyses to infer a robust backbone tree for *Discinaceae* and revise genus-level taxonomy. Fungarium specimens, including ones over 60 years old, were crucial in our broad sampling of *Discinaceae*. To our surprise, *Paradiscina melaleuca* and *Marcelleina donadinii* formed a monophyletic clade sister to *Morchellaceae* of unknown family identity. This scenario is an important reminder that proper outgroup selection is crucial for accurate phylogenetic analysis. If we had not included *Tuberaceae* and set *Morchellaceae* as the outgroup, *Paradiscina melaleuca* would have been inaccurately classified within *Discinaceae*. Broader sampling of this understudied lineage is needed to determine if they belong to *Morchellaceae* or constitute their own taxonomic family. In any case, *Marcelleina donadinii* will need to be revised as the type species of the genus, *Marcelleina persoonii*, belongs to *Pezizaceae* (Hansen et al., 2001; Hansen and Pfister, 2006; Moravec, 1987). The sequencing of genomes from *Geomoriaceae* and *Helvellaceae*, two closely related families with no publicly available genome data, would greatly benefit comparative genomics studies on the ecological diversification and evolution of sequestrate forms in the *Pezizales*. Indeed, the many independent origins of ectomycorrhizal truffloid fungi and associated changes in genome properties suggest potent selection pressure is at play, causing evolutionary convergence at several biological scales.

Acknowledgements

We extend our utmost gratitude to those mycologists—seeking, collecting, and laboring, often with no compensation and for no motive besides their curiosity and passion—who donated invaluable specimens for this research. In alphabetical order, these people are: Brittany Marcotte (*Gyromitra anthracobia*), Emmeline Lagrange (*G. venenata*), Eric Chandler (*G. infula*, *G. splendida*), Dave Ramos (*Gyromitra* sp. 9), Dave Wasilewski (*G. ancilis*), Drew T. Henderson (*G. mcknightii*, *G. venenata*), Garrett Taylor (*G. korfii*, *G. leucoxantha*), Huafang Su (*G. brunnea*), Jacques Guinberteau (*Paradiscina melaleuca*), Jeff Volpert (*G. brunnea*), Jen Chandler (*G. infula*, *G. splendida*), Jon Shaffer (*G. caroliniana*), Jonathan Frank (*G. montana*), Leah Bendlin (*G. melaleucoides*), Mical Moser (*G. venenata*), Michael W. Beug (*G. persicula*, *Gyromitra* sp. 4, *Gyromitra* sp. 6, *Gyromitra* sp. 11), Rubén Martínez Gil (*Marcelleina donadinii*), Steve Ness (*G. californica*), and Vail Paterson (*Gyromitra* sp. 2). We thank Nicolas Van Vooren for thoughtful discussions on *Discinaceae* taxonomy, facilitating the material transfer of *Paradiscina melaleuca*, and notifying us of the taxonomic relationship of *Marcelleina donadinii*. Fungarium curators and their collections are indispensable for this research. For managing safely and securely treasure troves of preserved mushrooms and living cultures, and for providing access to these collections for our studies, we thank Joseph Myers and Ming-Min Lee at the University of Arizona (ARIZ); Matthew Smith, Rosanne Healy, and Benjamin Lemmond at the University of Florida (FLAS); and Alison Harrington at the University of Michigan (MICH). Bettina Haberl and Till R. Lohmeyer graciously organized the donation of several exceptional specimens from the Technical University of Munich and the Inn-Salzach Mycology Working Group (pilze-insa.de) collected by Bernd Fellmann and others. This research was supported by the Michigan Translational Research and Commercialization (MTRAC)

Innovation Hub for AgBio; the University of Michigan Rackham Graduate School and Department of Ecology and Evolution; and the Canadian Institute for Advanced Research (CIFAR). Genome sequencing (proposals 10.46936/10.25585/60001019 and 10.46936/10.25585/60001060) conducted by the U.S. Department of Energy (DOE) Joint Genome Institute (https://ror.org/04xm1d337) was supported by the DOE Office of Science under contract number DE-AC02-05CH11231. T.Y.J. is a fellow of the CIFAR research program "Fungal Kingdom: Threats & Opportunities." During the preparation of this work, the authors used the Generative Expand tool by Adobe Photoshop 2024 to expand the background of closely cropped images in Figure 3.1. This had no effect on the presentation of the focal subject of the photograph. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Figures



Figure 3.1. Morphological diversity of *Discinaceae* and related taxa. A–M. Species belonging to *Discinaceae*. N–O. Species belonging to *familia incertae sedis*. Subgenus *Discina*: A) *Gyromitra ancilis* MICH352026; B) *Gyromitra mcknightii* MICH346490. Subgenus *Caroliniana*: C) *Gyromitra brunnea* MICH352018; D) *Gyromitra martinii* MICH346002. Subgenus *Gigas*: E) *Gyromitra korfii* MICH352049; F) *Gyromitra* sp. 2 MICH346448. Subgenus *Leucoxantha*: G) *Gyromitra leucoxantha* MICH352087. Subgenus *Melaleucoides*: H) *Gyromitra melaleucoides* iNat134134066. Subgenus *Hydnotrya*: I) *Gyromitra bailii* MICH345143. Subgenus *Paragyromitra*: J) *Gyromitra infula* MICH352090. Subgenus *Gyromitra*: K) *Gyromitra venenata* MICH352028. Subgenus *Pseudorhizina*: L) *Gyromitra californica* MICH352088. Subgenus *Pseudoverpa*: M) *Gyromitra anthracobia* MICH346478. *Familia incertae sedis*: N) *Marcelleina donadinii* VIT9785; O) *Paradiscina melaleuca* MICH346528, LY−NV2020.05.32. Photo credits: Alden Dirks (A, C, K), Bernd Fellmann (D, I), Brittany Marcotte (M), Drew Henderson (B), Eric and Jen Chandler (J), Garrett Taylor (E, G), Jacques Guinberteau (O), Rubén Martínez Gil (N), Steve Ness (L), Vail Paterson (F), and Yi-Min Wang (H).

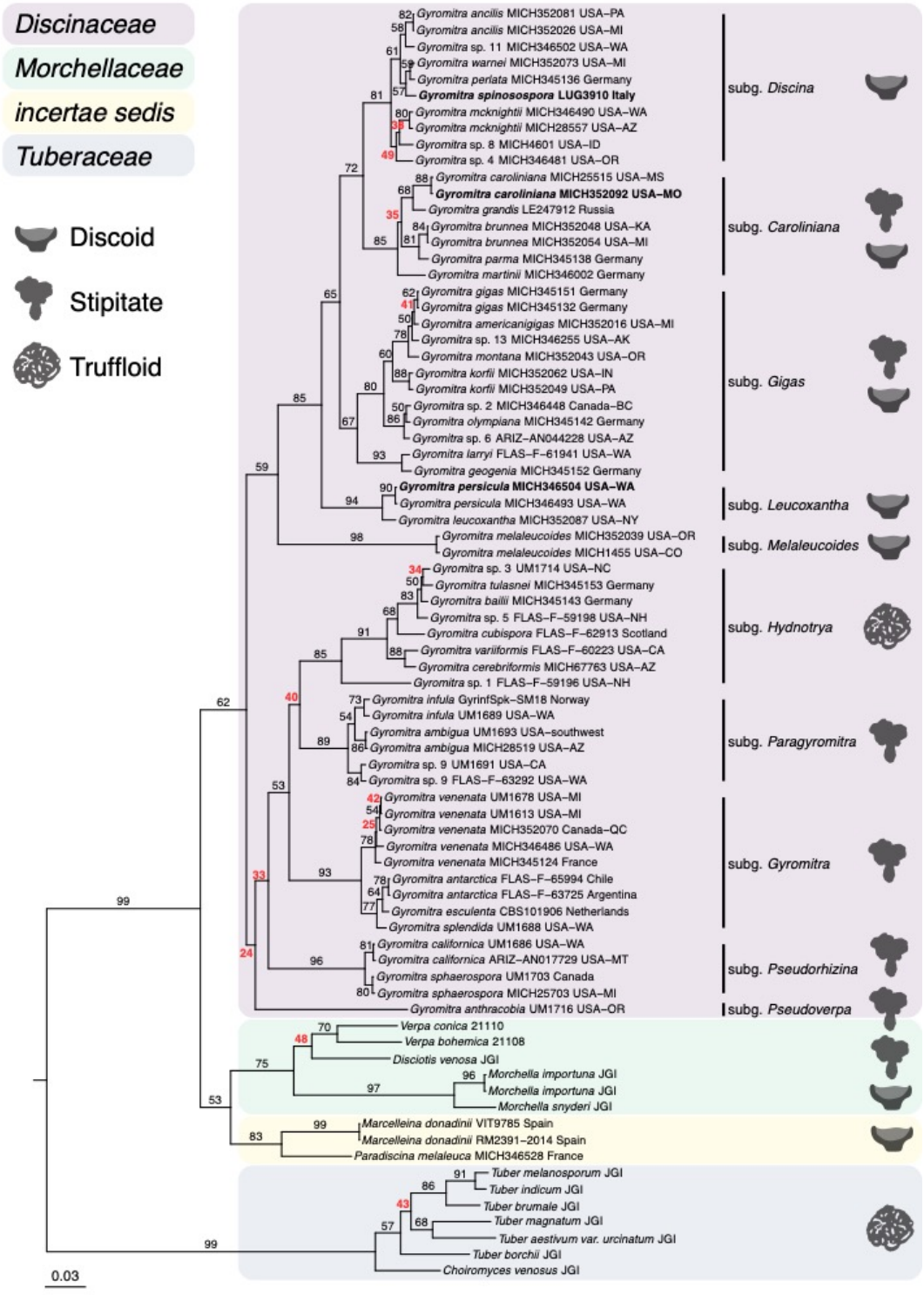


Figure 3.2. Phylogenomic tree of the lorchel family *Discinaceae* and related *Pezizales* families. The phylogenomic tree was inferred with 1542 concatenated BUSCOs that were analyzed using 1000 ultrafast bootstrap replicates and an edge-linked proportional partition model in IQ-TREE2. Gene concordance factors (gCF) are listed on the branches; values less than 50% are colored red. Families are colored according to the legend. *Discinaceae* subgenera are labelled and the diversity of gross morphology represented in our sampling within each subgenus or outgroup family is represented by icons. Sequences derived from type specimens are in **bold**.

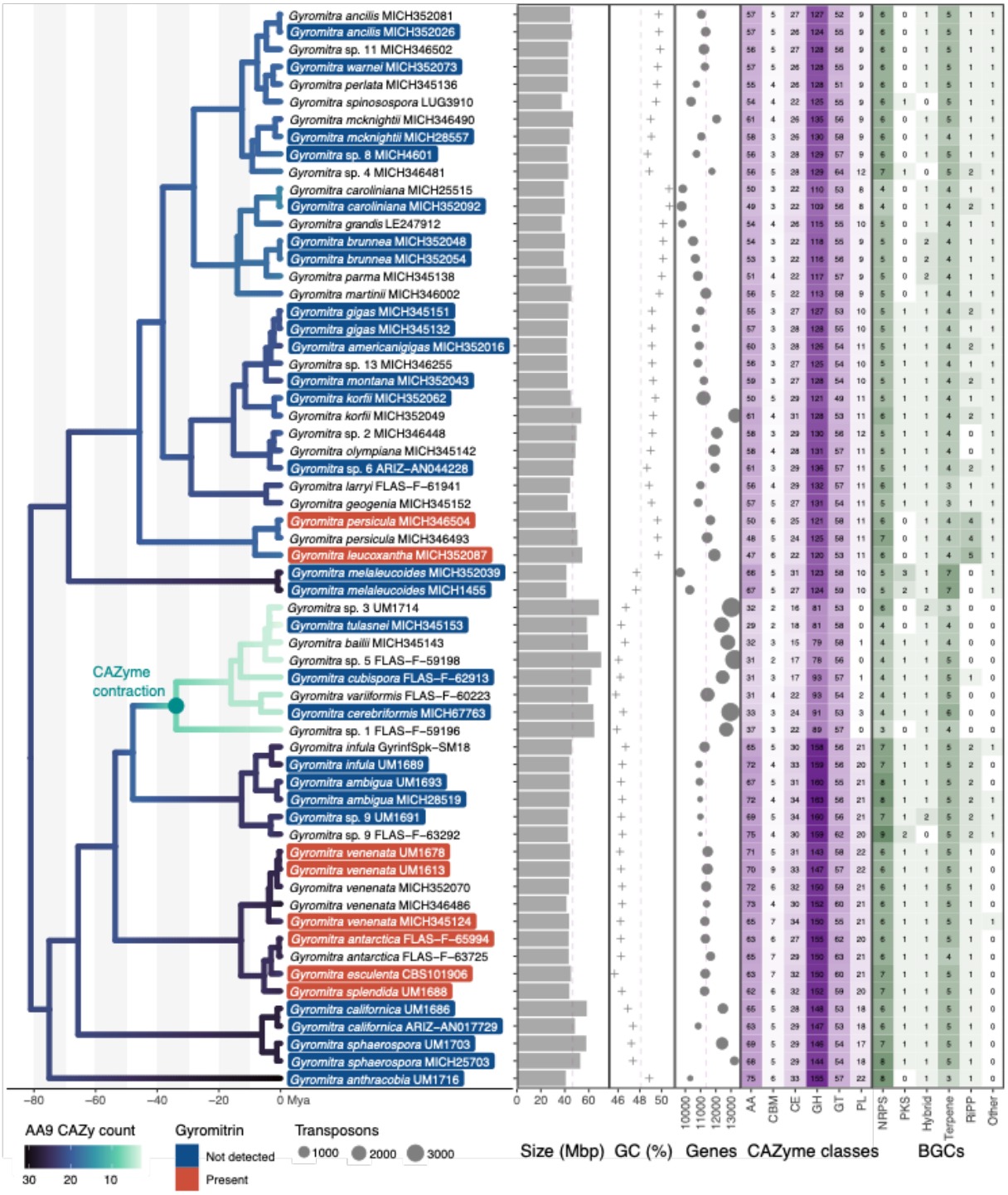


Figure 3.3. *Discinaceae* time divergence estimation, modeling of AA9 CAZyme family evolutionary history, gyromitrin status, and genome traits. Branches are colored according to the AA9 CAZyme count inferred through time with CAFE5. The background color of the tips of the tree indicates gyromitrin status (red: tested positive; blue: gyromitrin not detected; no background color: not tested). Genome size, average GC content, total number of genes sized by number of transposable elements, number of CAZyme classes, and number of biosynthesis gene clusters predicted by antiSMASH are shown for each sample. The dotted lines indicate mean values.

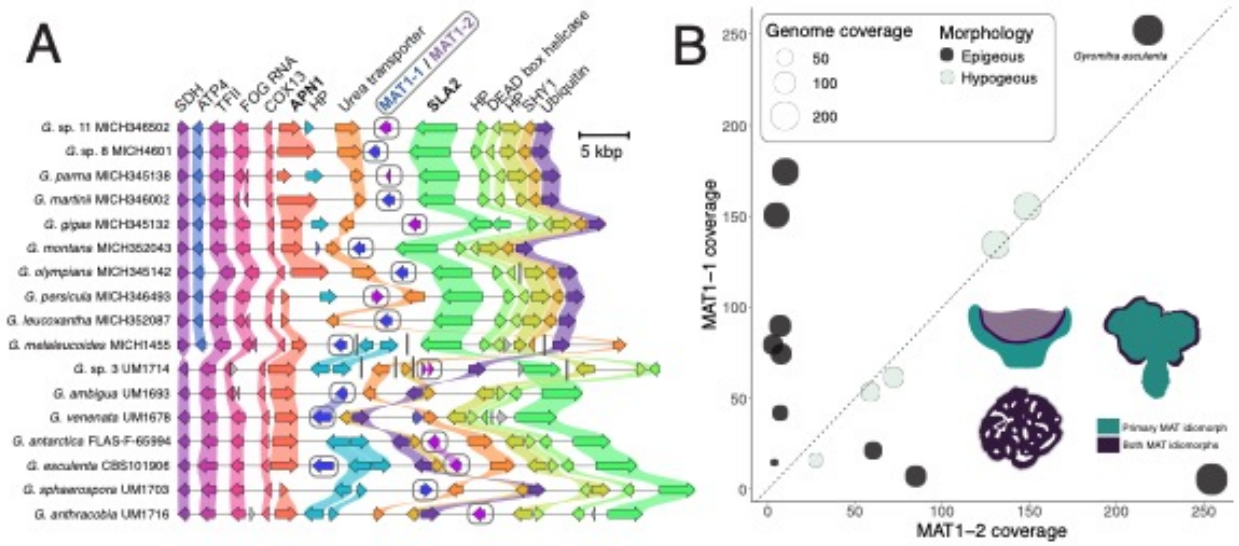


Figure 3.4. *Discinaceae* mating type (*MAT*) locus gene synteny diagram and scatterplot of *MAT* idiomorph coverage. A) *MAT* locus gene diagrams from representatives *Discinaceae* species. Genes are indicated by arrows and homologous genes have the same fill color. Conserved flanking genes of the *MAT* locus are in **bold** (*APN1* and *SLA2*). *MAT* idiomorphs are surrounded by boxes. Vertical lines show breaks in contigs. B) Scatterplot of *MAT* idiomorph coverage for genomes that contained two *MAT* idiomorphs. The size of dots indicates genome coverage. Darker and lighter circles correspond to species that produce epigeous or hypogeous ascomata, respectively. The diagonal dotted line shows a *MAT* idiomorph coverage ratio of one. The inset shows the expected distribution of *MAT* idiomorphs within *Discinaceae* ascomata across different macromorphological forms. Teal corresponds to tissue containing one primary *MAT* idiomorph; purple corresponds to tissue containing both idiomorphs in roughly equivalent numbers.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J Mol Biol 215, 403–410. https://doi.org/10.1016/S0022- 2836(05)80360-2

Amses, K.R., Davis, W.J., James, T.Y., 2020. SCGid: a consensus approach to contig filtering and genome prediction from single-cell sequencing libraries of uncultured eukaryotes. Bioinformatics 36, 1994–2000. https://doi.org/10.1093/bioinformatics/btz866

Andrew, C., Diez, J., James, T.Y., Kauserud, H., 2018. Fungarium specimens: a largely untapped source in global change biology and beyond. Phil. Trans. R. Soc. B 374, 20170392. https://doi.org/10.1098/rstb.2017.0392

Astashyn, A., Tvedte, E.S., Sweeney, D., Sapojnikov, V., Bouk, N., Joukov, V., Mozes, E., Strope, P.K., Sylla, P.M., Wagner, L., Bidwell, S.L., Clark, K., Davis, E.W., Smith-White, B., Hlavina, W., Pruitt, K.D., Schneider, V.A., Murphy, T.D., 2023. Rapid and sensitive detection of genome contamination at scale with FCS-GX. https://doi.org/10.1101/2023.06.02.543519

Bakker, F.T., Bieker, V.C., Martin, M.D., 2020. Editorial: herbarium collection-based plant evolutionary genetics and genomics. Front. Ecol. Evol. 8, 603948. https://doi.org/10.3389/fevo.2020.603948

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., Amend, A., Jumpponen, A., Unterseher, M., Kristiansson, E., Abarenkov, K., Bertrand, Y.J.K., Sanli, K., Eriksson, K.M., Vik, U., Veldre, V., Nilsson, R.H., 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Methods in Ecology and Evolution 4, 914–919. https://doi.org/10.1111/2041-210X.12073

Beug, M.W., 2014. False morels—age-old questions of edibility: a primer by Michael Beug. FUNGI 7, 28–31.

Binder, M., Justo, A., Riley, R., Salamov, A., Lopez-Giraldez, F., Sjökvist, E., Copeland, A., Foster, B., Sun, H., Larsson, E., Larsson, K.-H., Townsend, J., Grigoriev, I.V., Hibbett, D.S., 2013. Phylogenetic and phylogenomic overview of the *Polyporales*. Mycologia 105, 1350–1373. https://doi.org/10.3852/13-003

Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J.N., van Wezel, G.P., Medema, M.H., Weber, T.,

2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research 51, W46–W50. https://doi.org/10.1093/nar/gkad344

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Buchfink, B., Reuter, K., Drost, H.-G., 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18, 366–368. https://doi.org/10.1038/s41592- 021-01101-x

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., Huerta-Cepas, J., 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Molecular Biology and Evolution 38, 5825– 5829. https://doi.org/10.1093/molbev/msab293

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972– 1973. https://doi.org/10.1093/bioinformatics/btp348

Chai, H., Liu, P., Ma, Y., Chen, W., Tao, N., Zhao, Y., 2022. Organization and unconventional integration of the mating-type loci in *Morchella* species. J Fungi (Basel) 8, 746. https://doi.org/10.3390/jof8070746

Chang, Y., Desirò, A., Na, H., Sandor, L., Lipzen, A., Clum, A., Barry, K., Grigoriev, I.V., Martin, F.M., Stajich, J.E., Smith, M.E., Bonito, G., Spatafora, J.W., 2019. Phylogenomics of *Endogonaceae* and evolution of mycorrhizas within *Mucoromycota*. New Phytologist 222, 511–525. https://doi.org/10.1111/nph.15613

Chen, K.-H., Liao, H.-L., Arnold, A.E., Korotkin, H.B., Wu, S.H., Matheny, P.B., Lutzoni, F., 2022. Comparative transcriptomics of fungal endophytes in co-culture with their moss host *Dicranum scoparium* reveals fungal trophic lability and moss unchanged to slightly increased growth rates. New Phytologist 234, 1832–1847. https://doi.org/10.1111/nph.18078

Chernomor, O., von Haeseler, A., Minh, B.Q., 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Systematic Biology 65, 997–1008. https://doi.org/10.1093/sysbio/syw037

Díaz-Escandón, D., Tagirdzhanova, G., Vanderpool, D., Allen, C.C.G., Aptroot, A., Češka, O., Hawksworth, D.L., Huereca, A., Knudsen, K., Kocourková, J., Lücking, R., Resl, P., Spribille, T., 2022. Genome-level analyses resolve an ancient lineage of symbiotic

ascomycetes. Current Biology 32, 5209-5218.e5. https://doi.org/10.1016/j.cub.2022.11.014

Dirks, A.C., Mohamed, O.G., Schultz, P.J., Miller, A.N., Tripathi, A., James, T.Y., 2023. Not all bad: gyromitrin has a limited distribution in the false morels as determined by a new ultra high-performance liquid chromatography method. Mycologia 115, 1–15. https://doi.org/10.1080/00275514.2022.2146473

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., Terrapon, N., 2022. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Research 50, D571–D577. https://doi.org/10.1093/nar/gkab1045

Du, X.-H., Yang, Z.L., 2021. Mating systems in true morels (*Morchella*). Microbiol Mol Biol Rev 85, 1–23. https://doi.org/10.1128/MMBR.00220-20

Eddy, S.R., 2011. Accelerated profile HMM searches. PLOS Computational Biology 7, e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792–1797. https://doi.org/10.1093/nar/gkh340

Gilchrist, C.L.M., Chooi, Y.-H., 2021. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37, 2473–2475. https://doi.org/10.1093/bioinformatics/btab007

Godbout, C., Fortin, J.A., 1985. Synthesized ectomycorrhizae of aspen: fungal genus level of structural characterization. Can. J. Bot. 63, 252–262. https://doi.org/10.1139/b85-029

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Research 42, D699–D704. https://doi.org/10.1093/nar/gkt1183

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hansen, K., Læssøe, T., Pfister, D.H., 2001. Phylogenetics of the *Pezizaceae*, with an emphasis on *Peziza*. Mycologia 93, 958–990. https://doi.org/10.1080/00275514.2001.12063229

Hansen, K., Pfister, D.H., 2006. Systematics of the *Pezizomycetes*—the operculate discomycetes. Mycologia 98, 1029–1040. https://doi.org/10.1080/15572536.2006.11832631

Harmaja, H., 1976a. Scanning electron microscopy of the spores of *Gyromitra* subg. *Gyromitra* and subg. *Discina* (*Pezizales*). Karstenia 16, 6–9.

Harmaja, H., 1976b. *Paradiscina Benedix* – a synonym of *Gyromitra* Fr. Karstenia 15, 33– 35.

Harmaja, H., 1969. A wider and more natural concept of the genus *Gyromitra* Fr. Karstenia 2, 9–12. https://doi.org/10.29203/ka.1969.53

Hawksworth, D.L., 2010. Funga and fungarium. IMA Fungus 1, 9–9. https://doi.org/10.1007/BF03449321

Healy, R.A., Arnold, A.E., Bonito, G., Huang, Y., Lemmond, B., Pfister, D.H., Smith, M.E., 2022. Endophytism and endolichenism in *Pezizomycetes*: the exception or the rule? New Phytologist 233, 1974–1983. https://doi.org/10.1111/nph.17886

Hensen, N., Bonometti, L., Westerberg, I., Brännström, I.O., Guillou, S., Cros-Aarteil, S., Calhoun, S., Haridas, S., Kuo, A., Mondo, S., Pangilinan, J., Riley, R., LaButti, K., Andreopoulos, B., Lipzen, A., Chen, C., Yan, M., Daum, C., Ng, V., Clum, A., Steindorff, A., Ohm, R.A., Martin, F., Silar, P., Natvig, D.O., Lalanne, C., Gautier, V., Ament-Velásquez, S.L., Kruys, Å., Hutchinson, M.I., Powell, A.J., Barry, K., Miller, A.N., Grigoriev, I.V., Debuchy, R., Gladieux, P., Hiltunen Thorén, M., Johannesson, H., 2023. Genome-scale phylogeny and comparative genomics of the fungal order *Sordariales*. Molecular Phylogenetics and Evolution 189, 107938. https://doi.org/10.1016/j.ympev.2023.107938
Hess, J., Skrede, I., Wolfe, B.E., LaButti, K., Ohm, R.A., Grigoriev, I.V., Pringle, A., 2014. Transposable element dynamics among asymbiotic and ectomycorrhizal *Amanita* fungi. Genome Biology and Evolution 6, 1564–1578. https://doi.org/10.1093/gbe/evu121

Hobbie, E.A., Weber, N.S., Trappe, J.M., 2001. Mycorrhizal vs saprotrophic status of fungi: the isotopic evidence. New Phytologist 150, 601–610. https://doi.org/10.1046/j.1469-8137.2001.00134.x

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47, D309– D314. https://doi.org/10.1093/nar/gky1085

James, T.Y., 2023. Sex without sexes: can the cost of finding a mate explain diversity in fungal mating systems? Integrative And Comparative Biology 63, 922–935. https://doi.org/10.1093/icb/icad037

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome- scale protein function classification. Bioinformatics 30, 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jurenitsch, J., Aurada, E., Kubelka, W., Robien, W., Viernstein, H., 1988. 1-(2- Hydroxyacetyl)-pyrazol, eine neue, mit gyromitrin verwandte verbindung aus *Gyromitra fastigiata*. Zeitschrift für Mykologie - Journal of the German Mycological Society 54, 155–157.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285

Kempton, P.E., Wells, V.L., 1973. Studies on the fleshy fungi of Alaska. VI. Notes on *Gyromitra*. Mycologia 65, 396–400. https://doi.org/10.1080/00275514.1973.12019448

Koch, A., Christan, J., Lohmeyer, T.R., 1996. *Gyromitra spinosospora* (Lucch. & Pelland.) comb. nov. (=*Discina spinosospora* Lucch. & Pelland.) - erstnachweis für Deutschland. Mycologia Bavarica 1, 46–49.

Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., Buscot, F., Canbäck, B., Choi, C., Cichocki, N., Clum, A., Colpaert, J., Copeland, A., Costa, M.D., Doré, J., Floudas, D., Gay, G., Girlanda, M., Henrissat, B., Herrmann, S., Hess, J., Högberg, N., Johansson, T., Khouja, H.-R., LaButti, K., Lahrmann, U., Levasseur, A., Lindquist, E.A., Lipzen, A., Marmeisse, R., Martino, E., Murat, C., Ngan, C.Y., Nehls, U., Plett, J.M., Pringle, A., Ohm, R.A., Perotto, S., Peter, M., Riley, R., Rineau, F., Ruytinx, J., Salamov, A., Shah, F., Sun, H., Tarkka, M., Tritt, A., Veneault-Fourrey, C., Zuccaro, A., Tunlid, A., Grigoriev, I.V., Hibbett, D.S., Martin, F., 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat Genet 47, 410–415. https://doi.org/10.1038/ng.3223

Kraisitudomsook, N., Healy, R.A., Pfister, D.H., Truong, C., Nouhra, E., Kuhar, F., Mujic, A.B., Trappe, J.M., Smith, M.E., 2020. Resurrecting the genus *Geomorium*: systematic study of fungi in the genera *Underwoodia* and *Gymnohydnotrya* (*Pezizales*) with the description of three new South American species. Persoonia 44, 98–112. https://doi.org/10.3767/persoonia.2020.44.04

Lagrange, E., de la Cruz, E., Esselin, F., Vernoux, J.-P., Pageot, N., Taieb, G., Camu, W., 2022. Reversible sub-acute motor neuron syndrome after mushroom intoxication masquerading as amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 23, 496–499. https://doi.org/10.1080/21678421.2021.2008453

Lagrange, E., Vernoux, J.P., Reis, J., Palmer, V., Camu, W., Spencer, P.S., 2021. An amyotrophic lateral sclerosis hot spot in the French Alps associated with genotoxic fungi. Journal of the Neurological Sciences 427, 117558. https://doi.org/10.1016/j.jns.2021.117558

Landvik, S., Egger, K., Schumacher, T., 1997. Towards a subordinal classification of the *Pezizales* (*Ascomycota*): phylogenetic analyses of SSU rDNA sequences. Nordic Journal of Botany 17, 403–418. https://doi.org/10.1111/j.1756-1051.1997.tb00337.x

Lanfranco, L., Bonfante, P., 2023. Lessons from arbuscular mycorrhizal fungal genomes. Current Opinion in Microbiology 75, 102357. https://doi.org/10.1016/j.mib.2023.102357

Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M., Henrissat, B., 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels 6, 41. https://doi.org/10.1186/1754-6834-6-41

Li, L., Wan, S.-P., Wang, Y., Thongklang, N., Tang, S.-M., Luo, Z.-L., Li, S.-H., 2023. New species of *Hydnotrya* (*Ascomycota*, *Pezizomycetes*) from southwestern China with notes on morphological characteristics of 17 species of *Hydnotrya*. MycoKeys 100, 49–67. https://doi.org/10.3897/mycokeys.100.106709

List, P.H., Luft, P., 1969. Nachweis und gehaltsbestimmung von gyromitrin in frischen lorcheln 19. Mitt. über pilzinhaltsstoffe. Arch Pharm Ber Dtsch Pharm Ges 302, 143– 146. https://doi.org/10.1002/ardp.19693020209

List, P.H., Luft, P., 1968. Gyromitrin, das gift der frühjahrslorchel. 16. Mitt. über pilzinhaltsstoffe. Arch Pharm Ber Dtsch Pharm Ges 301, 294–305. https://doi.org/10.1002/ardp.19683010410

Melie, T., Pirro, S., Miller, A.N., Smith, S.D., Schutz, K.S., Quandt, C.A., 2023. Comparative genomics and phylogenomic investigation of the class *Geoglossomycetes* provide insights into ecological specialization and the systematics of *Pezizomycotina*. Mycologia 115, 499–512. https://doi.org/10.1080/00275514.2023.2186743

Mendes, F.K., Vanderpool, D., Fulton, B., Hahn, M.W., 2021. CAFE 5 models variation in evolutionary rates among gene families. Bioinformatics 36, 5516–5518. https://doi.org/10.1093/bioinformatics/btaa1022

Methven, A.S., Zelski, S.E., Miller, A.N., 2013. A molecular phylogenetic assessment of the genus *Gyromitra* in North America. Mycologia 105, 1306–1314. https://doi.org/10.3852/12-397

Michelot, D., Toth, B., 1991. Poisoning by *Gyromitra esculenta*—a review. J of Applied Toxicology 11, 235–243. https://doi.org/10.1002/jat.2550110403

Miller, A.N., Dirks, A.C., Filippova, N., Popov, E., Methven, A.S., 2022. Studies in *Gyromitra* II: cryptic speciation in the *Gyromitra gigas* species complex; rediscovery of *G. ussuriensis* and *G. americanigigas* sp. nov. Mycol Progress 21, 1–13. https://doi.org/10.1007/s11557-022-01832-x

Miller, A.N., Matlak, A., Raudabaugh, D., Pärtel, K., Tamm, H., Methven, A.S., Voitk, A., 2015. The genus *Gyromitra* in NL. Omphalina 6, 4–15.

Miller, A.N., Yoon, A., Gulden, G., Stensholt, Ø., Van Vooren, N., Ohenoja, E., Methven, A.S., 2020. Studies in *Gyromitra* I: the *Gyromitra gigas* species complex. Mycol Progress 19, 1459–1473. https://doi.org/10.1007/s11557-020-01639-8

Minh, B.Q., Hahn, M.W., Lanfear, R., 2020a. New methods to calculate concordance factors for phylogenomic datasets. Molecular Biology and Evolution 37, 2727–2733. https://doi.org/10.1093/molbev/msaa106

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020b. IQ-TREE 2: new models and efficient methods for phylogenetic

inference in the genomic era. Mol Biol Evol 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mo, Y.K., Lanfear, R., Hahn, M.W., Minh, B.Q., 2023. Updated site concordance factors minimize effects of homoplasy and taxon sampling. Bioinformatics 39, 1–2. https://doi.org/10.1093/bioinformatics/btac741

Moravec, J., 1987. A taxonomic revision of the genus *Marcelleina*. Mycotaxon 30, 473–499.
Mundry, R., 2014. Statistical issues and assumptions of phylogenetic generalized least squares, in: Garamszegi, L.Z. (Ed.), Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology. Springer, Berlin, Germany, pp. 131–153. https://doi.org/10.1007/978-3-662-43550-2_6

Murat, C., Payen, T., Noel, B., Kuo, A., Morin, E., Chen, J., Kohler, A., Krizsán, K., Balestrini, R., Da Silva, C., Montanini, B., Hainaut, M., Levati, E., Barry, K.W., Belfiori, B., Cichocki, N., Clum, A., Dockter, R.B., Fauchery, L., Guy, J., Iotti, M., Le Tacon, F., Lindquist, E.A., Lipzen, A., Malagnac, F., Mello, A., Molinier, V., Miyauchi, S., Poulain, J., Riccioni, C., Rubini, A., Sitrit, Y., Splivallo, R., Traeger, S., Wang, M., Žifčáková, L., Wipf, D., Zambonelli, A., Paolocci, F., Nowrousian, M., Ottonello, S., Baldrian, P., Spatafora, J.W., Henrissat, B., Nagy, L.G., Aury, J.-M., Wincker, P., Grigoriev, I.V., Bonfante, P., Martin, F.M., 2018. *Pezizomycetes* genomes reveal the molecular basis of ectomycorrhizal truffle lifestyle. Nat Ecol Evol 2, 1956–1965. https://doi.org/10.1038/s41559-018-0710-4

Nelson, A., Vandegrift, R., Carroll, G.C., Roy, B.A., 2020. Double lives: transfer of fungal endophytes from leaves to woody substrates. PeerJ 8, e9341. https://doi.org/10.7717/peerj.9341

O'Donnell, K., Cigelnik, E., Weber, N.S., Trappe, J.M., 1997. Phylogenetic relationships among ascomycetous truffles and the true and false morels inferred from 18S and 28S ribosomal DNA sequence analysis. Mycologia 89, 48–65. https://doi.org/10.1080/00275514.1997.12026754

Patron, N.J., Waller, R.F., Cozijnsen, A.J., Straney, D.C., Gardiner, D.M., Nierman, W.C., Howlett, B.J., 2007. Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol Biol 7, 174. https://doi.org/10.1186/1471-2148-7-174

Pfister, D.H., 1980. On "*Peziza*" *melaleucoides*—a species of *Gyromitra* from the Western United States. Mycologia 72, 614–619.

Pinheiro, J., Bates, D., R Core Team, 2023. nlme: linear and nonlinear mixed effects models.

Porter, T.M., Martin, W., James, T.Y., Longcore, J.E., Gleason, F.H., Adler, P.H., Letcher, P.M., Vilgalys, R., 2011. Molecular phylogeny of the *Blastocladiomycota* (*Fungi*) based on nuclear ribosomal DNA. Fungal Biology 115, 381–392. https://doi.org/10.1016/j.funbio.2011.02.004

Promputtha, I., Lumyong, S., Dhanasekaran, V., McKenzie, E.H.C., Hyde, K.D., Jeewon, R., 2007. A phylogenetic evaluation of whether endophytes become saprotrophs at host senescence. Microb Ecol 53, 579–590. https://doi.org/10.1007/s00248-006- 9117-x

Revell, L.J., 2024. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ 12, e16505. https://doi.org/10.7717/peerj.16505

Reynolds, H.T., Slot, J.C., Divon, H.H., Lysøe, E., Proctor, R.H., Brown, D.W., 2017. Differential retention of gene functions in a secondary metabolite cluster. Molecular Biology and Evolution 34, 2002–2015. https://doi.org/10.1093/molbev/msx145

Rokas, A., Carroll, S.B., 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Molecular Biology and Evolution 22, 1337–1344. https://doi.org/10.1093/molbev/msi121

Rydholm, C., Dyer, P., Lutzoni, F., 2007. DNA sequence characterization and molecular evolution of MAT1 and MAT2 mating-type loci of the self-compatible ascomycete mold *Neosartorya fischeri*. Eukaryotic cell 6, 868–874. https://doi.org/10.1128/EC.00319-06

Salgado, J., Quandt, C.A., Stajich, J., 2023. Phylogenetics and phylogenomics to understand fungal diversity. https://doi.org/10.22541/au.167845826.61933731/v1

Sánchez-Vallet, A., Fouché, S., Fudal, I., Hartmann, F.E., Soyer, J.L., Tellier, A., Croll, D., 2018. The genome biology of effector gene evolution in filamentous plant pathogens. Annual Review of Phytopathology 56, 21–40. https://doi.org/10.1146/annurev-phyto-080516-035303

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B.W., Pruitt, K.D., Sherry, S.T., 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50, D20–D26. https://doi.org/10.1093/nar/gkab1112

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single- copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sow, A., Van Wyk, J., Lemmond, B., Healy, R., Smith, M.E., Bonito, G., 2024. Effective field collection of *Pezizales* ascospores for procuring diverse fungal isolates. Diversity 16, 1–14. https://doi.org/10.3390/d16030165

Steenwyk, J.L., Li, Y., Zhou, X., Shen, X.-X., Rokas, A., 2023. Incongruence in the phylogenomics era. Nat Rev Genet 24, 834–850. https://doi.org/10.1038/s41576- 023-00620-x

Steenwyk, J.L., Shen, X.-X., Lind, A.L., Goldman, G.H., Rokas, A., 2019. A robust phylogenomic time tree for biotechnologically and medically important fungi in the

genera *Aspergillus* and *Penicillium*. mBio 10, e00925-19. https://doi.org/10.1128/mbio.00925-19

Steinegger, M., Söding, J., 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol 35, 1026–1028. https://doi.org/10.1038/nbt.3988

Sun, W., Liu, W., Cai, Y., Shi, X., Wu, L., Zhang, J., Er, L., Huang, Q., Yin, Q., Zhao, Z., He, P., Yu, F., 2023. Structure of the mating-type genes and mating systems of *Verpa bohemica* and *Verpa conica* (*Ascomycota*, *Pezizomycotina*). Journal of Fungi 9, 1– 16. https://doi.org/10.3390/jof9121202

Szink, I., Davis, E.L., Ricks, K.D., Koide, R.T., 2016. New evidence for broad trophic status of leaf endophytic fungi of *Quercus gambelii*. Fungal Ecology 22, 2–9. https://doi.org/10.1016/j.funeco.2016.04.003

Taylor, J.W., Berbee, M.L., 2006. Dating divergences in the Fungal Tree of Life: review and new analyses. Mycologia 98, 838–849. https://doi.org/10.1080/15572536.2006.11832614

Tedersoo, L., May, T.W., Smith, M.E., 2010. Ectomycorrhizal lifestyle in fungi: global diversity, distribution, and evolution of phylogenetic lineages. Mycorrhiza 20, 217– 263. https://doi.org/10.1007/s00572-009-0274-x

The UniProt Consortium, 2023. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research 51, D523–D531. https://doi.org/10.1093/nar/gkac1052

To, T.-H., Jung, M., Lycett, S., Gascuel, O., 2016. Fast dating using least-squares criteria and algorithms. Systematic Biology 65, 82–97. https://doi.org/10.1093/sysbio/syv068

Tremble, K., Henkel, T., Bradshaw, A., Domnauer, C., Brown, L.M., Thám, L.X., Furci, G., Aime, M.C., Moncalvo, J.-M., Dentinger, B., 2024. A revised phylogeny of *Boletaceae* using whole genome sequences. Mycologia 0, 1–17. https://doi.org/10.1080/00275514.2024.2314963

Vandepol, N., Liber, J., Desirò, A., Na, H., Kennedy, M., Barry, K., Grigoriev, I.V., Miller, A.N., O'Donnell, K., Stajich, J.E., Bonito, G., 2020. Resolving the *Mortierellaceae* phylogeny through synthesis of multi-gene phylogenetics and phylogenomics. Fungal Diversity 104, 267–289. https://doi.org/10.1007/s13225-020-00455-5

Vellinga, E.C., Kuyper, T.W., Ammirati, J., Desjardin, D.E., Halling, R.E., Justo, A., Læssøe, T., Lebel, T., Lodge, D.J., Matheny, P.B., Methven, A.S., Moreau, P.-A., Mueller, G.M., Noordeloos, M.E., Nuytinck, J., Ovrebo, C.L., Verbeken, A., 2015. Six simple guidelines for introducing new genera of fungi. IMA Fungus 6, A65–A68. https://doi.org/10.1007/BF03449356

Wang, X.-C., Yang, Z.-L., Chen, S.-L., Bau, T., Li, T.-H., Li, L., Fan, L., Zhuang, W.-Y., 2023. Phylogeny and taxonomic revision of the family *Discinaceae* (*Pezizales*, *Ascomycota*). Microbiology Spectrum 11, e00207-23. https://doi.org/10.1128/spectrum.00207-23

Wang, X.-C., Zhuang, W.-Y., 2019. A three-locus phylogeny of *Gyromitra* (*Discinaceae*, *Pezizales*) and discovery of two cryptic species. Mycologia 111, 69–77. https://doi.org/10.1080/00275514.2018.1515456

Weber, N.S., 1995. A Morel Hunter's Companion: A Guide to True and False Morels. Thunder Bay Press, Lansing, Michigan.

Wells, V.L., Kempton, P.E., 1968. Studies on the fleshy fungi of Alaska. II. Mycologia 60, 888–901. https://doi.org/10.1080/00275514.1968.12018645

Yoshida, K., Burbano, H.A., Krause, J., Thines, M., Weigel, D., Kamoun, S., 2014. Mining herbaria for plant pathogen genomes: back to the future. PLOS Pathogens 10, e1004028. https://doi.org/10.1371/journal.ppat.1004028

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28–36. https://doi.org/10.1111/2041-210X.12628

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19, 15– 32. https://doi.org/10.1186/s12859-018-2129-y

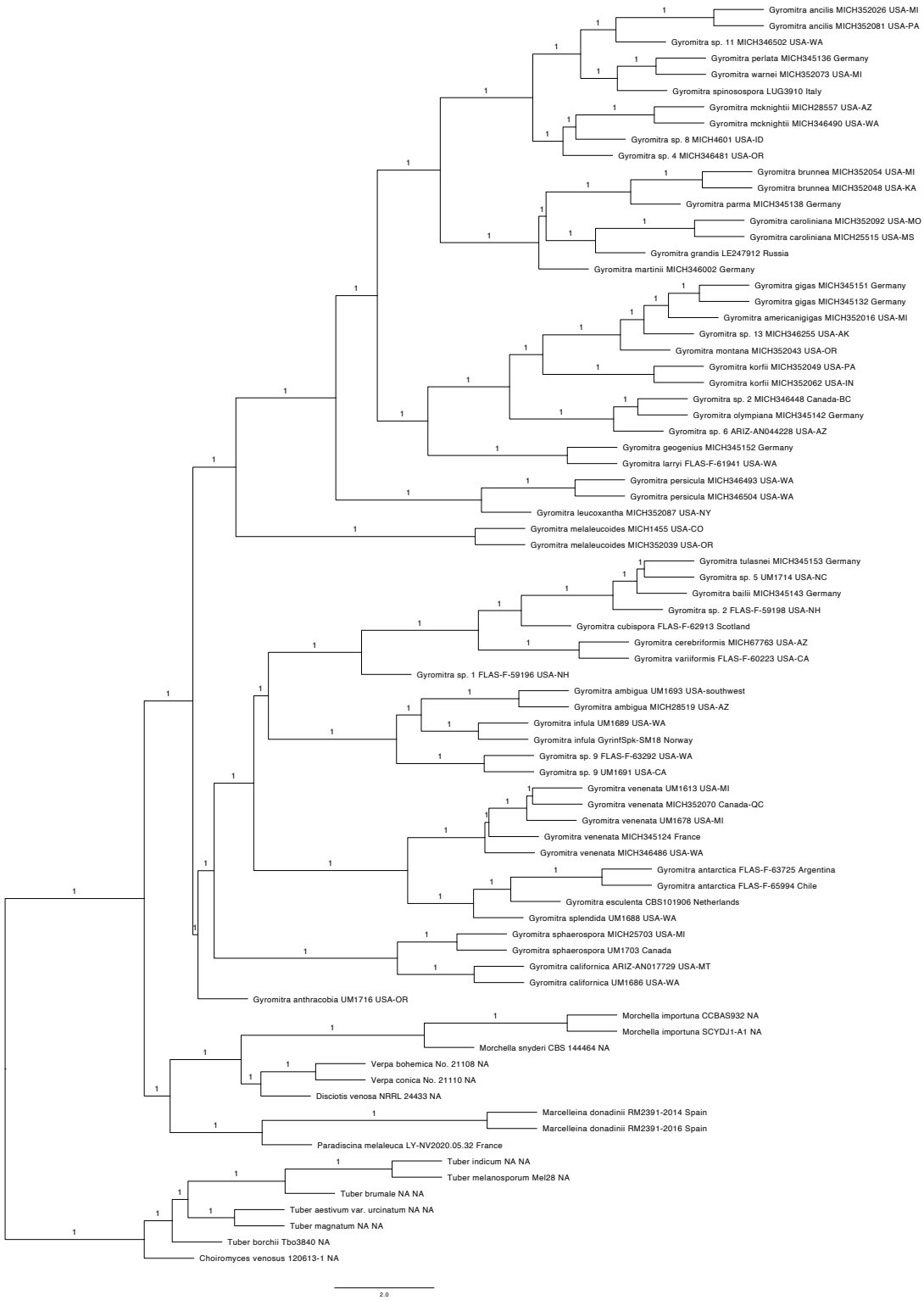
Zhang, N., Luo, J., Bhattacharya, D., 2017. Advances in fungal phylogenomics and their impact on fungal systematics, in: Townsend, J.P., Wang, Z. (Eds.), Advances in Genetics, Fungal Phylogenetics and Phylogenomics. Academic Press, Cambridge, Massachusetts, pp. 309–328. https://doi.org/10.1016/bs.adgen.2017.09.004

Zhao, H., Dai, Y.-C., Wu, F., Liu, X.-Y., Maurice, S., Krutovsky, K.V., Pavlov, I.N., Lindner, D.L., Martin, F.M., Yuan, Y., 2023. Insights into the ecological diversification of the *Hymenochaetales* based on comparative genomics and phylogenomics with an emphasis on *Coltricia*. Genome Biology and Evolution 15, evad136. https://doi.org/10.1093/gbe/evad136

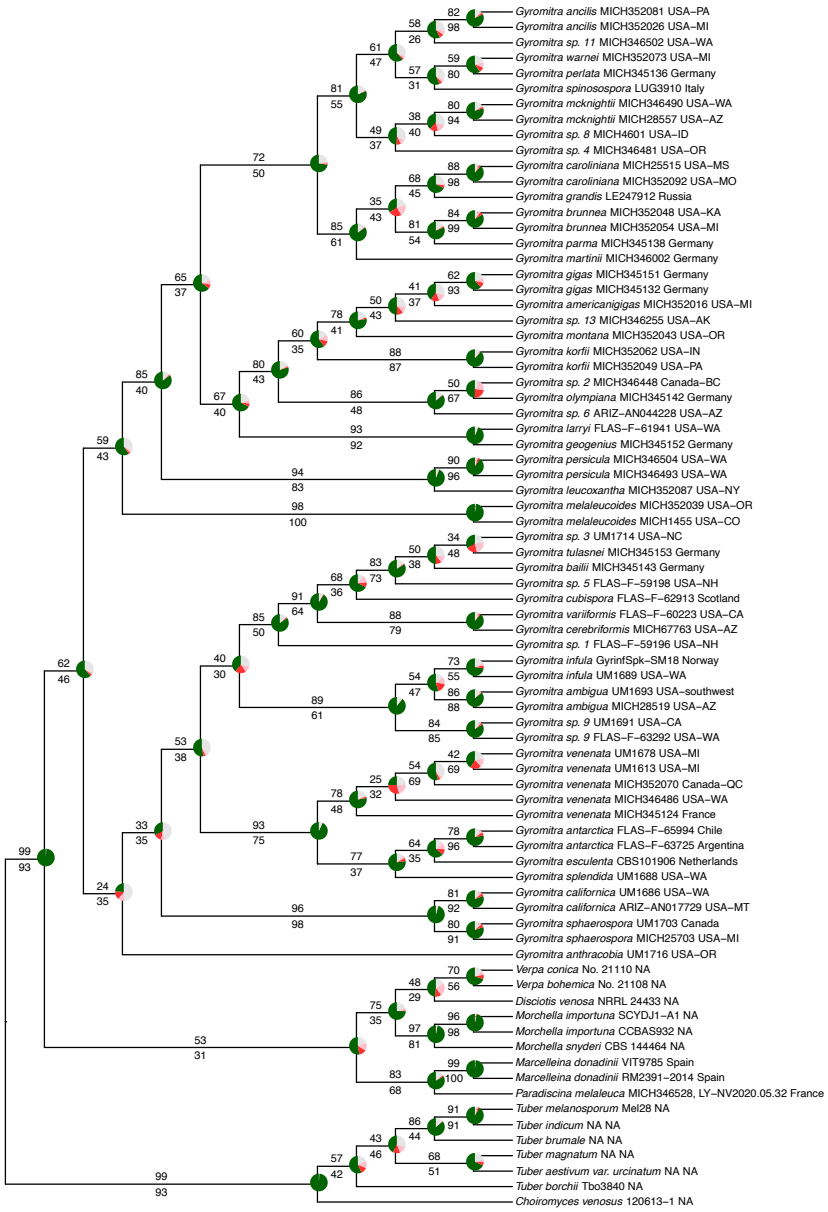
Zhao, R.-L., Zhou, J.-L., Chen, J., Margaritescu, S., Sánchez-Ramírez, S., Hyde, K.D., Callac, P., Parra, L.A., Li, G.-J., Moncalvo, J.-M., 2016. Towards standardizing taxonomic ranks using divergence times – a case study for reconstruction of the *Agaricus* taxonomic system. Fungal Diversity 78, 239–292. https://doi.org/10.1007/s13225- 016-0357-x

Zuev, A.G., Kudrin, A.A., Leonov, V.D., Saraeva, A.K., Tsurikov, S.M., Tiunov, A.V., 2023. Good shelter, bad food: invertebrates inhabiting fruiting bodies of spring ascomycetes in a mixed forest. Russ J Ecol 54, 307–310. https://doi.org/10.1134/S1067413623040136

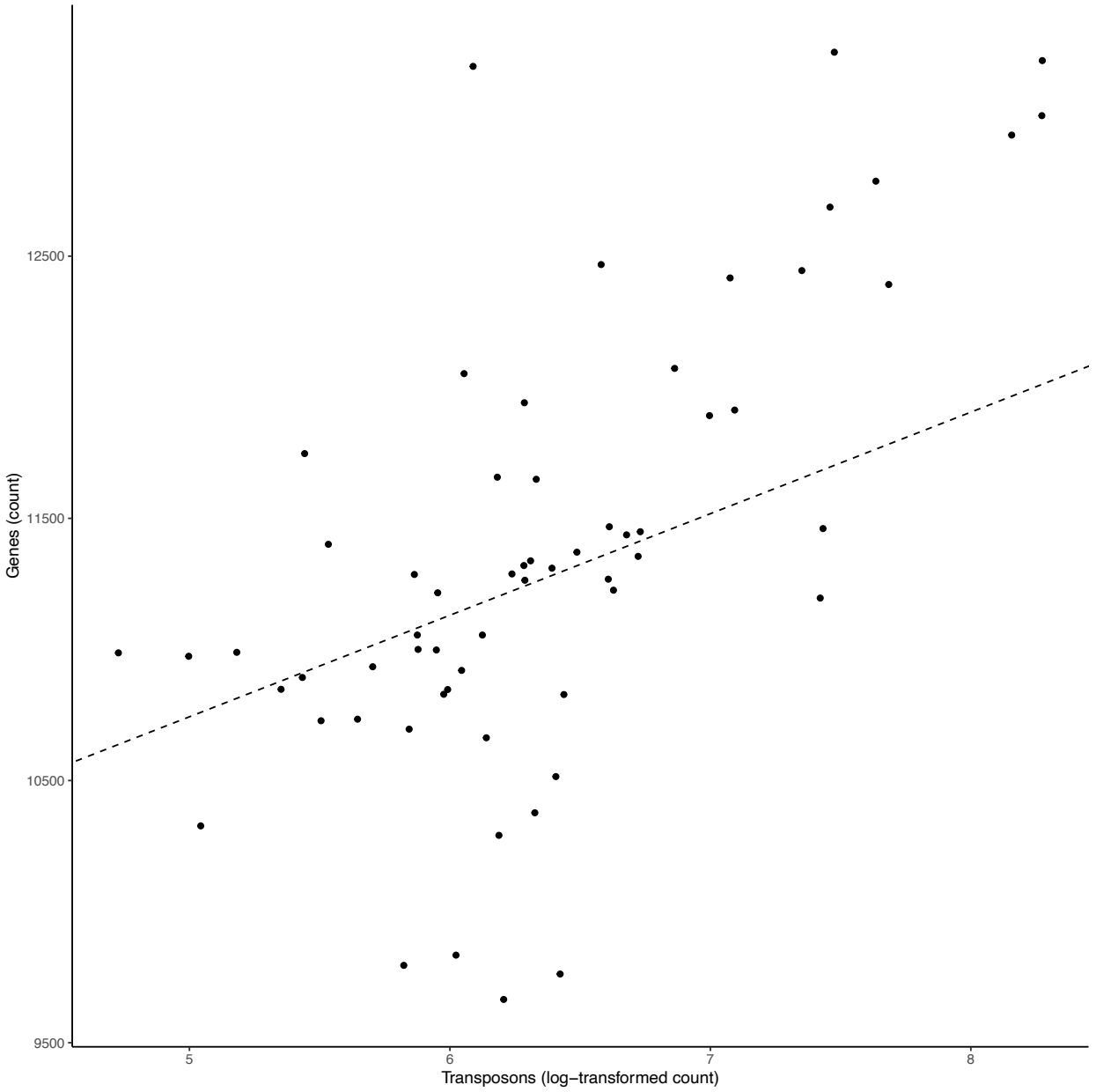
Supplement



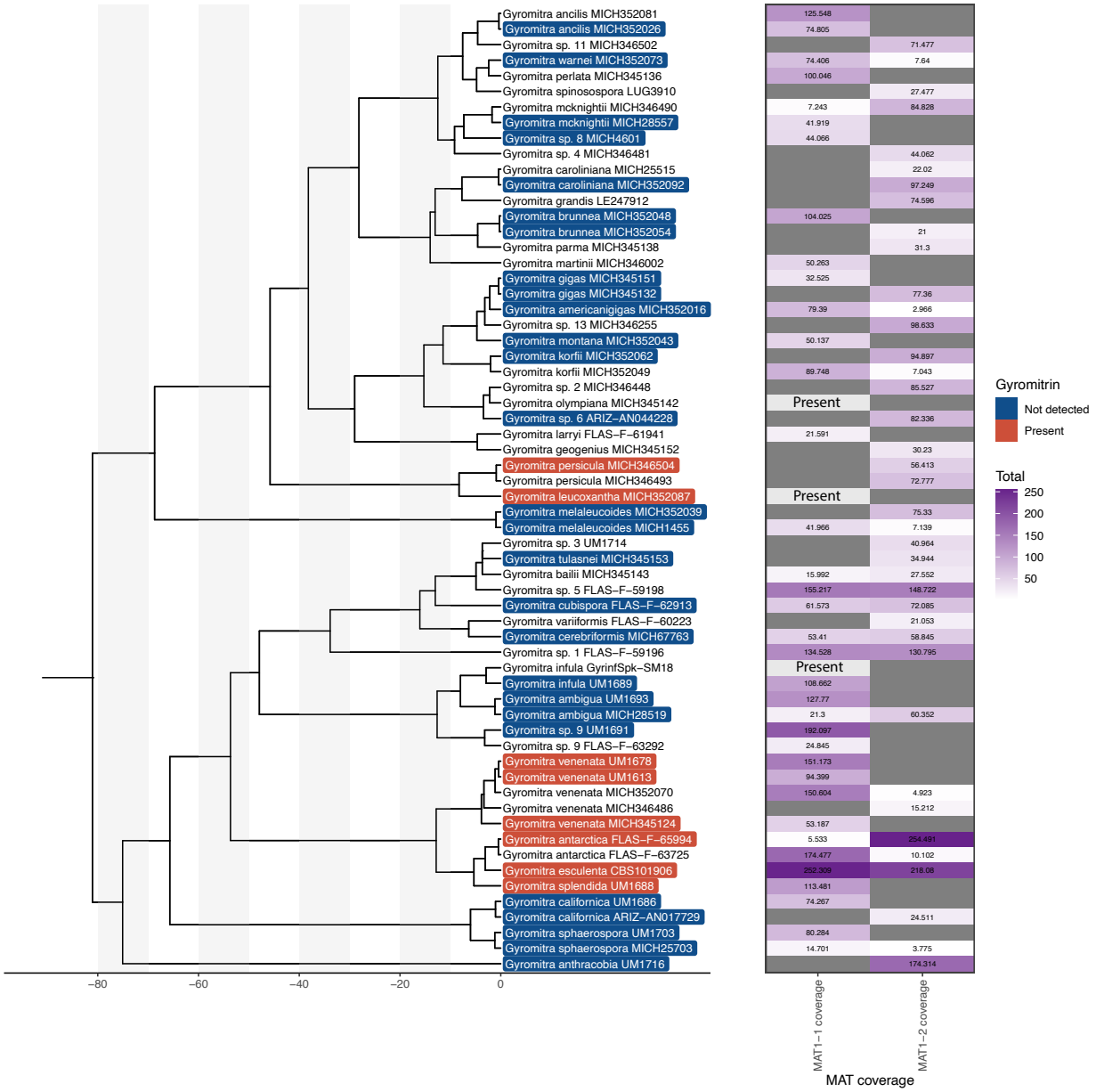
Supplemental Figure 3.1. Coalescent phylogenomic tree of *Discinaceae* and related *Pezizales* families. A total of 1645 individual gene trees were inferred with IQ-TREE2 and a coalescent phylogenomic tree was created with ASTRAL-III. Numbers on branches are quartet probabilities.



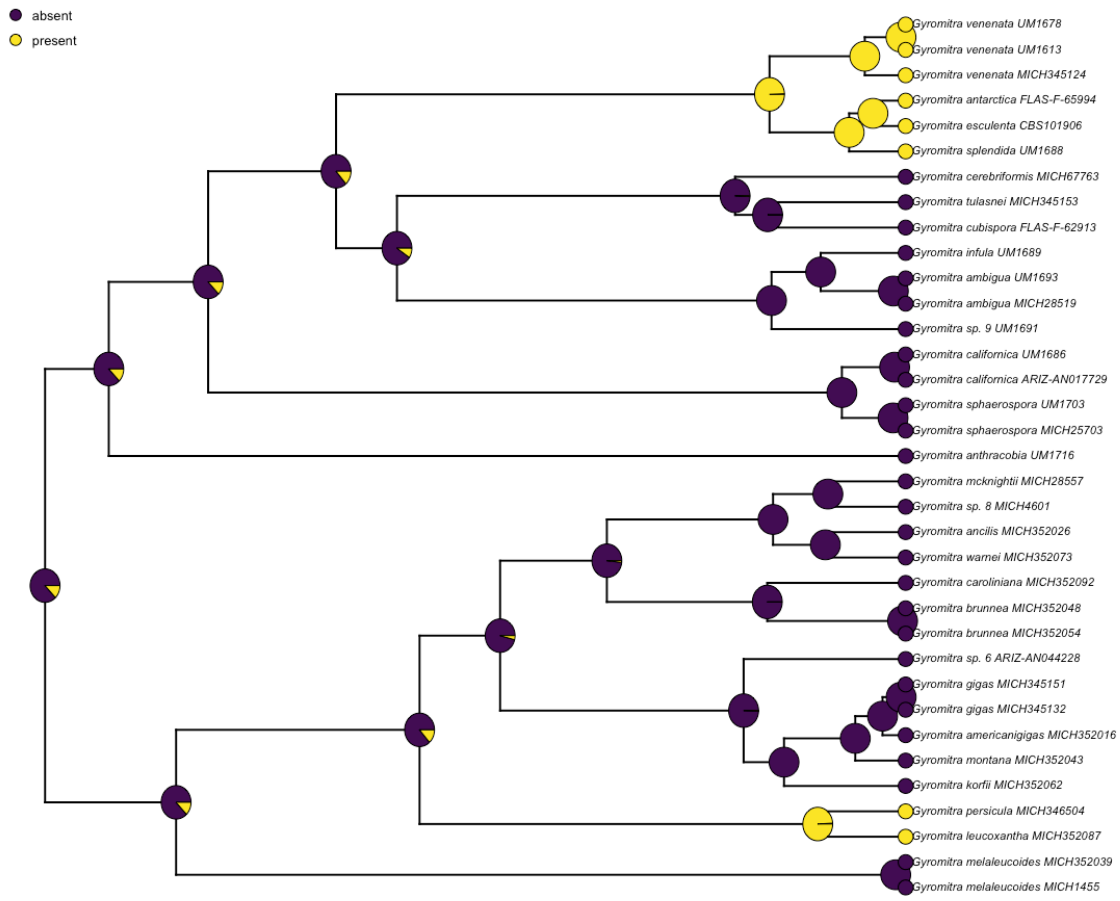
Supplemental Figure 3.2. Cladogram of *Discinaceae* and related *Pezizales* families with concordance pie charts. Gene concordance and site concordance values are listed above and below the branches, respectively. For each branch, the pie charts show the proportion of decisive gene trees that supported the species tree (dark green), the most likely alternate arrangement of the quarter (red), and the second most likely alternate arrangement of the quarter (light red); grey pie wedges indicate the proportion of decisive gene trees that were polyphyletic for a branch, meaning one or more of the groupings of the species tree quartet at that branch were not recovered as monophyletic.



Supplemental Figure 3.3. Scatterplot of log-transformed number of transposable elements and total gene number. The dotted line is the best fit phylogenetic generalized least squares regression.



Supplemental Figure 3.4. Heatmap of *MAT* idiomorph coverage. *Discinaceae* phylogram tip labels are colored according to gyromitrin status. *MAT* idiomorph coverage was calculated by mapping raw reads to the *MAT* gene sequence with BBTools.



Supplemental Figure 3.5. Ancestral state reconstruction of gyromitrin production in *Discinaceae*. Circles at the tips show observed gyromitrin production (yellow means present, purple means absent). Ancestral nodes show posterior probabilities of gyromitrin production from stochastic character mapping with phytools 2.1-1.

Supplemental Table 3.1. Metadata for samples whose genomes were sequenced.

Species	Accession	Location	Date	Genome Source
Gyromitra ancilis	MICH352026	USA-MI	2020	this study
Gyromitra venenata	UM1613	USA-MI	culture	this study
Gyromitra melaleucoides	MICH352039	USA-OR	2020	this study
Gyromitra montana	MICH352043	USA-OR	2020	this study
Gyromitra brunnea	MICH352054	USA-MI	2020	this study
Gyromitra korfii	MICH352062	USA-IN	2021	this study
Gyromitra splendida	MICH25554	USA-MI	1984	this study
Gyromitra sphaerospora	UM1703	Canada	culture	this study
Gyromitra californica	UM1686	USA-WA	culture	this study
Gyromitra splendida	UM1688	USA-WA	culture	this study
Gyromitra infula	UM1689	USA-WA	culture	this study
Gyromitra sp. 9	UM1691	USA-CA	culture	this study
Gyromitra americanigigas	MICH352016	USA-MI	2020	this study
Gyromitra leucoxantha	MICH352087	USA-NY	2020	this study
Gyromitra caroliniana	MICH352092	USA-MO	2021	this study
Gyromitra cubispora	FLAS-F-62913	Scotland	2018	this study
Gyromitra antarctica	FLAS-F-65994	Chile	2019	this study
Gyromitra sphaerospora	MICH25703	USA-MI	1970	this study
Gyromitra ambigua	MICH28519	USA-AZ	1990	this study
Gyromitra mcknightii	MICH28557	USA-AZ	1983	this study
Gyromitra sp. 6	MICH4614	USA-ID	1972	this study
Gyromitra cerebriformis	MICH67763	USA-AZ	1996	this study
Gyromitra venenata	UM1678	USA-MI	culture	this study
Gyromitra korfii	MICH352049	USA-PA	2020	this study
Gyromitra venenata	MICH352070	Canada-QC	2021	this study
Gyromitra warnei	MICH352073	USA-MI	2020	this study
Gyromitra ancilis	MICH352081	USA-PA	2021	this study
Gyromitra sp. 2	MICH346448	Canada-BC	2022	this study
Gyromitra sp. 4	MICH346481	USA-OR	2022	this study
Gyromitra sp. 6	MICH346452	USA-ID	2011	this study
Gyromitra sp. 11	MICH346502	USA-WA	2011	this study
Gyromitra persicula	MICH346504	USA-WA	2022	this study
Gyromitra venenata	MICH346486	USA-WA	2022	this study
Gyromitra anthracobia	UM1716	USA-OR	culture	this study
Gyromitra ambigua	UM1693	USA	culture	this study
Gyromitra sp. 1	FLAS-F-59196	USA-NH	2015	this study
Gyromitra sp. 5	FLAS-F-59198	USA-NH	2014	this study
Gyromitra variiformis	FLAS-F-60223	USA-CA	2016	this study
Gyromitra larryi	FLAS-F-61941	USA-WA	2017	this study
Gyromitra antarctica	FLAS-F-63725	Argentina	2015	this study
Gyromitra sp. 8	MICH4601	USA-ID	1962	this study
Gyromitra melaleucoides	MICH1455	USA-CO	1983	this study
Gyromitra leucoxantha	MICH25407	USA-MI	1984	this study
Gyromitra caroliniana	MICH25515	USA-MS	1984	this study
Gyromitra brunnea	MICH352048	USA-KA	2020	this study
Gyromitra korfii	CUP-K-4801	USA-NY	1953	this study
Gyromitra spinosospora	LUG3910	Italy	1984	this study
Gyromitra sichuanensis	CUP-CH002522	China	1981	this study
Gyromitra grandis	LE247912	Russia	2011	this study
Gyromitra grandis	TNS-F-80830	Japan	2017	this study
Gyromitra olympiana	MICH345142	Germany	2012	this study

Gyromitra geogenia	ZT-Myc0067647	Switzerland	1964	this study
Gyromitra mcknightii	MICH346490	USA-WA	2022	this study
Gyromitra sp. 3	UM1714	USA-NC	culture	this study
Gyromitra venenata	MICH345124	France	2022	this study
Gyromitra gigas	MICH345132	Germany	2004	this study
Gyromitra parma	MICH345134	Germany	2004	this study
Gyromitra perlata	MICH345136	Germany	2007	this study
Gyromitra parma	MICH345138	Germany	2014	this study
Gyromitra martinii	MICH346002	Germany	2014	this study
Gyromitra bailii	MICH345143	Germany	2011	this study
Gyromitra gigas	MICH345151	Germany	2015	this study
Gyromitra geogenius	MICH345152	Germany	2019	this study
Gyromitra tulasnei	MICH345153	Germany	2019	this study
Gyromitra splendida	MICH346243	USA-AK	1960	this study
Gyromitra californica	ARIZ-AN017729	USA-MT	1977	this study
Paradiscina melaleuca	MICH346528	France	2020	this study
Gyromitra persicula	MICH346493	USA-WA	2011	this study
Gyromitra sp. 6	ARIZ-AN044228	USA-AZ	2017	this study
Gyromitra sp. 9	FLAS-F-63292	USA-WA	2018	this study
Gyromitra sp. 13	MICH346255	USA-AK	1960	this study
Marcelleina donadinii	NA	Spain	2014	this study
Marcelleina donadinii	VIT9785	Spain	2016	this study
Gyromitra esculenta	CBS101906	Netherlands	culture	JGI
Gyromitra infula	GyrinfSpk-SM18	Norway	culture	JGI
Choiromyces venosus	120613-1	NA	NA	Murat et al. 2018
Disciotis venosa	NRRL 24433	NA	NA	NA
Morchella importuna	CCBAS932	NA	NA	Murat et al. 2018
Morchella importuna	SCYDJ1-A1	NA	NA	Tan et al. 2019
Morchella snyderi	CBS 144464	NA	NA	Steindorff et al. 2022
Tuber aestivum	NA	NA	NA	Murat et al. 2018
Tuber borchii	Tbo3840	NA	NA	Murat et al. 2018
Tuber brumale	NA	NA	NA	Morin et al. 2021
Tuber indicum	NA	NA	NA	Morin et al. 2021
Tuber magnatum	NA	NA	NA	Murat et al. 2018
Tuber melanosporum	Mel28	NA	NA	Martin et al. 2010
Verpa bohemica	No. 21108	NA	NA	Sun et al. 2023
Verpa conica	No. 21110	NA	NA	Sun et al. 2023

JGI genome sequencing and assembly

The *Gyromitra esculenta* CBS101906 genome was sequenced using fragment and long mate-pair (LMP) Illumina libraries. Fragment 300 bp insert site library was produced from 100 ng of genomic DNA, sheared to 300 bp using the Covaris LE220-Plus and size selected with SPRI using TotalPure NGS beads (Omega Bio-tek). The fragments were treated with end-repair, A-tailing, and ligation of Illumina compatible adapters (IDT, Inc) using the KAPA- HyperPrep kit (KAPA Biosystems). For 4kmp LMP libraries, 5 µg of DNA was sheared using the Covaris g-TUBE (Covaris) and gel size selected for 4 kb. The sheared DNA was treated with end repair and ligated with biotinylated adapters containing loxP. The adapter ligated DNA fragments were circularized via recombination by a Cre excision reaction (NEB). The circularized DNA templates were then randomly sheared using the Covaris LE220 (Covaris). The sheared fragments were treated with end repair and A-tailing using the KAPA- Illumina library creation kit (KAPA biosystems) followed by immobilization of mate pair fragments on strepavidin beads (Invitrogen). Illummina compatible adapters (IDT, Inc) were ligated to the mate pair fragments and 10 cycles of PCR was used to enrich for the final library (KAPA Biosystems). The prepared library was quantified using KAPA Biosystem's next-generation sequencing library qPCR kit (Roche) and run on a Roche LightCycler 480 real-time PCR instrument. The quantified library was then multiplexed with other libraries, and the pool of libraries was then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3 or v4, and Illumina's cBot instrument to generate a clustered flow cell for sequencing. Sequencing of the flow cell was performed on the Illumina HiSeq 2000 or 2500 sequencer using HiSeq TruSeq SBS sequencing kits, v3 or v4, following a 2x100 bp (LPM) or 2x150 bp (fragments) indexed run recipe. Each fastq file was QC filtered for artifact/process contamination

and subsequently assembled together with AllPathsLG version R49403, (Gnerre et al., 2010), to produce a 83.6 X coverage main assembly with 271 scagolds. Mitochondria was assembled separately with AllPathsLG version R47710.

The *Gyromitra infula* GyrinfSpk-SM18 genome was sequenced using PacBio sequencing. For this, 5 ug of genomic DNA was sheared to > 10 kb using Covaris g-Tubes. The sheared DNA was treated with exonuclease to remove single-stranded ends and DNA damage repair mix followed by end repair and ligation of blunt adapters using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences). The library was purified with AMPure PB beads. PacBio Sequencing primer was then annealed to the SMRTbell template library and sequencing polymerase was bound to them using Sequel Binding kit 3.0. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosystem's Sequel sequencer using v3 sequencing primer, 1M v3 SMRT cells, and Version 3.0 sequencing chemistry with 1x360 min sequencing movie run times. Filtered subread data was processed with the JGI QC pipeline. (genome.jgi.doe.gov/lookup?keyName=jgiProjectId&keyValue=1240576) to remove artifacts and assembled together with Falcon version pb-assembly=0.0.2|falcon- kit=1.2.3|pypeflow=2.1.0 (github.com/PacificBiosciences/FALCON) to generate an initial assembly. The mitochondrial genome was assembled separately from the Falcon pre-assembled reads (preads) using an in-house tool (assemblemito.py), used to filter the preads, and polished with Arrow version SMRTLink v7.0.1.66975 (github.com/PacificBiosciences/GenomicConsensus). A secondary Falcon assembly was generated using the mitochondria-filtered preads, improved with finisherSC version 2.1 (Lam et al.,2014), and polished with Arrow version SMRTLINK v7.0.1.66975 (github.com/PacificBiosciences/GenomicConsensus). Main genome contamination of possible mitochondrial origin was identified using BLAST vs. the NCBI nt database, and

analysis of GC content, read coverage, and tetranucleotide profiles. These contaminant scagolds were removed from the main genome and used, along with the initial mitochondrial genome assembly, to recruit preads using bbtools version 38.73 [bbduk.sh k=31 mm=f mkf=0.05] (sourceforge.net/projects/bbmap/). Matching preads were then assembled with Flye version 2.5 [--pacbio-corr -g 200k --asm-coverage 50], and the resulting mitochondrial scagold was polished with Arrow version SMRTLINK v7.0.1.66975. Contigs less than 1000 bp were excluded.

JGI transcriptome sequencing and assembly

For *Gyromitra esculenta* CBS101906 transcriptome, stranded cDNA libraries were generated using the Illumina Truseq Stranded, RNA LT kit. mRNA was purified from 1 ug of total RNA using magnetic beads containing poly-T oligos. mRNA was fragmented and reversed transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 8 cycles of PCR.

For *Gyromitra infula* GyrinfSpk-SM18 transcriptome, plate-based RNA sample prep was performed on the PerkinElmer Sciclone NGS robotic liquid handling system using Illumina's TruSeq Stranded mRNA HT sample prep kit utilizing poly-A selection of mRNA following the protocol outlined by Illumina in their user guide: support.illumina.com/sequencing/sequencing_kits/truseq-stranded-mrna.html, and with the following conditions: total RNA starting material was 1 ug per sample and 8 cycles of PCR was used for library amplification.

Both prepared libraries were quantified using KAPA Biosystem’s next-generation sequencing library qPCR kit (Roche) and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then multiplexed with other libraries, and the pool of

libraries was sequenced on either the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v4, and Illumina's cBot instrument (*Gyromitra esculenta* CBS101906) or the Illumina NovaSeq 6000 sequencing platform using NovaSeq XP v1 reagent kits, S4 flow cell (*Gyromitra infula* GyrinfSpk-SM18), both following a 2x150 bp indexed run recipe.

Raw fastq file reads were filtered and trimmed using the JGI QC pipeline resulting in the filtered fastq file (*.filter-RNA.fastq.gz files). Using BBDuk (https://sourceforge.net/projects/bbmap/), raw reads were evaluated for artifact sequence by kmer matching (kmer=25), allowing 1 mismatch and detected artifact was trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads and reads containing any Ns were removed. Quality trimming was performed using the phred trimming method set at Q6. Finally, following trimming, reads under the length threshold were removed (minimum length 25 bases or 1/3 of the original read length - whichever was longer). Filtered fastq files were used as input for de novo assembly of RNA contigs. Reads were assembled into consensus sequences using Trinity ver. 2.3.2 [Graber et al., 2011].

JGI annotation

Both genomes were annotated using the JGI Annotation pipeline (Grigoriev et al., NAR 2014).

Chapter 4 Bioinformatics Identification of a Putative Gyromitrin Biosynthesis Gene Cluster in Lorchel Mushrooms (*Gyromitra* spp., *Pezizales*)

Abstract

Gyromitrin is a deadly mycotoxin produced by ascomycete mushrooms in the genus *Gyromitra* (*Discinaceae*, *Pezizomycetes*), also known as false morels or lorchels. It belongs to a rare and interesting class of secondary metabolites defined by the presence of a nitrogen-nitrogen (N–N) bond. Great strides have been made in understanding N–N bond formation in bacteria, but nothing is known of the biosynthetic machinery responsible for its formation in fungi. Gyromitrin has a discontinuous distribution in the lorchel family, *Discinaceae*, suggestive of horizontal gene transfer. We hypothesized that the gyromitrin biosynthesis genes would reflect this distribution in the form of phylogenetic incongruence among the gene trees, prompting a bioinformatics search for putative gyromitrin biosynthesis gene clusters (BGCs). Using a recently generated, robust dataset of *Gyromitra* genomes, we searched for homologs of known bacterial N–N bond-forming enzymes and patterns of phylogenetic incongruence whereby the gyromitrin-producing taxa were more closely related than expected. This approach yielded three gyromitrin BGC candidates. Two of these BGCs likely encode other secondary metabolites, namely an iron-scavenging siderophore and an uncharacterized cyclic peptide mycotoxin, although their association with gyromitrin production cannot be ruled out. The third, a noncanonical BGC devoid of conventional backbone biosynthesis genes, is our leading hypothesis for gyromitrin production. It consists of four to five colocalized genes whose inferred phylogenies show the gyromitrin-producing species forming a monophyletic clade. These

phylogenies suggest a horizontal gene transfer event occurred between an ancestor of subg. *Gyromitra* to an ancestor of subg. *Leucoxantha.* The putative gyromitrin BGC is composed of two iron-dependent oxygenases, a peptidase, an aldolase, and a didomain protein with fused aminoacyl-tRNA synthetase and acetyltransferase domains. Several components of this cluster are analogous to enzymes involved in bacterial N–N bond biosynthesis. Our candidate gyromitrin BGCs are exceptional targets for future validation studies involving labelled isotope analyses, heterologous expression, and gene knockouts. Elucidation of the gyromitrin BGC promises to augment genome mining and identification of new N–N bond-containing secondary metabolites in fungi.

4.1 Introduction

Gyromitrin (acetaldehyde *N*-methyl-*N*-formylhydrazone) is a deadly mycotoxin produced by lorchel mushrooms (*Gyromitra* spp.) in the family *Discinaceae* (*Pezizomycetes*, *Ascomycota*) (Figure 4.1). It is an unstable, heat- and acid-labile molecule that breaks down into monomethylhydrazine, which is responsible for its deleterious health impacts (Dirks et al., 2023; Lagrange et al., 2021; Michelot and Toth, 1991). Ironically, in Scandinavia, lorchels are enjoyed as a delicacy after first being detoxified via a laborious boiling procedure (Sitta et al., 2021; Svanberg and Lindh, 2019). They are also sometimes mistaken for prized morel mushrooms (*Morchella* spp.) and inadvertently consumed without proper detoxification, resulting in poisoning (Leathem and Dorran, 2007). List and Luft (1968, 1969) first isolated and characterized the chemical structure of gyromitrin, and it was later found to coexist with eight higher homologs present in smaller quantities (Figure 4.1) (List and Luft, 1969, 1968; Pyysalo, 1975; Pyysalo and Niskanen, 1977). A survey of gyromitrin in *Discinaceae* revealed a limited and phylogenetically discontinuous distribution (Dirks et al., 2023). Gyromitrin was detected in

all tested samples from subgenus *Gyromitra* and subgenus *Leucoxantha* but not in any other subgenera (Figure 4.2). This distribution is suggestive of horizontal gene transfer but may be the result of extensive gene loss or could simply be a mirage caused by transcriptionally silent gene clusters in negative samples.

As a hydrazone, gyromitrin belongs to a rare class of secondary metabolites called nitrogen-nitrogen (N–N) bond-containing natural products (N2NPs). Only about 300 N2NPs have been discovered—less than 0.1% of the total number of documented secondary metabolites (Blair and Sperry, 2013; Cano-Prieto et al., 2024; He et al., 2022; Le Goff and Ouazzani, 2014). N2NPs possess remarkable structural diversity that imparts interesting bioactive properties, making them attractive to chemists and valuable targets for bioprospecting. Great strides have been made over the past decade in understanding the biosynthetic machinery that catalyzes the N–N bond in bacteria. Two principal strategies are the generation of nitrite and the conversion of amines to hydroxylamines, often with the aid of metal-dependent proteins (Chen et al., 2021; He et al., 2022; Katsuyama and Matsuda, 2020). Studying fungal N2NP biosynthesis could reveal alternative biosynthetic pathways for N–N bond formation and yield additional insights into the evolution of N2NP biochemistry.

While they have received less attention than bacteria, fungi are also known to produce a wide array of N2NPs that function as antibiotics, mycotoxins, and pigments (He et al., 2022; Le Goff and Ouazzani, 2014). Within *Discinaceae*, an additional N2NP has been documented: 1-(2-hydroxyacetyl)-pyrazole, produced by *Gyromitra grandis* (Jurenitsch et al., 1988). Beyond the lorchels, several N2NPs have been identified from the model molds *Aspergillus* spp. and *Penicillium* spp. (*Aspergillaceae*, *Eurotiomycetes*, *Ascomycota*), including farylhydrazone, NG-061, and penoxalin, among others (Bhandari et al., 1999; Ding et al., 2020; Khalil et al., 2022;

Minato, 1979; Tang et al., 2023; Zhang et al., 2016). The lichens *Lichina pygmea* (*Lichinaceae, Lichinomycetes, Ascomycota*) and *Ramalina terebrata* (*Ramalinaceae, Lecanoromycetes, Ascomycota*) belong to two separates classes but produce an identical N2NP called pygmeine or romalin (Paudel et al., 2011; Roullier et al., 2010). Entomopathogenic fungi in the family *Cordycipitaceae* (*Sordariomycetes*, *Ascomycota*) produce *N*-amino-L-proline methyl ester and farylhydrazone (Humber et al., 2013; C. Ma et al., 2011; Mohamed et al., 2021, 2018; Weng et al., 2019). Gilled mushrooms (*Agaricus* spp.) and puffballs (*Calvatia* spp.) of the *Agaricaceae* (*Agaricomycetes*, *Basidiomycota*) produce agaritine and calvatic acid, among other N2NPs, while fungi in *Stephanosporaceae* and *Tricholomataceae* produce stephanosporin and lyophyllin (Chulia et al., 1988; Fugmann et al., 2001; Fugmann and Steglich, 1984; Le Goff and Ouazzani, 2014; Roupas et al., 2010; Schulzová et al., 2009; Umezawa et al., 1975; Ye et al., 1997). Another entomopathogenic fungus, *Entomophthora virulenta* (*Entomophthoraceae*, *Entomophthoromycetes*) from the phylum *Zoopagomycota* produces azoxybenzol-4,40-dicarboxylic acid (Claydon and Grove, 1978; He et al., 2022). Clearly, N2NPs are diverse and phylogenetically widespread in fungi—having been documented from three different phyla—yet the enzymatic machinery responsible for N–N bond formation in fungi remains completely unknown.

In this study, we took advantage of a recently generated, well-sampled genomic dataset for the gyromitrin-producing family *Discinaceae* and identified putative gyromitrin biosynthesis gene clusters (BGCs) using bioinformatics approaches. Our aims were to identify gene targets for future validation studies and to explore the potential of comparative genomics to identify genes of interest based on phenotypic data and a robust species tree. We hypothesized that the observed phylogenetic distribution of gyromitrin would be reflected in the evolutionary history

of the corresponding biosynthesis genes. The bioinformatics analyses consisted of (1) a homology search of known bacterial N2NP BGCs and (2) identification of phylogenetic incongruence between gene phylogenies and the inferred *Discinaceae* species phylogeny (Figure 4.2). Our search was guided by the fact that secondary metabolite biosynthesis genes are typically clustered or colocalized in the genome as a BGC (Cano-Prieto et al., 2024; Rokas et al., 2020). Discovery of a gyromitrin BGC would facilitate in-depth studies on the evolutionary history of gyromitrin and allow for genetic mining to identify other species with the potential to produce gyromitrin or related N2NPS.

4.2 Materials and methods

A total of 62 *Discinaceae* genomes generated in Chapter 3 were used in this study. Gene prediction and annotation were executed using the Funannotate 1.8.14 pipeline (github.com/nextgenusfs/funannotate) employing antiSMASH 7.1.0, eggNOG-mapper 2.1.9, and InterProScan 5.60.92.0, among other software and databases (Blin et al., 2023; Buchfink et al., 2021; Cantalapiedra et al., 2021; Drula et al., 2022; Eddy, 2011; Huerta-Cepas et al., 2019; Jones et al., 2014; Steinegger and Söding, 2017; The UniProt Consortium, 2023). Annotations for select genes of interest were checked against the Conserved Domain database of the National Center for Biotechnology Information (NCBI) (ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (J. Wang et al., 2022). The genome of *Gyromitra esculenta* CBS101906 hosted by JGI MycoCosm was the primary reference used throughout this manuscript due to its high quality, public accessibility, and the sequenced strain's production of gyromitrin. The *Gyromitra esculenta* MycoCosm portal can be accessed at mycocosm.jgi.doe.gov/Gyresc1/Gyresc1.home.html (Grigoriev et al., 2014). DNA and protein sequences and functional annotations of genes listed in this study can be viewed at mycocosm.jgi.doe.gov/cgi-

bin/dispGeneModel?db=Gyresc1&id=[protein_id], where "[protein_id]" is to be replaced with the protein ID number. Gene synteny diagrams were created with clinker 0.0.28 (Gilchrist and Chooi, 2021). A synopsis of the bioinformatics analyses conducted for this study is shown in Figure 4.3.

4.2.1 Homology search

A list of enzymes representing the known diversity of N–N bond formation biosynthetic pathways—largely related to secondary metabolism but also involving primary nitrogen metabolism (*e.g.,* anaerobic ammonium oxidation and denitrification)—was manually curated (Table 4.1). Protein sequences were downloaded from NCBI GenBank (Sayers et al., 2022). *Discinaceae* homologs were identified by using BLASTP and searching a custom *Discinaceae* proteomes database for matches with an E-value cutoff of 1E-5 (Camacho et al., 2009).

4.2.2 Phylogenetic incongruence search

Protein sequences were clustered into orthogroups with OrthoFinder 2.5.2 (Emms and Kelly, 2019). Orthogroups that did not contain at least one representative from both gyromitrin-producing clades (subg. *Gyromitra* and subg. *Leucoxantha*) or had fewer than four sequences were excluded. For each orthogroup, sequences were aligned with MAFFT 7.310 and trimmed with trimAl 1.4.rev22 using default settings (Capella-Gutiérrez et al., 2009; Katoh and Standley, 2013). Automatic model selection was conducted with ModelFinder and phylogenies were inferred with IQ-TREE2 2.2.6 using 1000 ultrafast bootstrap replicates (Hoang et al., 2018; Kalyaanamoorthy et al., 2017; Minh et al., 2020). Orthogroup trees were imported into R 4.2 to search for patterns of phylogenetic incongruence whereby gyromitrin-producing taxa were more closely related than expected according to the *Discinaceae* species tree (R Core Team, 2022).

For our first search pattern, the function "is.monophyletic" from *ape* 5.7-1 was used to identify orthogroup trees in which species belonging to subg. *Gyromitra* and subg. *Leucoxantha* formed a monophyletic clade (Paradis et al., 2004; Paradis and Schliep, 2019). For our second search pattern, the analysis was limited to orthogroup trees corresponding to antiSMASH-predicted BGC genes of our reference genome *Gyromitra esculenta* CBS101906. The most recent common ancestor of the gyromitrin-producing clades was identified using the function "getMRCA" from *ape*. Trees whose descendants of that node that did not include other non-gyromitrin taxa from the top half (subgenera *Discina*, *Caroliniana*, *Gigas*, and *Melaleucoides*) or the bottom half (subgenera *Hydnotrya*, *Paragyromitra*, *Pseudorhizina*, and *Pseudoverpa*) of the species tree were selected. This was achieved using the function "descendants" from *ips* 0.0.11 (github.com/heibl/ips). Phylogenetic trees were midpoint rooted and visualized with *ggtree* 3.6.2 (Yu et al., 2017). In short, we searched for a stringent pattern of monophyletic gyromitrin clades among all gene trees, and a more general pattern of phylogenetic incongruence among a subset of gene trees corresponding to the predicted *Gyromitra esculenta* BGC genes.

4.3 Results

4.3.1 Homology search

We queried 698 protein sequences belonging to 26 bacterial N2NP BGCs and three primary nitrogen metabolism enzymes. At an E-value cutoff of 1E-5, 323 queried proteins (46%) had BLASTP hits to *Discinaceae* proteins. No BLASTP hits were exclusively to the gyromitrin-producing species. Homologs of core N–N bond-producing enzymes from bacteria were largely absent in *Discinaceae*. There were, however, BLASTP hits to the fungal cytochrome P450 nitric oxide reductase, the methionyl-tRNA synthetase-like domain of the didomain protein PyrM, and diazo-forming enzyme CreM, a long-chain fatty acid CoA-ligase. Their low percent similarity,

widespread distribution in *Discinaceae* or absence from one of the gyromitrin-producing clades, lack of clustering with other interesting genes, and conserved nature reduced their appeal as putative gyromitrin genes. Most intriguing, all *Discinaceae* genomes possessed a single BLASTP hit with an E-value less than 1E-63 and similarity around 34% to the flavin-dependent ornithine *N*-monooxygenase KtzI and its homologs (JGI MycoCosm *G. esculenta* protein ID 528488). These enzymes produce an *N*-hydroxylamine that couples with an amino acid to form the N–N bond (Matsuda et al., 2024; Mügge et al., 2020). *Discinaceae* homologs of the *N*-monooxygenase occurred in a nonribosomal peptide synthase (NRPS) BGC, our first candidate gyromitrin BGC (putative gyromitrin BGC 1).

Putative gyromitrin BGC 1 (Figure 4.4) was present in all *Discinaceae* species, had a conserved gene synteny, and did not show prominent phylogenetic incongruence (Supplemental Figures 4.1–4.2). It consisted of an ATP-binding cassette (ABC) superfamily transporter (1.1; 613027), an NRPS with four complete and two incomplete modules (1.2; 477280), a hypothetical protein (HP) (1.3; 593487), the ornithine *N*-monooxygenase (1.4; 528488), another HP (1.5; 509158), a protein arginine *N*-methyltransferase (1.6; 593497), and a DNA damage inducible protein (1.7; 509160). Gene 1.4 was more than 50% identical to *Aspergillus fumigatus* SidA (GenBank AY586511), an *L*-ornithine-N^5-monooxygenase involved in siderophore production (Haas, 2014; Wallner et al., 2009).

4.3.2 Phylogenetic analysis

The 62 *Discinaceae* genomes contained a total of 690,475 genes, 97.2% of which were assigned to orthogroups. OrthoFinder inferred a total 21,232 orthogroups. Of these, 7823 (37%) met our filtering criteria—that is, they contained at least one representative from each of the gyromitrin-producing clades and had more than three taxa. For our first search pattern,

gyromitrin-producing species formed a monophyletic clade in 74 orthogroups (< 1%); 27 included *Gyromitra esculenta* (Supplemental Table 4.1), and only 11 included every sample belonging to the gyromitrin-positive subgenera (Supplemental Figure 4.3). Among these 27 monophyletic *G. esculenta* orthogroups, 22 were scattered throughout the genome. Two genes on scaffold 119 were colocalized (574706, 438843) but they had no known predicted function, nor did the flanking genes; this cluster was not considered further. However, another set of orthogroups on scaffold 12 emerged as a candidate biosynthesis cluster due to the genes' colocalization and their predicted functions (putative gyromitrin BGC 2).

Putative gyromitrin BGC 2 (Figure 4.4) consisted of five genes with prominent phylogenetic incongruence—four orthogroups with a gyromitrin-monophyletic pattern and one additional nonmonophyletic but similarly incongruent orthogroup, found via manual inspection of the phylogenies of the neighboring genes (Figures 4.5–4.6, Supplemental Figure 4.4). All five were colocalized in subg. *Leucoxantha* in one broadly conserved region across *Discinaceae*; four were colocalized in subg. *Gyromitra* in a different broadly conserved region, and a smaller subset was colocalized in subg. *Hydnotrya* in yet another broadly conserved region (Supplemental Figures 4.5–4.8). The single gene that was apart in subg. *Gyromitra* was in the same syntenic region as the colocalized cluster in subg. *Leucoxantha*. These genes were predicted to encode an amidohydrolase or peptidase (2.1; 529821); an iron-containing redox enzyme (2.2; 511368); an iron-dependent dioxygenase (2.3; 44179); a didomain protein with tyrosyl-tRNA synthetase and GCN5-related *N*-acetyltransferase (GNAT) domains (2.4; 546118); and an aldolase (2.5; 531962), this latter gene being the one that was apart from the main cluster in subg. *Gyromitra*. The GNAT domain of 2.4 was homologous to spb43, an acetyltransferase necessary for N–N bond formation in the hydrazone N2NP s56-p1 (Matsuda et al., 2018, 2017).

At an E-value cutoff of 1E-4, all samples from the gyromitrin-producing subgenera were hits to spb43 (along with *G. anthracobia*, which is not known to produce gyromitrin). The tyrosyl-tRNA synthetase and GNAT domains of 2.4 were present as full-length didomain genes exclusively in the gyromitrin-producing taxa; didomain proteins were also present in some truffloid taxa of subg. *Hydnotrya*, although the tyrosyl-tRNA synthetase domain appeared fragmented (Figure 4.6). The tyrosyl-tRNA synthetase domain of subg. *Leucoxantha* taxa was 54–60% identical to that of *Gyromitra esculenta* CBS101906, whereas subg. *Hydnotrya* taxa showed 36-44% similarity. The GNAT domain was more consistent with both groups around 63% identical to *G. esculenta*.

For our second search pattern, 94 orthogroups corresponding to *Gyromitra esculenta* BGC genes were analyzed. Nine orthogroups across seven BGCs matched our search pattern (Supplemental Figure 4.9). Upon manual inspection, five orthogroups were deemed spurious hits caused by the positioning of subg. *Gyromitra* next to the top-clade subgenera in a ladder-like topology; another two orthogroups were missing multiple gyromitrin-producing taxa and were not considered further. The two remaining orthogroups corresponded to three UstYa-family genes at the heart of a putative ribosomally synthesized and post-translationally modified peptide (RiPP) BGC on scaffold 33 (putative gyromitrin BGC 3). Putative gyromitrin BGC 3 (Figure 4.4) had a broadly conserved synteny among *Discinaceae* fungi (Supplemental Figure 4.10). The BGC was estimated to contain 16 genes predicted to encode an HP/2OG-Fe(II)-dependent oxygenase (3.1; 502429/559555), a mitochondrial C oxidase assembly factor (3.2; 559554), a transcriptional regulator (3.3; 524899), a HEAT repeat like-protein (3.4; 524898), an ABC transporter (3.5; 559543), a UstYa-family mycotoxin biosynthesis protein (3.6; 514692), an HP (3.7; 581295), two additional UstYa-family proteins (3.8–3.9; 559536, 514689), a glucose-

inhibited division-like protein (3.10; 581292), an ankyrin repeat domain-containing protein (3.11; 559520), a translational repressor of the PUF3 family (3.12; 467662), three HPs (3.13–1.15; NA, NA, 467628), and a rap/ran GTPase-activating protein (3.16; 199115). The inferred orthogroup phylogenies corresponding to this BGC largely corresponded to the species phylogeny (Supplemental Figure 4.11), except for the UstYa-family genes, which displayed notable incongruence with the *Discinaceae* species tree (Figure 4.7). The UstYa-like genes formed multiple paralogous clades with similar evolutionary histories. They were particularly abundant in subg. *Leucoxantha*, which contained up to five homologous genes.

In manually inspecting BGC 3 genes for a polypeptide that could serve as a RiPP precursor, we identified gene 3.7 as a likely candidate due to a predicted N-terminus signal peptide and repeated amino acid motifs (Nagano et al., 2016). Gene 3.7 consists of 11 (up to 19 in the case of *G. californica*) repeats of the eight-unit amino acid sequence DADSHQSY separated usually by the five-unit amino acid motif KRSEA but also ARSEA, RRSEA, RRGEA, and KRSEL, among other variations. Homologs of gene 3.7 are only present in subgenera *Gyromitra*, *Paragyromitra*, and *Pseudorhizina* (Supplemental Figure 4.10). For taxa in subg. *Leucoxantha*, no other potential RiPP precursor polypeptides were found in the region even though these species contain two RiPP-associated UstYa-family genes in BGC 3. A BLASTP search of the gene 3.7 orthogroup against the NCBI nonredundant protein database showed a complete absence from other *Pezizomycetes* genomes but a close affinity to *Dothideomycetes*. For example, gene 3.7 is 72–92% identical with 100% query coverage to XP_033378734 from *Aaosphaeria arxii* strain CBS 175.79. While the core repeat domains are identical (XP_033378734 consists of 23), differences in the arrangements of the variable spacer motifs affected their percent similarity.

4.4 Discussion

Bioinformatics analyses combining homology searches and phylogenetic comparisons resulted in three putative gyromitrin biosynthesis gene clusters. The pros and cons of each candidate BGC as they relate to gyromitrin biosynthesis are discussed below.

4.4.1 Putative gyromitrin BGC 1 likely produces an iron-chelating siderophore

Putative gyromitrin BGC 1, the predicted metallophore NRPS, is a candidate due to the presence of an ornithine *N*-monooxygenase (gene 1.4), which shares homology with *N*-hydroxylating enzymes that catalyze N–N bond formation. BGC 1 also possesses a methyltransferase (gene 1.6)—we would expect a gyromitrin BGC to contain a methyltransferase based on gyromitrin's chemical structure. However, BGC1 has several downsides that make it an unlikely candidate for gyromitrin production. The BGC is conserved among all *Discinaceae* species, which is inconsistent with the observed distribution of gyromitrin in the family (Dirks et al., 2023). Furthermore, the NRPS (gene 1.2) is unexpectedly large—an NRPS with one or two modules, not four or more, would be expected for the small size of gyromitrin. Moreover, while homologs of the monooxygenase do play a role in N–N bond formation, these enzymes are more commonly associated with the biosynthesis of siderophores—secondary metabolites used for the uptake, transport, and storage of iron from the environment. In the end, this may be the most likely function of BGC 1, especially since siderophores are produced by most fungal species and thus would be expected to occur in *Discinaceae* (Haas, 2014; Matsuda et al., 2024).

4.4.2 Putative gyromitrin BGC 2 is our leading hypothesis for gyromitrin biosynthesis

Putative gyromitrin BGC 2 consists of a substantial number of genes with inferred phylogenies showing a close relationship between the gyromitrin-positive subgenera (subg.

Gyromitra and subg. *Leucoxantha*) (Figure 4.5). This contrasts with the *Discinaceae* species tree in which they are only distantly related, sharing a most recent common ancestor some 80 million years ago (Figure 4.2). In searching thousands of gene trees for a particular phylogenetic pattern, false positives would be expected. The colocalization of five genes with similar patterns of phylogenetic incongruence, however, suggests a real biological signal.

A common pattern in these incongruent phylogenies—besides the clustering of gyromitrin-producing species—is the gyromitrin taxa showing a close relationship with truffloid taxa in subg. *Hydnotrya*. Two gene trees (genes 2.1 and 2.3) consist only of subg. *Hydnotrya* in addition to subg. *Gyromitra* and *Leucoxantha* and show the truffloid taxa as sister to the gyromitrin-producing clades. A third gene tree, the one corresponding to tyrosyl-tRNA synthetase and GNAT didomain gene 2.4, contains additional taxa, but subg. *Hydnotrya* is again sister to the gyromitrin-producing species and is the only other clade with both domains present in a single gene, albeit inconsistently and in a fragmented manner. A fourth, gene 2.2, shows a clade consisting of these same three subgenera with 100% bootstrap support, but here subg. *Leucoxantha* and subg. *Hydnotrya* were inferred to be more closely related. The branch supporting that relationship, however, only received 65% bootstrap support, meaning we are not confident about the exact relationship between subgenera *Gyromitra*, *Hydnotrya*, and *Leucoxantha*, even though we are confident that together they form a monophyletic clade. Taken together, these consistent patterns of stark phylogenetic incongruence are suggestive of horizontal gene transfer between ancestral populations of the gyromitrin trees.

The close relationship between the gyromitrin producers in the inferred phylogenies is what alerted us to BGC 2, but the predicted functions of the genes are also suggestive of a potential role in gyromitrin production. One or more oxidoreductases would be expected in the

gyromitrin BGC based on gyromitrin's chemical structure, which we find in BGC 2 (genes 2.2 and 2.3). Iron-dependent oxidoreductases in particular are interesting since they have been shown to catalyze oxidative rearrangements to form the N–N bond in streptozotocin biosynthesis (Ng et al., 2019; Yang and Chen, 2021). The GNAT domain of the didomain gene 2.4 is intriguing due to its homology to spb43, a gene necessary for s56-p1 hydrazone formation (Matsuda et al., 2017). GCN5-related *N*-acetyltransferases are also known to attach fatty acyl units to hydroxyornithine or lysine side chains to form lipopeptides (Hubrich et al., 2022). Gyromitrin's fatty acid moiety could be catalyzed by this GNAT domain, or perhaps by the aldolase (gene 2.5). The aldehyde-derived fatty acid side chain likely comes from the primary metabolic pool, as is typical for lipopeptides, resulting in the observed congener mixture of gyromitrin homologs (Hubrich et al., 2022; Pyysalo and Niskanen, 1977). Regardless of their exact function, both iron-dependent oxygenases and acetyltransferase-like enzymes frequently cooccur with N–N bond-forming genes, making BGC 2 a promising candidate (Matsuda et al., 2024).

The tyrosyl-tRNA synthetase domain of gene 2.4 is also an attractive target because one family of N–N bond-forming enzymes, the hydrazine synthetases, are didomain proteins that consist of an aminoacyl-tRNA synthetase (Matsuda and Wakimoto, 2024). The coexistence of the tyrosyl-tRNA synthetase with the GNAT domain in a didomain protein implies it is no longer involved in normal protein translation. Genomic duplication of aminoacyl-tRNA synthetases is a widespread and frequent phenomenon that facilitates the evolution of interesting new catalytic functions, like ones involved in secondary metabolite biosynthesis (Krahn et al., 2022). If BGC 2 were indeed the gyromitrin BGC, the lack of any stereotypical secondary metabolite backbone biosynthesis genes (such as an NRPS) would make it a noncanonical secondary metabolite BGC

(Rokas et al., 2020). However, N–N bond-forming genes are usually not colocalized with typical secondary metabolite backbone biosynthesis genes, and as a result most are overlooked by secondary metabolite prediction software (Matsuda et al., 2024). With a refined understanding of the types of aminoacyl-tRNA synthetases involved in secondary metabolite biosynthesis, like those belonging to multidomain proteins such as gene 2.4, aminoacyl-tRNA synthetases could become new targets for BGC prediction in the future.

A notable downside of BGC 2 as it relates to gyromitrin production is the lack of a methyltransferase, both among the phylogenetically incongruent genes as well as flanking genes. Interestingly, one of the 11 monophyletic orthogroups corresponds to a predicted S-adensyl-L-methionine (SAM)-dependent tetrapyrrole methylase. This raises the possibility that the genes responsible for gyromitrin biosynthesis are not wholly colocalized. If this were the case, putative gyromitrin BGC 1 could serve as an excellent companion to BGC 2 because of the ornithine *N*-monooxygenase and the presence of a methyltransferase. One could speculate that chemical intermediates of siderophore production are shunted into the gyromitrin biosynthesis pathway, linking iron scavenging with the iron-dependent enzymatic processes involved in gyromitrin production. These conjectures will need to be tested with isotopic labelling, gene knockout, and heterologous expression studies.

4.4.3 Putative gyromitrin BGC 3 likely encodes a RiPP cyclic peptide mycotoxin

The genetic makeup of BGC 3—a polypeptide precursor with repeating amino acid motifs (gene 3.7), RiPP-associated UstYa-family genes (genes 3.6, 3.8, and 3.9), and an ABC transporter (gene 3.5)—strongly suggests a role in the biosynthesis of a ribosomally synthesized and post-translationally modified peptide. Unlike NRPSs, which biosynthesize secondary metabolites without the use of ribosomal translation of genetically encoded polypeptides, RiPP

biosynthesis begins with the translation of a small gene. The linear peptide is cleaved by peptidases, and then a short polypeptide is cyclized and further modified by colocalized accessory enzymes to produce the final natural product (Cano-Prieto et al., 2024). The absence of a polypeptide precursor gene in subg. *Leucoxantha* implies production of the RiPP natural product is limited to the subgenera *Gyromitra*, *Paragyromitra*, and *Pseudorhizina*. The presence of RiPP-associated UstYa-family genes but absence of a precursor polypeptide coding sequence in subg. *Leucoxantha* BGC 3 is a phenomenon that has been observed in *Aspergillus* spp. via genome mining studies, but its biological implications are unknown (Nagano et al., 2016). As a result of this distribution, and due to the cyclic nature of RiPP natural products, BGC 3 would appear unlikely to be directly responsible for gyromitrin production but may instead produce an uncharacterized cyclic peptide. Characterized fungal RiPPs are potent mycotoxins, raising the possibility that an additional, undescribed mycotoxin is produced by the three above-mentioned subgenera. This could explain the credible poisoning reports associated with subg. *Paragyromitra*, which is not known to produce gyromitrin but is regarded as toxic (there are no data on the edibility of subg. *Pseudorhizina* as this group is rarely encountered, let alone consumed) (Dirks et al., 2023; Harmaja, 1976; Wells and Kempton, 1968). Additional RiPP BGCs are present across *Discinaceae*—including multiple additional RiPPs in subg. *Leucoxantha*—and warrant detailed study to characterize their associated natural products (Dirks et al. in review).

BGC 3 was selected due to the close relationship between the gyromitrin producers in the inferred phylogenies of the UstYa-family genes. The name of this gene family comes from its association with the mycotoxin ustiloxin, first characterized from the rice pathogenic fungus *Villosiclava virens* (= *Ustilaginoidea virens*) (*Sordariomycetes*) and later linked to a BGC in

Aspergillus spp. (*Eurotiomycetes*) (Koiso et al., 1998, 1994, 1992; Umemura et al., 2014, 2013; Ye et al., 2016). Follow-up studies showed that the UstYa family is a widespread, fungus-specific class of enzymes that catalyze diverse chemical reactions such as oxidative cyclization, hydroxylation, and desaturation (Nagano et al., 2016; Sogahata et al., 2021; Ye et al., 2019). Given their diversity and understudied nature, sequence-based prediction of UstYa-family protein function is currently not possible (Ozaki et al., 2023). The UstYa phylogeny appears to show a duplication of the gene in the early evolution of *Discinaceae*, resulting in two paralogs. The unexpected placement of paralogs from subg. *Leucoxantha* could be explained by a similar evolutionary history as the one conjectured for BGC 2. There appears to have been a horizontal gene transfer event to an ancestral population of subg. *Leucoxantha* and then an additional duplication in this clade. The precursor polypeptide's lack of homologs in other *Pezizomycetes* genomes and high similarity to *Dothideomycetes* homologs suggests BGC 3 has a convoluted, reticulate evolutionary history.

Contrary to its expected role in RiPP biosynthesis, BGC 3 lacks any predicted peptidases. Notably, one peptidase is present in putative gyromitrin BGC 2 and another is present among the flanking genes in *G. esculenta*, both of which could theoretically be involved in the cleaving of the RiPP precursor peptide. The asperipin-2a RiPP BGC lacks peptidase genes and is expected to make use of other peptidases outside of the gene cluster. Specifically, the Kex2 endopeptidase is thought to cleave the precursor polypeptide at the C-terminus sides of "KR" and "RR" motifs after it is exported to the Golgi apparatus, followed by further processing by the Kex1 carboxypeptidease (Mizuno et al., 1989; Nagano et al., 2016). Functional studies will be required to elucidate the chemical structure and biosynthesis pathway of this potentially novel cyclic peptide mycotoxin.

4.4.4 Conclusions

We employed bioinformatics analyses to identify putative gyromitrin biosynthesis gene clusters. Three candidate BGCs were identified. The NRPS cluster (BGC 1) was selected due to the presence of an *N*-hydroxylating monooxygenase with homology to known N–N bond-forming enzymes, but also to enzymes involved in siderophore production, which is the expected function of this cluster. The RiPP cluster (BGC 3) was selected due to phylogenetic incongruence in the UstYa-family gene trees. This BGC likely produces a cyclic peptide that could be an undescribed mycotoxin. BGC 2 is our leading candidate cluster and was selected due to prominent phylogenetic incongruence among four to five colocalized genes whereby the gyromitrin-producing subgenera were inferred to share a sister relationship. The phylogenies suggest horizontal gene transfer occurred between the ancestors of the gyromitrin-producing clades. If indeed associated with gyromitrin production, BGC 2 would represent a noncanonical BGC due to the lack of stereotypical backbone synthase genes. However, the tyrosyl-tRNA synthetase domain of the didomain protein is an intriguing candidate backbone synthase gene given that aminoacyl-tRNA synthetases are known to catalyze N–N bond formation in bacteria. Our study provides candidates for gene knockout and heterologous expression studies for validation of the gyromitrin BGC. As a member of the rare and interesting class of secondary metabolites with N–N bonds, elucidation of the gyromitrin BGC could greatly expand discovery of novel N2NPs among the chemodiverse and understudied nonmodel fungi.

Acknowledgements

We are grateful to Antonis Rokas, David Sherman, Fengan Yu, and Susan Egbert for insightful conversations on the biosynthesis of gyromitrin. Taylor Tai provided thoughtful comments on earlier drafts of the manuscript. This research was supported by the Michigan Translational Research and Commercialization (MTRAC) Innovation Hub for AgBio; the University of Michigan Rackham Graduate School and Department of Ecology and Evolution; and the Canadian Institute for Advanced Research (CIFAR).

Figures

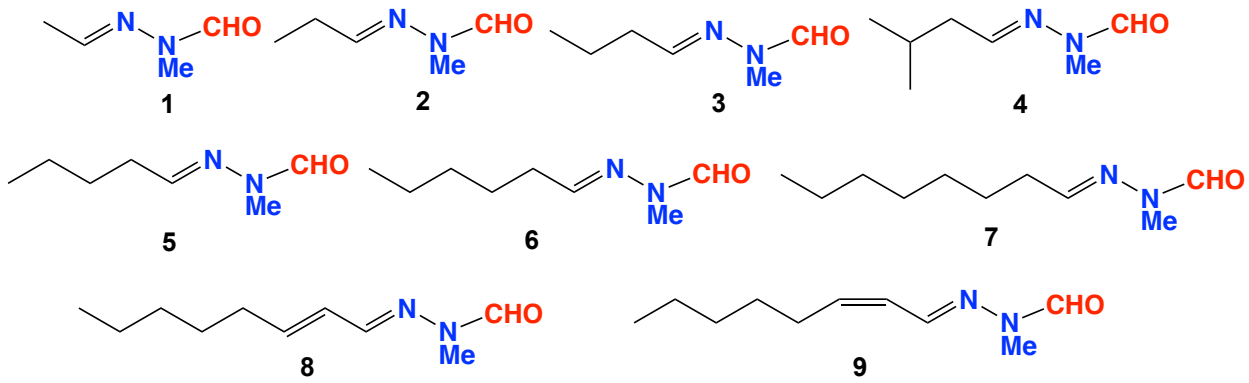


Figure 4.1. Chemical structures of gyromitrin (**1**) and its higher homologs (**2**–**9**).

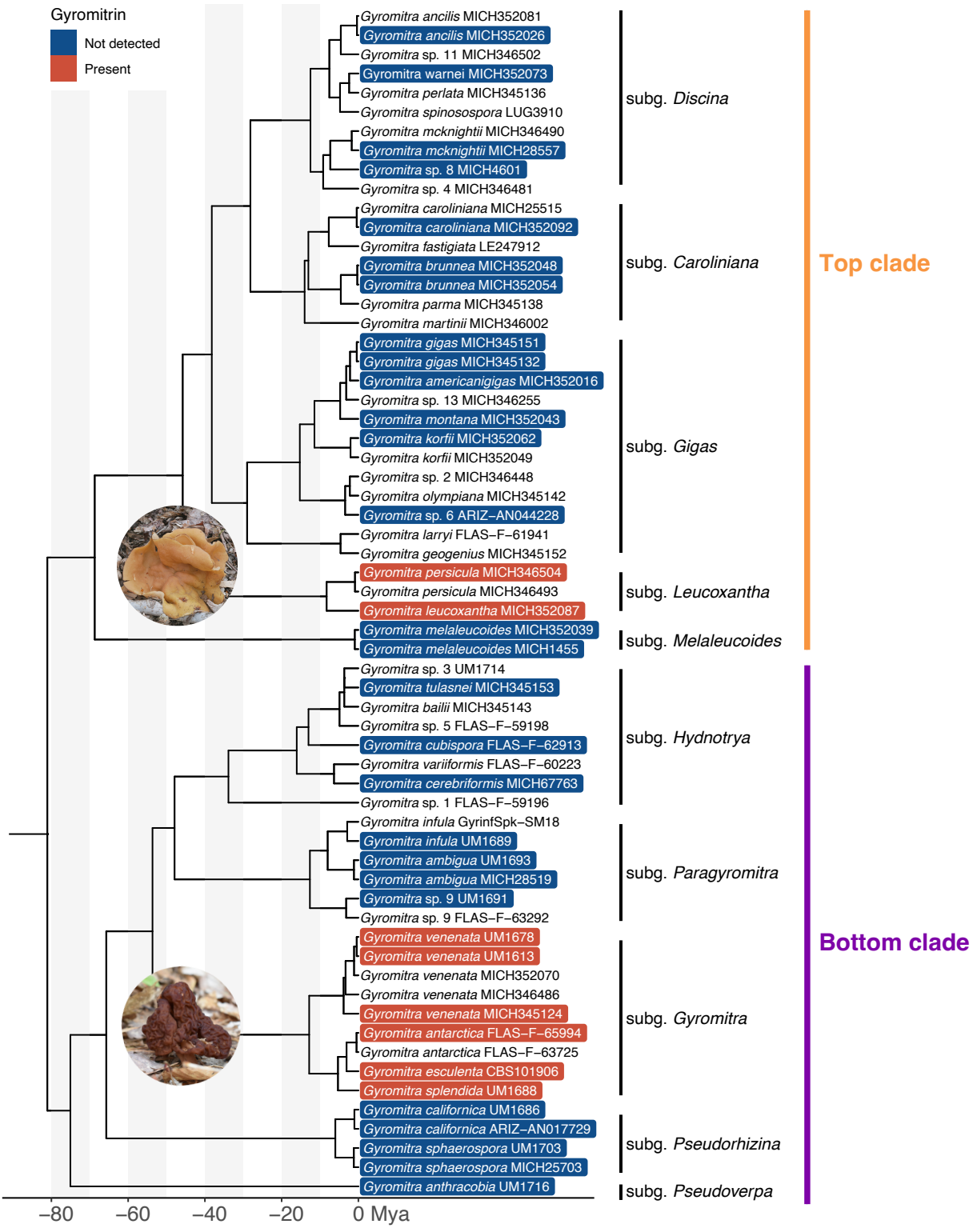


Figure 4.2. *Discinaceae* phylogenomic time divergence species tree and distribution of gyromitrin production. The species tree and taxonomy are from Chapter 2 and the gyromitrin information from Chapter 2 and Lagrange *et al.* (in review). Tip labels in red correspond to specimens that tested positive for gyromitrin; those in blue did not have detectable levels of gyromitrin; and ones without color were not tested. Colors corresponding to "top-clade taxa" and "bottom-clade taxa" (orange and purple, respectively) are used in phylogenetic trees throughout the study.

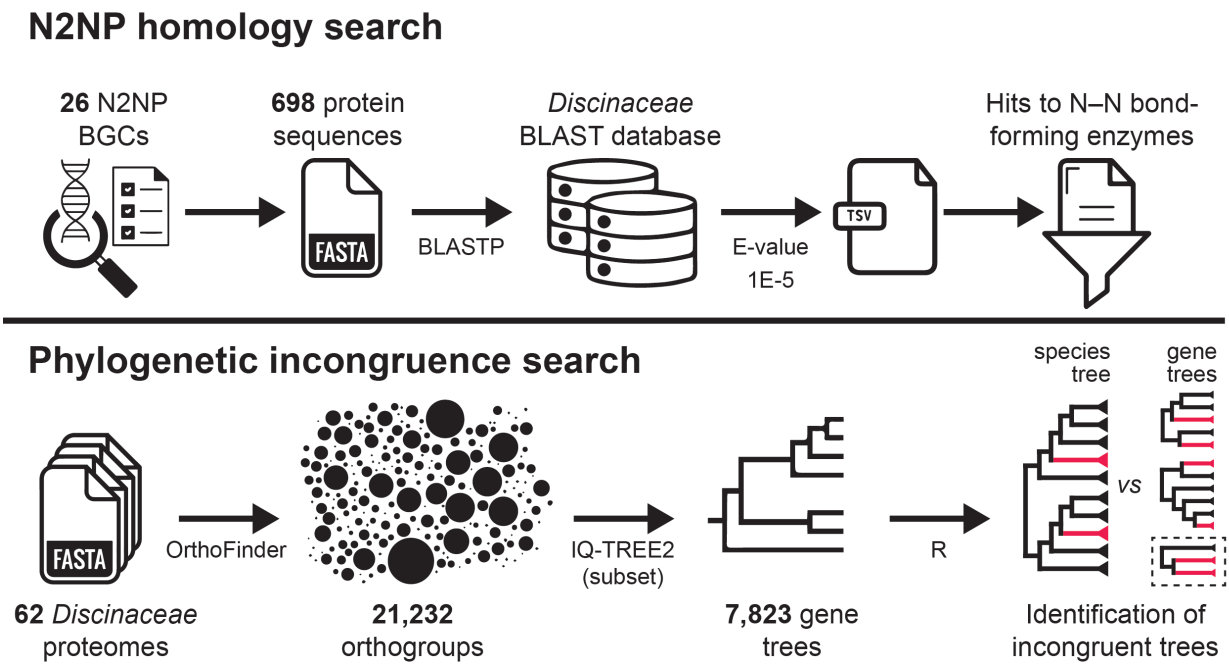
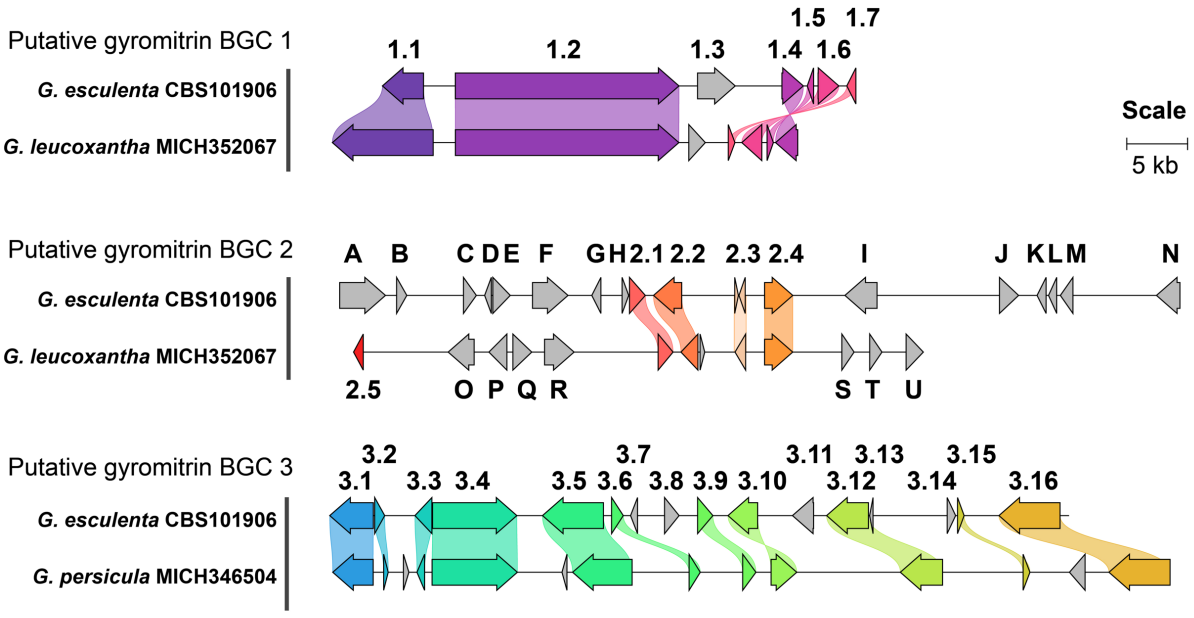


Figure 4.3. Schematic of bioinformatics analyses conducted in this study for the identification of putative gyromitrin biosynthesis gene clusters. *N2NP homology search.* A dataset of known N2NP biosynthesis gene clusters was manually curated. The protein sequences were extracted from NCBI GenBank and queried against a custom *Discinaceae* proteomes database with BLASTP. An E-value cutoff of 1E-5 was used to identify homologs. *Phylogenetic incongruence search*. 62 Discinaceae proteomes were clustered into homologous with OrthoFinder. Orthogroups with four or more taxa and consisting of at least one representative from both gyromitrin-producing clades (subg. *Gyromitra* and subg. *Leucoxantha*) had their protein phylogenies inferred with IQ-TREE2. Gene trees that showed the gyromitrin-producers as more closely related than expected, particularly when they formed a monophyletic clade, were selected (dashed box). Red branches indicate gyromitrin-producing clades.



Putative gyromitrin BGC 1
1.1: ABC transporter (613027)
1.2: NRPS (477280)
1.3: HP (593487)
1.4: monooxygenase (528488)
1.5: HP (509158)
1.6: methyltransferase (593497)
1.7: DNA damage inducible protein (509160)

Putative gyromitrin BGC 2
2.1: peptidase (529821)
2.2: iron-containing redox enzyme (511368)
2.3: iron-dependent dioxygenase (44179)
2.4: tyrosine-tRNA ligase and *N*-acetyltransferase (546118)
2.5: aldolase (531962)

BGC 2, conserved region A, flanking genes
A: heavy metal transporting ATPase (497149)
B: RNA recognition motif domain (497153)
C: HP (511362)
D: SNARE-associated protein (546087)
E: nitrogen permease regulator (546089)
F: serine protease (497157)
G: RNA recognition motif domain (43858)
H, I: HP (NA, 44421)
J: urease/HMG box-containing protein (44515)
K–M: pyridoxal-dependent decarboxylase (511374, 522115)
N: histone acetyltransferase (511377)

BGC 2, conserved region B, flanking genes
O: monooxygenase (561660)
P: aminotransferase (515232)
Q: MFS transporter (503193)
R: acyl-CoA synthetase (561667)
S: integral membrane protein (515235)
T: HP (581773)
U: dihydrolipoamide acetyltransferase (531970)

Putative gyromitrin BGC 3
3.1: HP/2OG-Fe(II)-dependent oxygenase (502429, 559555)
3.2: mitochondrial cytochorome C oxidase assembly factor (559554)
3.3: transciptional regulator (524899)
3.4: HEAT repeat protein-like (524898)
3.5: multidrug-resistance protein, ABC transporter (559543)
3.6: mycotoxin biosynthesis protein UstYa-family (514692)
3.7: HP/RiPP polypeptide precursor (581295)
3.8: UstYa-family (559536)
3.9: UstYa-family (514689)
3.10: glucose-inhibited division protein A-related (581292)
3.11: ankyrin repeat domain (559520)
3.12: translational repressor PUF3 (467662)
3.13, 3:14: HP
3.15: HP (467628)
3.16: rap/ran-GTPase-activating protein (199115)

Figure 4.4. Putative gyromitrin biosynthesis gene clusters (BGC) and their predicted functions. The gene maps were created with Clinker and are sized according to the scale. Genes belonging to the putative gyromitrin BGCs are colored. Homologous genes are linked by the corresponding color. Genes are listed with their predicted functioned according to antiSMASH and the annotated JGI *Gyromitra esculenta* CBS101906 reference genome. Protein IDs correspond to the *G. esculenta* JGI genome and can be accessed at "mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=[protein_id]", where "[protein_id]" is to be replaced with the protein ID number given in parentheses.

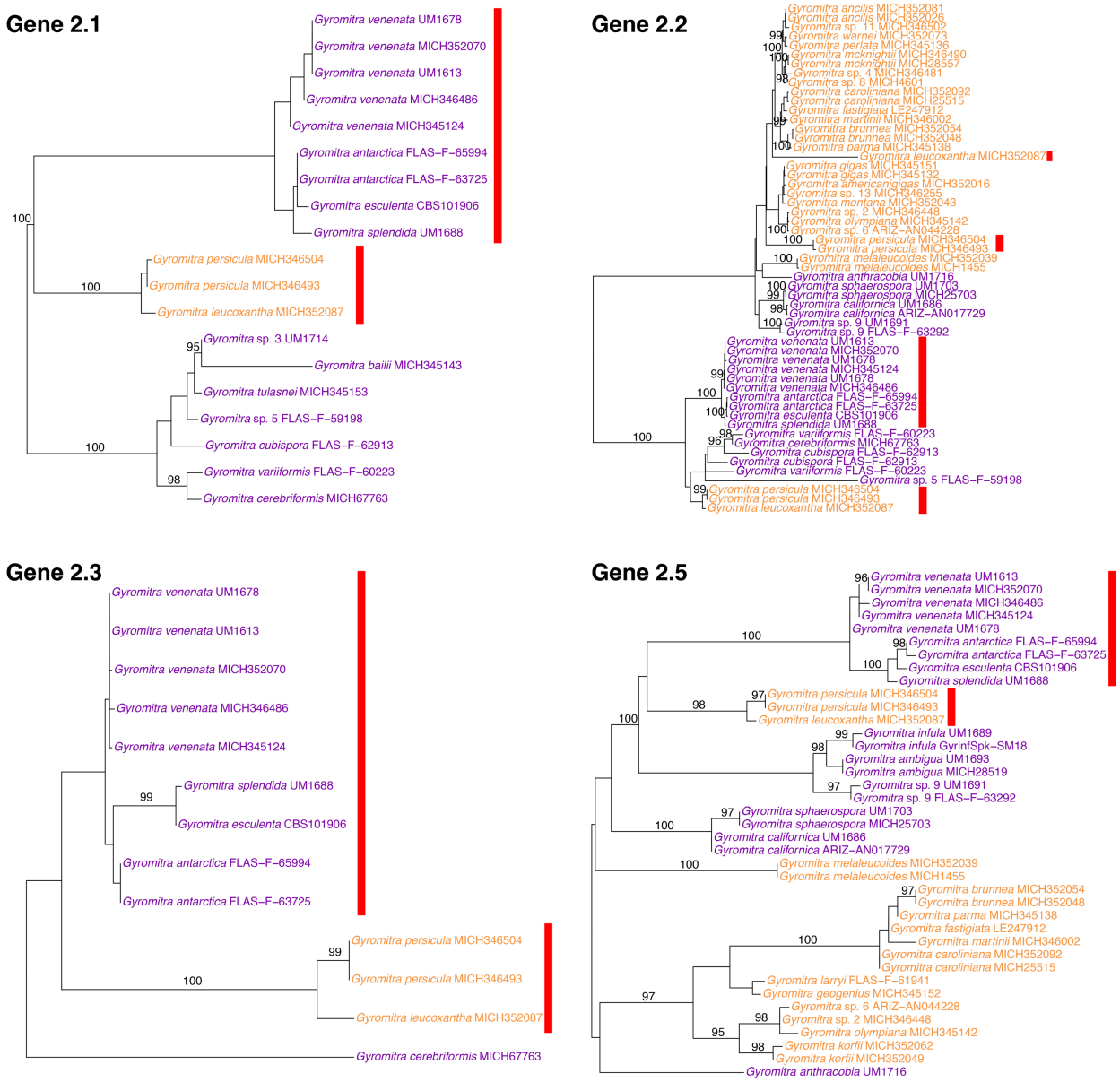


Figure 4.5. Phylogenetic tree of proteins homologous to genes 2.1, 2.2, 2.3, and 2.5 of putative gyromitrin biosynthesis gene cluster 2. The phylogenetic tree was inferred with IQ-TREE2 and midpoint rooted. Ultrafast bootstrap approximation support values greater than or equal to 95% are shown on the branches. Gyromitrin-producing clades are indicated by red bars. "Top-clade" taxa are colored orange, and "bottom-clade" taxa are colored purple (see Figure 4.2). See Figure 6 for the phylogenetic tree corresponding to gene 2.4. See the main text for the associated functions and *Gyromitra esculenta* CBS101906 MycoCosm protein ID.

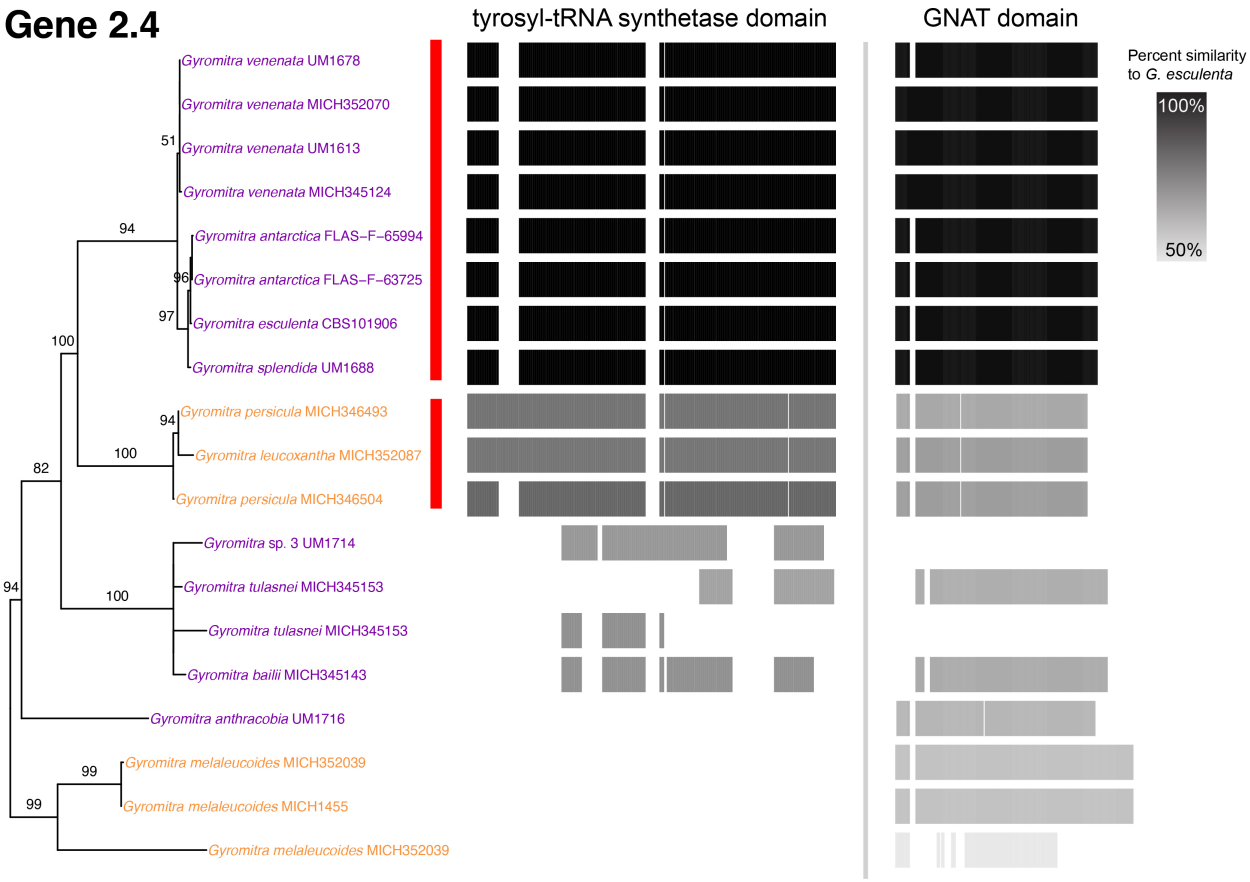


Figure 4.6. Phylogenetic tree of proteins homologous to gene 2.4 of putative gyromitrin biosynthesis gene cluster 2 and an alignment of the associated amino acid sequences. The phylogenetic tree was inferred with IQ-TREE2 and midpoint rooted. Branch values show ultrafast bootstrap approximation support. Gyromitrin-producing clades are indicated by red bars. "Top-clade" taxa are colored orange, and "bottom-clade" taxa are colored purple (see Figure 4.2). The bounds of the tyrosyl-tRNA synthetase and GCN5-related *N*-acetyltransferase (GNAT) domains of the didomain protein (gene 2.4) were identified with the NCBI Conserved Domains search and aligned separately with MUSCLE. The amino acid alignments are shaded according to their average percent similarity to the *Gyromitra esculenta* CBS101906 sequence. *Gyromitra venenata* MICH346486 was excluded due to erroneous gene prediction resulting in an artificially fragmented gene.

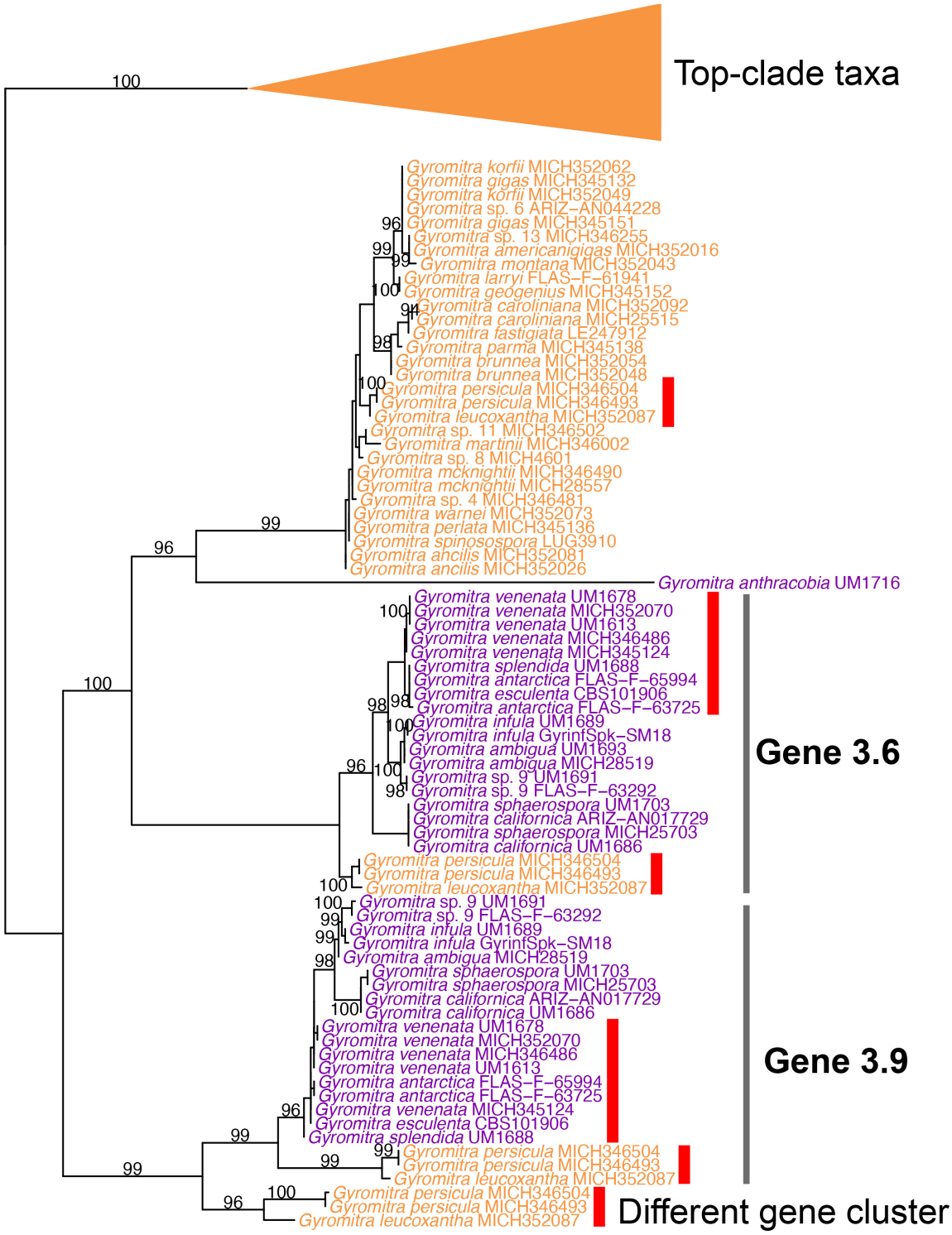


Figure 4.7. Phylogenetic tree of proteins homologous to UstYa-family genes 3.6 and 3.9 from putative gyromitrin biosynthesis gene cluster 3. The phylogenetic tree was inferred with IQ-TREE2 and midpoint rooted. Ultrafast bootstrap approximation support values greater than or equal to 95% are shown on the branches. Gyromitrin-producing clades are indicated by red bars. "Top-clade" taxa are colored orange, and "bottom-clade" taxa are colored purple (see Figure 4.2).

Tables

Table 4.1. Nitrogen-nitrogen bond-containing natural product biosynthesis gene clusters and nitrogen primary metabolism enzymes that were used as queries in a BLASTP homology search against a custom *Discinaceae* proteomes database.

Secondary metabolite/enzyme	**Organism**	**NCBI accession**	**Source**
Actinopyridazinone	*Streptomyces sp.* MSD090630SC-05	LC660652	(Matsuda et al., 2022)
Alazopeptin cluster 1	*Streptacidiphilus griseoplanus* JCM4300	LC593325	(Kawai et al., 2021)
Alazopeptin cluster 2	*Streptacidiphilus griseoplanus* JCM4300	LC593326	(Kawai et al., 2021)
Albofungin	*Streptomyces tumemacerans* JCM5050	ON399210	(Z.-C. Wang et al., 2022)
Aureothin	*Streptomyces thioluteus* HKI-22	AJ575648	(He and Hertweck, 2003)
Azaserine	*Streptomyces fragilis*	WP108953661–79	(Shikai et al., 2023; Van Cura et al., 2023; Wei et al., 2023)
Azodyrecin	*Streptomyces* sp. A1C6	LC712332	(Choirunnisa et al., 2022)
Cremeomycin	*Streptomyces cremeus* NRRL3241	KT381192	(Waldman et al., 2015)
Dixiamycin	*Streptomyces* sp. FXJ7.023	KR230086	(Uhong Lü et al., 2016)
Formycin	*Streptomyces kaniharaensis* SF-557	MK830994	(Wang et al., 2019)
Fosfazinomycin	*Streptomyces* sp. XY332	KF386860	(Yu et al., 2013)
Himastatin	*Streptomyces himastatinicus* ATCC53653	FR823394	(J. Ma et al., 2011)
Hydrazine synthase	*Kuenenia stuttgartiensis*	Q1Q0T2–4	(Strous et al., 2006)
Kinamycin	*Streptomyces murayamaensis* ATCC21414	MH703729	(Wang et al., 2018)
Kutznerides	*Kutzneria* sp. 744	EU074211	(Fujimori et al., 2007)
L-alanosine	*Streptomyces alanosinicus* DSM40606	MN626712	(Wang et al., 2020)
Lomaiviticin	*Salinispora pacifica* DPJ-0019	KF515737	(Janso et al., 2014)
Nitric oxide reductase (bacterial)	*Pseudomonas aeruginosa* PAO1	Q59646–7	(Arai et al., 1995)
Nitric oxide reductase (fungal)	*Fusarium oxysporum*	P23295	(Kizawa et al., 1991)
Nocuolin	*Nodularia* sp. LEGE 06071	MW071172	(Figueiredo et al., 2021; Martins et al., 2022)
Padanamide	*Streptomyces* sp. RJA2928	KC915040	(Du et al., 2013)
Pathocidin (8-azaguanine)	*Streptomyces pathocidini* ATCC 14510	MN707952	(Zhao et al., 2020a)
Pyrazomycin	*Streptomyces candidus* NRRL3601	MN170532	(Zhao et al., 2020b)
s56-p1	*Streptomyces* sp. SoC090715LN-17	LC177423	(Matsuda et al., 2017)
Sanglifehrin	*Streptomyces flaveolus* DSM 9954	FJ809786	(Qu et al., 2011)
Streptozotocin	*Streptomyces achromogenes* NRRL 2697	MK303572	(Ng et al., 2019)
Tasikamide	*Streptomyces tasikensis* P46	MZ964597	(Ma et al., 2022)
Toxoflavin	*Burkholderia glumae* BGR1	AY641455	(Kim et al., 2004; Suzuki et al., 2004)
Triacsins	*Kitasatospora aureofaciens* ATCC31442	MK932017	(Del Rio Flores et al., 2021; Twigg et al., 2019)
Valanimycin	*Streptomyces viridifaciens* MG456-hf10	AY116644	(Garg et al., 2002)

References

Arai, H., Igarashi, Y., Kodama, T., 1995. The structural genes for nitric oxide reductase from *Pseudomonas aeruginosa*. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1261, 279–284. https://doi.org/10.1016/0167-4781(95)00018-C

Bhandari, R., Eguchi, T., Sekine, A., Ohashi, Y., Kakinuma, K., Ito, M., Mizoue, K., 1999. Structure of NG-061, a novel potentiator of nerve growth factor (NGF) isolated from Penicillium minioluteum F-4627. J Antibiot (Tokyo) 52, 231–234. https://doi.org/10.7164/antibiotics.52.231

Blair, L.M., Sperry, J., 2013. Natural products containing a nitrogen–nitrogen bond. J. Nat. Prod. 76, 794–812. https://doi.org/10.1021/np400124n

Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B.R., Metcalf, W.W., Helfrich, E.J.N., van Wezel, G.P., Medema, M.H., Weber, T., 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research 51, W46–W50. https://doi.org/10.1093/nar/gkad344

Buchfink, B., Reuter, K., Drost, H.-G., 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18, 366–368. https://doi.org/10.1038/s41592-021-01101-x

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Cano-Prieto, C., Undabarrena, A., de Carvalho, A.C., Keasling, J.D., Cruz-Morales, P., 2024. Triumphs and challenges of natural product discovery in the postgenomic era. Annu Rev Biochem 93, 20.1-20.35. https://doi.org/10.1146/annurev-biochem-032620-104731

Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., Huerta-Cepas, J., 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Molecular Biology and Evolution 38, 5825–5829. https://doi.org/10.1093/molbev/msab293

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen, L., Deng, Z., Zhao, C., 2021. Nitrogen–nitrogen bond formation reactions involved in natural product biosynthesis. ACS Chem. Biol. 16, 559–570. https://doi.org/10.1021/acschembio.1c00052

Choirunnisa, A.R., Arima, K., Abe, Y., Kagaya, N., Kudo, K., Suenaga, H., Hashimoto, J., Fujie, M., Satoh, N., Shin-ya, K., Matsuda, K., Wakimoto, T., 2022. New azodyrecins identified

by a genome mining-directed reactivity-based screening. Beilstein J. Org. Chem. 18, 1017–1025. https://doi.org/10.3762/bjoc.18.102

Chulia, A.J., Bernillon, J., Favre-Bonvin, J., Kaouadji, M., Arpin, N., 1988. Isolation of β-*N*-(γ-glutamyl)-4-formylphenylhydrazine (agaritinal) from *Agaricus campestris*. Phytochemistry 27, 929–930. https://doi.org/10.1016/0031-9422(88)84122-0

Claydon, N., Grove, J.F., 1978. Metabolic products of *Entomophthora virulenta*. J. Chem. Soc., Perkin Trans. 1 171–173. https://doi.org/10.1039/P19780000171

Del Rio Flores, A., Twigg, F.F., Du, Y., Cai, W., Aguirre, D.Q., Sato, M., Dror, M.J., Narayanamoorthy, M., Geng, J., Zill, N.A., Zhai, R., Zhang, W., 2021. Biosynthesis of triacsin featuring an N-hydroxytriazene pharmacophore. Nat Chem Biol 17, 1305–1313. https://doi.org/10.1038/s41589-021-00895-3

Ding, T., Zhou, Y., Qin, J., Yang, L., Zhang, W., Shen, Y., 2020. Chemical constituents from wetland soil fungus *Penicillium oxalicum* GY1. Fitoterapia 142, 104530. https://doi.org/10.1016/j.fitote.2020.104530

Dirks, A.C., Mohamed, O.G., Schultz, P.J., Miller, A.N., Tripathi, A., James, T.Y., 2023. Not all bad: gyromitrin has a limited distribution in the false morels as determined by a new ultra high-performance liquid chromatography method. Mycologia 115, 1–15. https://doi.org/10.1080/00275514.2022.2146473

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., Terrapon, N., 2022. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Research 50, D571–D577. https://doi.org/10.1093/nar/gkab1045

Du, Y.-L., Dalisay, D.S., Andersen, R.J., Ryan, K.S., 2013. N-carbamoylation of 2,4-diaminobutyrate reroutes the outcome in padanamide biosynthesis. Chem Biol 20, 1002–1011. https://doi.org/10.1016/j.chembiol.2013.06.013

Eddy, S.R., 2011. Accelerated profile HMM searches. PLOS Computational Biology 7, e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Emms, D.M., Kelly, S., 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol 20, 238. https://doi.org/10.1186/s13059-019-1832-y

Figueiredo, S.A.C., Preto, M., Moreira, G., Martins, T.P., Abt, K., Melo, A., Vasconcelos, V.M., Leão, P.N., 2021. Discovery of cyanobacterial natural products containing fatty acid residues. Angewandte Chemie International Edition 60, 10064–10072. https://doi.org/10.1002/anie.202015105

Fugmann, B., Arnold, S., Steglich, W., Fleischhauer, J., Repges, C., Koslowski, A., Raabe, G., 2001. Pigments from the puffball *Calvatia rubro-flava* − isolation, structural elucidation and synthesis. European Journal of Organic Chemistry 2001, 3097–3104. https://doi.org/10.1002/1099-0690(200108)2001:16<3097::AID-EJOC3097>3.0.CO;2-V

Fugmann, B., Steglich, W., 1984. Unusual components of the toadstool *Lyophyllum connatum* (*Agaricales*). Angewandte Chemie International Edition in English 23, 72–73. https://doi.org/10.1002/anie.198400721

Fujimori, D.G., Hrvatin, S., Neumann, C.S., Strieker, M., Marahiel, M.A., Walsh, C.T., 2007. Cloning and characterization of the biosynthetic gene cluster for kutznerides. Proceedings of the National Academy of Sciences 104, 16498–16503. https://doi.org/10.1073/pnas.0708242104

Garg, R.P., Ma, Y., Hoyt, J.C., Parry, R.J., 2002. Molecular characterization and analysis of the biosynthetic gene cluster for the azoxy antibiotic valanimycin. Molecular Microbiology 46, 505–517. https://doi.org/10.1046/j.1365-2958.2002.03169.x

Gilchrist, C.L.M., Chooi, Y.-H., 2021. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37, 2473–2475. https://doi.org/10.1093/bioinformatics/btab007

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Research 42, D699–D704. https://doi.org/10.1093/nar/gkt1183

Haas, H., 2014. Fungal siderophore metabolism with a focus on *Aspergillus fumigatus*. Natural Product Reports 31, 1266–1276. https://doi.org/10.1039/C4NP00071D

Harmaja, H., 1976. Another poisonous species discovered in the genus *Gyromitra*: *G. ambigua*. Karstenia 15, 36–37.

He, H.-Y., Niikura, H., Du, Y.-L., Ryan, K.S., 2022. Synthetic and biosynthetic routes to nitrogen–nitrogen bonds. Chem. Soc. Rev. 51, 2991–3046. https://doi.org/10.1039/C7CS00458C

He, J., Hertweck, C., 2003. Iteration as programmed event during polyketide assembly; molecular analysis of the aureothin biosynthesis gene cluster. Chemistry & Biology 10, 1225–1232. https://doi.org/10.1016/j.chembiol.2003.11.009

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Molecular Biology and Evolution 35, 518–522. https://doi.org/10.1093/molbev/msx281

Hubrich, F., Bösch, N.M., Chepkirui, C., Morinaka, B.I., Rust, M., Gugger, M., Robinson, S.L., Vagstad, A.L., Piel, J., 2022. Ribosomally derived lipopeptides containing distinct fatty acyl moieties. Proceedings of the National Academy of Sciences 119, e2113120119. https://doi.org/10.1073/pnas.2113120119

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology

resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47, D309–D314. https://doi.org/10.1093/nar/gky1085

Humber, R.A., Rocha, L.F.N., Inglis, P.W., Kipnis, A., Luz, C., 2013. Morphology and molecular taxonomy of *Evlachovaea*-like fungi, and the status of this unusual conidial genus. Fungal Biology 117, 1–12. https://doi.org/10.1016/j.funbio.2012.10.001

Janso, J.E., Haltli, B.A., Eustáquio, A.S., Kulowski, K., Waldman, A.J., Zha, L., Nakamura, H., Bernan, V.S., He, H., Carter, G.T., Koehn, F.E., Balskus, E.P., 2014. Discovery of the lomaiviticin biosynthetic gene cluster in *Salinispora pacifica*. Tetrahedron 70, 4156–4164. https://doi.org/10.1016/j.tet.2014.03.009

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jurenitsch, J., Aurada, E., Kubelka, W., Robien, W., Viernstein, H., 1988. 1-(2-Hydroxyacetyl)-pyrazol, eine neue, mit gyromitrin verwandte verbindung aus *Gyromitra fastigiata* [1-(2-Hydroxyacetyl)-pyrazol, a new gyromitrin related compound of *Gyromitra fastigiata*]. Zeitschrift für Mykologie 54, 155–157.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30, 772–780. https://doi.org/10.1093/molbev/mst010

Katsuyama, Y., Matsuda, K., 2020. Recent advance in the biosynthesis of nitrogen–nitrogen bond–containing natural products. Current Opinion in Chemical Biology 59, 62–68. https://doi.org/10.1016/j.cbpa.2020.05.002

Kawai, S., Sugaya, Y., Hagihara, R., Tomita, H., Katsuyama, Y., Ohnishi, Y., 2021. Complete biosynthetic pathway of alazopeptin, a tripeptide consisting of two molecules of 6-diazo-5-oxo-L-norleucine and one molecule of alanine. Angewandte Chemie International Edition 60, 10319–10325. https://doi.org/10.1002/anie.202100462

Khalil, Z.G., Kankanamge, S., Capon, R.J., 2022. Structure revision of penipacids A–E reveals a putative new cryptic natural product, *N*-aminoanthranilic acid, with potential as a transcriptional regulator of silent secondary metabolism. Marine Drugs 20, 339. https://doi.org/10.3390/md20060339

Kim, J., Kim, J.-G., Kang, Y., Jang, J.Y., Jog, G.J., Lim, J.Y., Kim, S., Suga, H., Nagamatsu, T., Hwang, I., 2004. Quorum sensing and the LysR-type transcriptional activator ToxR

regulate toxoflavin biosynthesis and transport in Burkholderia glumae. Molecular Microbiology 54, 921–934. https://doi.org/10.1111/j.1365-2958.2004.04338.x

Kizawa, H., Tomura, D., Oda, M., Fukamizu, A., Hoshino, T., Gotoh, O., Yasui, T., Shoun, H., 1991. Nucleotide sequence of the unique nitrate/nitrite-inducible cytochrome P-450 cDNA from Fusarium oxysporum. Journal of Biological Chemistry 266, 10632–10637. https://doi.org/10.1016/S0021-9258(18)99270-7

Koiso, Y., Li, Y., Iwasaki, S., Hanaoka, K., Kobayashi, T., Sonoda, R., Fujita, Y., Yaegashi, H., Sato, Z., 1994. Ustiloxins, antimitotic cyclic peptides from false smut balls on rice panicles caused by *Ustilaginoidea virens*. J Antibiot (Tokyo) 47, 765–773. https://doi.org/10.7164/antibiotics.47.765

Koiso, Y., Morisaki, N., Yamashita, Y., Mitsui, Y., Shirai, R., Hashimoto, Y., Iwasaki, S., 1998. Isolation and structure of an antimitotic cyclic peptide, ustiloxin F: chemical interrelation with a homologous peptide, ustiloxin B. J. Antibiot. 51, 418–422. https://doi.org/10.7164/antibiotics.51.418

Koiso, Y., Natori, M., Iwasaki, S., Sato, S., Sonoda, R., Fujita, Y., Yaegashi, H., Sato, Z., 1992. Ustiloxin: a phytotoxin and a mycotoxin from false smut balls on rice panicles. Tetrahedron Letters 33, 4157–4160. https://doi.org/10.1016/S0040-4039(00)74677-6

Krahn, N., Söll, D., Vargas-Rodriguez, O., 2022. Diversification of aminoacyl-tRNA synthetase activities via genomic duplication. Front Physiol 13, 983245. https://doi.org/10.3389/fphys.2022.983245

Lagrange, E., Vernoux, J.P., Reis, J., Palmer, V., Camu, W., Spencer, P.S., 2021. An amyotrophic lateral sclerosis hot spot in the French Alps associated with genotoxic fungi. Journal of the Neurological Sciences 427, 117558. https://doi.org/10.1016/j.jns.2021.117558

Le Goff, G., Ouazzani, J., 2014. Natural hydrazine-containing compounds: biosynthesis, isolation, biological activities and synthesis. Bioorganic & Medicinal Chemistry 22, 6529–6544. https://doi.org/10.1016/j.bmc.2014.10.011

Leathem, A.M., Dorran, T.J., 2007. Poisoning due to raw *Gyromitra esculenta* (false morels) west of the Rockies. Canadian Journal of Emergency Medicine 9, 127–130. https://doi.org/10.1017/S1481803500014937

List, P.H., Luft, P., 1969. Nachweis und gehaltsbestimmung von gyromitrin in frischen lorcheln 19. Mitt. über pilzinhaltsstoffe. Arch Pharm Ber Dtsch Pharm Ges 302, 143–146. https://doi.org/10.1002/ardp.19693020209

List, P.H., Luft, P., 1968. Gyromitrin, das gift der frühjahrslorchel. 16. Mitt. über pilzinhaltsstoffe. Archiv Der Pharmazie 301, 294–305. https://doi.org/10.1002/ardp.19683010410

Ma, C., Li, Y., Niu, S., Zhang, H., Liu, X., Che, Y., 2011. *N*-hydroxypyridones, phenylhydrazones, and a quinazolinone from *Isaria farinosa*. Journal of natural products 74. https://doi.org/10.1021/np100568w

Ma, G.-L., Candra, H., Pang, L.M., Xiong, J., Ding, Y., Tran, H.T., Low, Z.J., Ye, H., Liu, M., Zheng, J., Fang, M., Cao, B., Liang, Z.-X., 2022. Biosynthesis of tasikamides via pathway coupling and diazonium-mediated hydrazone formation. J. Am. Chem. Soc. 144, 1622–1633. https://doi.org/10.1021/jacs.1c10369

Ma, J., Wang, Z., Huang, H., Luo, M., Zuo, D., Wang, B., Sun, A., Cheng, Y.-Q., Zhang, C., Ju, J., 2011. Biosynthesis of himastatin: assembly line and characterization of three cytochrome P450 enzymes involved in the post-tailoring oxidative steps. Angewandte Chemie International Edition 50, 7797–7802. https://doi.org/10.1002/anie.201102305

Martins, T.P., Glasser, N.R., Kountz, D.J., Oliveira, P., Balskus, E.P., Leão, P.N., 2022. Biosynthesis of the unusual carbon skeleton of nocuolin A. ACS Chem. Biol. 17, 2528–2537. https://doi.org/10.1021/acschembio.2c00464

Matsuda, K., Arima, K., Akiyama, S., Yamada, Y., Abe, Y., Suenaga, H., Hashimoto, J., Shin-ya, K., Nishiyama, M., Wakimoto, T., 2022. A natural dihydropyridazinone scaffold generated from a unique substrate for a hydrazine-forming enzyme. J. Am. Chem. Soc. 144, 12954–12960. https://doi.org/10.1021/jacs.2c05269

Matsuda, K., Hasebe, F., Shiwa, Y., Kanesaki, Y., Tomita, T., Yoshikawa, H., Shin-ya, K., Kuzuyama, T., Nishiyama, M., 2017. Genome mining of amino group carrier protein-mediated machinery: discovery and biosynthetic characterization of a natural product with unique hydrazone unit. ACS Chem. Biol. 12, 124–131. https://doi.org/10.1021/acschembio.6b00818

Matsuda, K., Nakahara, Y., Choirunnisa, A.R., Arima, K., Wakimoto, T., 2024. Phylogeny-guided characterization of bacterial hydrazine biosynthesis mediated by cupin/methionyl tRNA synthetase-like enzymes. Chembiochem e202300838. https://doi.org/10.1002/cbic.202300838

Matsuda, K., Tomita, T., Shin-ya, K., Wakimoto, T., Kuzuyama, T., Nishiyama, M., 2018. Discovery of unprecedented hydrazine-forming machinery in bacteria. J. Am. Chem. Soc. 140, 9083–9086. https://doi.org/10.1021/jacs.8b05354

Matsuda, K., Wakimoto, T., 2024. Bacterial hydrazine biosynthetic pathways featuring cupin/methionyl tRNA synthetase-like enzymes. ChemBioChem e202300874. https://doi.org/10.1002/cbic.202300874

Michelot, D., Toth, B., 1991. Poisoning by *Gyromitra esculenta*—a review. J of Applied Toxicology 11, 235–243. https://doi.org/10.1002/jat.2550110403

Minato, S., 1979. Isolation of anthglutin, an inhibitor of γ-glutamyl transpeptidase from *Penicillum oxalicum*. Archives of Biochemistry and Biophysics 192, 235–240. https://doi.org/10.1016/0003-9861(79)90088-2

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S., Matsuo, H., 1989. Characterization of KEX2-encoded endopeptidase from yeast *Saccharomyces cerevisiae*. Biochemical and Biophysical Research Communications 159, 305–311. https://doi.org/10.1016/0006-291X(89)92438-8

Mohamed, O.G., Khalil, Z.G., Capon, R.J., 2021. *N*-amino-L-proline methyl ester from an Australian fish gut-derived fungus: challenging the distinction between natural product and artifact. Marine Drugs 19, 151. https://doi.org/10.3390/md19030151

Mohamed, O.G., Khalil, Z.G., Capon, R.J., 2018. Prolinimines: *N*-amino-L-pro-methyl ester (hydrazine) Schiff bases from a fish gastrointestinal tract-derived fungus, *Trichoderma* sp. CMB-F563. Org. Lett. 20, 377–380. https://doi.org/10.1021/acs.orglett.7b03666

Mügge, C., Heine, T., Baraibar, A.G., van Berkel, W.J.H., Paul, C.E., Tischler, D., 2020. Flavin-dependent *N*-hydroxylating enzymes: distribution and application. Appl Microbiol Biotechnol 104, 6481–6499. https://doi.org/10.1007/s00253-020-10705-w

Nagano, N., Umemura, M., Izumikawa, M., Kawano, J., Ishii, T., Kikuchi, M., Tomii, K., Kumagai, T., Yoshimi, A., Machida, M., Abe, K., Shin-Ya, K., Asai, K., 2016. Class of cyclic ribosomal peptide synthetic genes in filamentous fungi. Fungal Genet Biol 86, 58–70. https://doi.org/10.1016/j.fgb.2015.12.010

Ng, T.L., Rohac, R., Mitchell, A.J., Boal, A.K., Balskus, E.P., 2019. An *N*-nitrosating metalloenzyme constructs the pharmacophore of streptozotocin. Nature 566, 94–99. https://doi.org/10.1038/s41586-019-0894-z

Ozaki, T., Minami, A., Oikawa, H., 2023. Recent advances in the biosynthesis of ribosomally synthesized and posttranslationally modified peptides of fungal origin. J Antibiot 76, 3–13. https://doi.org/10.1038/s41429-022-00576-w

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Paudel, B., Bhattarai, H.D., Koh, H.Y., Lee, S.G., Han, S.J., Lee, H.K., Oh, H., Shin, H.W., Yim, J.H., 2011. Ramalin, a novel nontoxic antioxidant compound from the Antarctic lichen *Ramalina terebrata*. Phytomedicine 18, 1285–1290. https://doi.org/10.1016/j.phymed.2011.06.007

Pyysalo, H., 1975. Some new toxic compounds in false morels, *Gyromitra esculenta*. Naturwissenschaften 62, 395. https://doi.org/10.1007/BF00625355

Pyysalo, H., Niskanen, A., 1977. On the occurrence of *N*-methyl-*N*-formylhydrazones in fresh and processed false morel, *Gyromitra esculenta*. J. Agric. Food Chem. 25, 644–647. https://doi.org/10.1021/jf60211a006

Qu, X., Jiang, N., Xu, F., Shao, L., Tang, G., Wilkinson, B., Liu, W., 2011. Cloning, sequencing and characterization of the biosynthetic gene cluster of sanglifehrin A, a potent cyclophilin inhibitor. Mol. BioSyst. 7, 852–861. https://doi.org/10.1039/C0MB00234H

R Core Team, 2022. R: a language and environment for statistical computing. R Foundation for Statistical Computing.

Rokas, A., Mead, M.E., Steenwyk, J.L., Raja, H.A., Oberlies, N.H., 2020. Biosynthetic gene clusters and the evolution of fungal chemodiversity. Nat Prod Rep 37, 868–878. https://doi.org/10.1039/c9np00045c

Roullier, C., Chollet-Krugler, M., Weghe, P. van de, Devehat, F.L.-L., Boustie, J., 2010. A novel aryl-hydrazide from the marine lichen *Lichina pygmaea*: isolation, synthesis of derivatives, and cytotoxicity assays. Bioorganic & Medicinal Chemistry Letters 20, 4582–4586. https://doi.org/10.1016/j.bmcl.2010.06.013

Roupas, P., Keogh, J., Noakes, M., Margetts, C., Taylor, P., 2010. Mushrooms and agaritine: a mini-review. Journal of Functional Foods 2, 91–98. https://doi.org/10.1016/j.jff.2010.04.003

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Connor, R., Funk, K., Kelly, C., Kim, S., Madej, T., Marchler-Bauer, A., Lanczycki, C., Lathrop, S., Lu, Z., Thibaud-Nissen, F., Murphy, T., Phan, L., Skripchenko, Y., Tse, T., Wang, J., Williams, R., Trawick, B.W., Pruitt, K.D., Sherry, S.T., 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50, D20–D26. https://doi.org/10.1093/nar/gkab1112

Schulzová, V., Hajslova, J., Peroutka, R., Hlavasek, J., Gry, J., Andersson, H.C., 2009. Agaritine content of 53 *Agaricus* species collected from nature. Food Additives & Contaminants: Part A 26, 82–93. https://doi.org/10.1080/02652030802039903

Shikai, Y., Kawai, S., Katsuyama, Y., Ohnishi, Y., 2023. In vitro characterization of nonribosomal peptide synthetase-dependent O -(2-hydrazineylideneacetyl)serine synthesis indicates a stepwise oxidation strategy to generate the α-diazo ester moiety of azaserine. Chemical Science 14, 8766–8776. https://doi.org/10.1039/D3SC01906C

Sitta, N., Davoli, P., Floriani, M., Suriano, E., 2021. Guida ragionata alla commestibilità dei funghi, 1st ed. Regione Piemonte, Italy.

Sogahata, K., Ozaki, T., Igarashi, Y., Naganuma, Y., Liu, C., Minami, A., Oikawa, H., 2021. Biosynthetic studies of phomopsins unveil posttranslational installation of dehydroamino

acids by UstYa family proteins. Angew. Chem. Int. Ed. 60, 25729–25734. https://doi.org/10.1002/anie.202111076

Steinegger, M., Söding, J., 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol 35, 1026–1028. https://doi.org/10.1038/nbt.3988

Strous, M., Pelletier, E., Mangenot, S., Rattei, T., Lehner, A., Taylor, M.W., Horn, M., Daims, H., Bartol-Mavel, D., Wincker, P., Barbe, V., Fonknechten, N., Vallenet, D., Segurens, B., Schenowitz-Truong, C., Médigue, C., Collingro, A., Snel, B., Dutilh, B.E., Op den Camp, H.J.M., van der Drift, C., Cirpus, I., van de Pas-Schoonen, K.T., Harhangi, H.R., van Niftrik, L., Schmid, M., Keltjens, J., van de Vossenberg, J., Kartal, B., Meier, H., Frishman, D., Huynen, M.A., Mewes, H.-W., Weissenbach, J., Jetten, M.S.M., Wagner, M., Le Paslier, D., 2006. Deciphering the evolution and metabolism of an anammox bacterium from a community genome. Nature 440, 790–794. https://doi.org/10.1038/nature04647

Suzuki, F., Sawada, H., Azegami, K., Tsuchiya, K., 2004. Molecular characterization of the tox operon involved in toxoflavin biosynthesis of Burkholderia glumae. J Gen Plant Pathol 70, 97–107. https://doi.org/10.1007/s10327-003-0096-1

Svanberg, I., Lindh, H., 2019. Mushroom hunting and consumption in twenty-first century post-industrial Sweden. Journal of Ethnobiology and Ethnomedicine 15, 42. https://doi.org/10.1186/s13002-019-0318-z

Tang, C., Gao, R., Tang, X., Zhang, Y., Feng, W., Feng, B., Lu, X., 2023. Metabolites isolated from *Penicillium* HDS-Z-1E, an endophytic fungal strain isolated from *Taxus cuspidata* and their activation effect of catalase. Chinese Herbal Medicines. https://doi.org/10.1016/j.chmed.2022.12.009

The UniProt Consortium, 2023. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research 51, D523–D531. https://doi.org/10.1093/nar/gkac1052

Twigg, F.F., Cai, W., Huang, W., Liu, J., Sato, M., Perez, T.J., Geng, J., Dror, M.J., Montanez, I., Tong, T.L., Lee, H., Zhang, W., 2019. Identifying the biosynthetic gene cluster for triacsins with an *N*-hydroxytriazene moiety. ChemBioChem 20, 1145–1149. https://doi.org/10.1002/cbic.201800762

Uhong Lü, Y., Liu, X., Wang, M., Li, Y., Liu, N., Bao, Y., Liu, M., Li, X., Wang, Y., Qian, S., Yue, C., Huang, Y., 2016. Activation and comparative analysis of cryptic xiamycin gene cluster from marine-derived Streptomyces sp. FXJ 7.388. Pak J Pharm Sci 29, 1811–1815.

Umemura, M., Koike, H., Nagano, N., Ishii, T., Kawano, J., Yamane, N., Kozone, I., Horimoto, K., Shin-ya, K., Asai, K., Yu, J., Bennett, J.W., Machida, M., 2013. MIDDAS-M: motif-independent *de novo* detection of secondary metabolite gene clusters through the integration of genome sequencing and transcriptome data. PLOS ONE 8, e84028. https://doi.org/10.1371/journal.pone.0084028

Umemura, M., Nagano, N., Koike, H., Kawano, J., Ishii, T., Miyamura, Y., Kikuchi, M., Tamano, K., Yu, J., Shin-ya, K., Machida, M., 2014. Characterization of the biosynthetic gene cluster for the ribosomally synthesized cyclic peptide ustiloxin B in *Aspergillus flavus*. Fungal Genetics and Biology 68, 23–30. https://doi.org/10.1016/j.fgb.2014.04.011

Umezawa, H., Takeuchi, T., Iinuma, H., Ito, M., Ishizuka, M., Kurakata, Y., Umeda, Y., Nakanishi, Y., Nakamura, T., Obayashi, A., Tanabe, O., 1975. A new antibiotic, calvatic acid. J. Antibiot. 28, 87–90. https://doi.org/10.7164/antibiotics.28.87

Van Cura, D., Ng, T.L., Huang, J., Hager, H., Hartwig, J.F., Keasling, J.D., Balskus, E.P., 2023. Discovery of the azaserine biosynthetic pathway uncovers a biological route for α-diazoester production. Angewandte Chemie International Edition 62, e202304646. https://doi.org/10.1002/anie.202304646

Waldman, A.J., Pechersky, Y., Wang, P., Wang, J.X., Balskus, E.P., 2015. The cremeomycin biosynthetic gene cluster encodes a pathway for diazo formation. Chembiochem 16, 2172–2175. https://doi.org/10.1002/cbic.201500407

Wallner, A., Blatzer, M., Schrettl, M., Sarg, B., Lindner, H., Haas, H., 2009. Ferricrocin, a siderophore involved in intra- and transcellular iron distribution in *Aspergillus fumigatus*. Applied and Environmental Microbiology 75, 4194–4196. https://doi.org/10.1128/AEM.00479-09

Wang, J., Chitsaz, F., Derbyshire, M.K., Gonzales, N.R., Gwadz, M., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C.J., Marchler-Bauer, A., 2022. The conserved domain database in 2023. Nucleic Acids Res 51, D384–D388. https://doi.org/10.1093/nar/gkac1096

Wang, K.-K.A., Ng, T.L., Wang, P., Huang, Z., Balskus, E.P., van der Donk, W.A., 2018. Glutamic acid is a carrier for hydrazine during the biosyntheses of fosfazinomycin and kinamycin. Nat Commun 9, 3687. https://doi.org/10.1038/s41467-018-06083-7

Wang, M., Niikura, H., He, H.-Y., Daniel-Ivad, P., Ryan, K.S., 2020. Biosynthesis of the N–N-bond-containing compound L-alanosine. Angewandte Chemie International Edition 59, 3881–3885. https://doi.org/10.1002/anie.201913458

Wang, S.-A., Ko, Y., Zeng, J., Geng, Y., Ren, D., Ogasawara, Y., Irani, S., Zhang, Y., Liu, H., 2019. Identification of the formycin A biosynthetic gene cluster from *Streptomyces kaniharaensis* illustrates the interplay between biological pyrazolopyrimidine formation and de novo purine biosynthesis. J. Am. Chem. Soc. 141, 6127–6131. https://doi.org/10.1021/jacs.9b00241

Wang, Z.-C., Lo, I.-W., Lin, K.-H., Cheng, A.N., Zadeh, S.M., Huang, Y.-H., Li, T.-L., 2022. Genetic and biochemical characterization of halogenation and drug transportation genes encoded in the albofungin biosynthetic gene cluster. Applied and Environmental Microbiology 88, e00806-22. https://doi.org/10.1128/aem.00806-22

Wei, Z.-W., Niikura, H., Wang, M., Ryan, K.S., 2023. Identification of the azaserine biosynthetic gene cluster implicates hydrazine as an intermediate to the diazo moiety. Org. Lett. 25, 4061–4065. https://doi.org/10.1021/acs.orglett.3c01229

Wells, V.L., Kempton, P.E., 1968. Studies on the fleshy fungi of Alaska. II. Mycologia 60, 888–901. https://doi.org/10.1080/00275514.1968.12018645

Weng, Q., Zhang, X., Chen, W., Hu, Q., 2019. Secondary metabolites and the risks of *Isaria fumosorosea* and *Isaria farinosa*. Molecules 24. https://doi.org/10.3390/molecules24040664

Yang, H., Chen, S.-L., 2021. Enzymatic N–N bond formation: mechanism for the *N*-nitroso synthesis catalyzed by non-heme iron SznF enzyme. Journal of Catalysis 398, 44–53. https://doi.org/10.1016/j.jcat.2021.04.010

Ye, Y., Aulinger, K., Arnold, N., Spahl, W., Steglich, W., 1997. Biosynthesis of the azoxycarboxamide lyophyllin and formation of some of its unnatural analogues in fruit-bodies of *Lyophyllum connatum*. Tetrahedron Letters 38, 8013–8016. https://doi.org/10.1016/S0040-4039(97)10177-0

Ye, Y., Minami, A., Igarashi, Y., Izumikawa, M., Umemura, M., Nagano, N., Machida, M., Kawahara, T., Shin-ya, K., Gomi, K., Oikawa, H., 2016. Unveiling the biosynthetic pathway of the ribosomally synthesized and post-translationally modified peptide ustiloxin B in filamentous fungi. Angewandte Chemie 128, 8204–8207. https://doi.org/10.1002/ange.201602611

Ye, Y., Ozaki, T., Umemura, M., Liu, C., Minami, A., Oikawa, H., 2019. Heterologous production of asperipin-2a: proposal for sequential oxidative macrocyclization by a fungi-specific DUF3328 oxidase. Org. Biomol. Chem. 17, 39–43. https://doi.org/10.1039/C8OB02824A

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28–36. https://doi.org/10.1111/2041-210X.12628

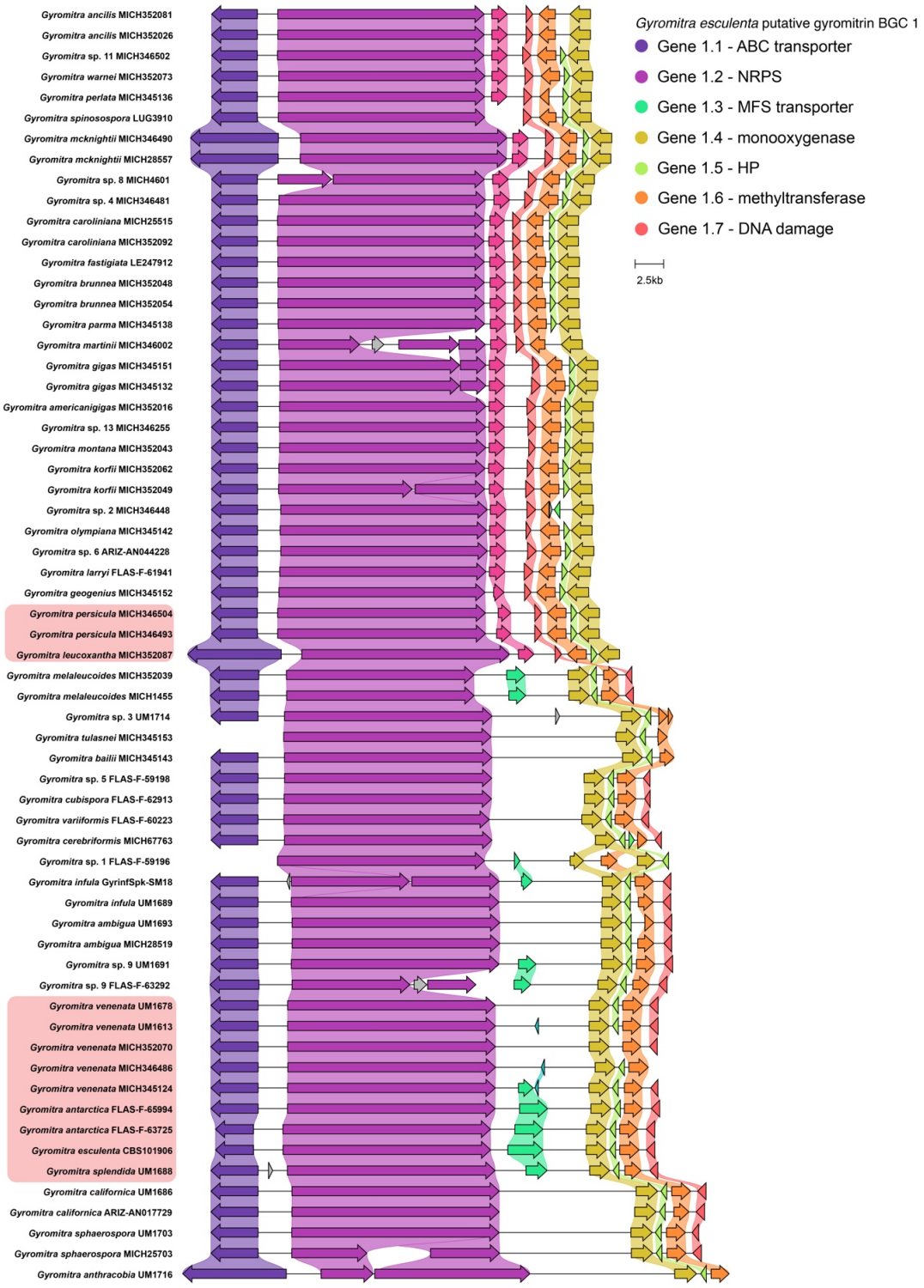
Yu, X., Doroghazi, J.R., Janga, S.C., Zhang, J.K., Circello, B., Griffin, B.M., Labeda, D.P., Metcalf, W.W., 2013. Diversity and abundance of phosphonate biosynthetic genes in nature. Proceedings of the National Academy of Sciences 110, 20759–20764. https://doi.org/10.1073/pnas.1315107110

Zhang, T., Zhu, M.-L., Sun, G.-Y., Li, N., Gu, Q.-Q., Li, D.-H., Che, Q., Zhu, T.-J., 2016. Exopisiod B and farylhydrazone C, two new alkaloids from the Antarctic-derived fungus *Penicillium* sp. HDN14-431. Journal of Asian Natural Products Research 18, 959–965. https://doi.org/10.1080/10286020.2016.1174699

Zhao, G., Guo, Y.-Y., Yao, S., Shi, X., Lv, L., Du, Y.-L., 2020a. Nitric oxide as a source for bacterial triazole biosynthesis. Nat Commun 11, 1614. https://doi.org/10.1038/s41467-020-15420-8

Zhao, G., Yao, S., Rothchild, K.W., Liu, T., Liu, Y., Lian, J., He, H.-Y., Ryan, K.S., Du, Y.-L., 2020b. The biosynthetic gene cluster of pyrazomycin—a C-nucleoside antibiotic with a rare pyrazole moiety. ChemBioChem 21, 644–649. https://doi.org/10.1002/cbic.201900449

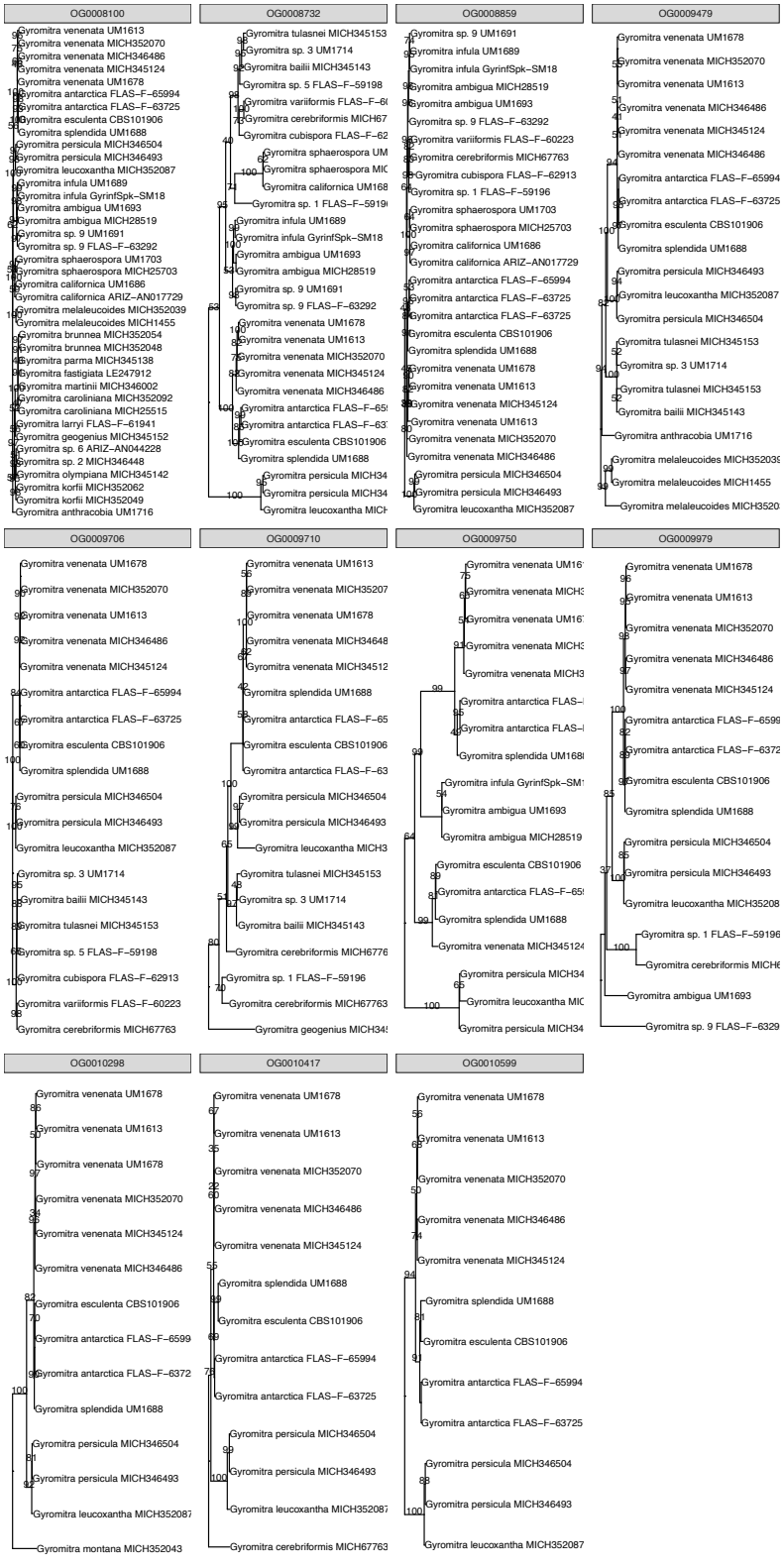
Supplement



Supplemental Figure 4.1. Gene synteny map of putative gyromitrin biosynthesis gene cluster 1. Red highlights indicate gyromitrin-producing clades.



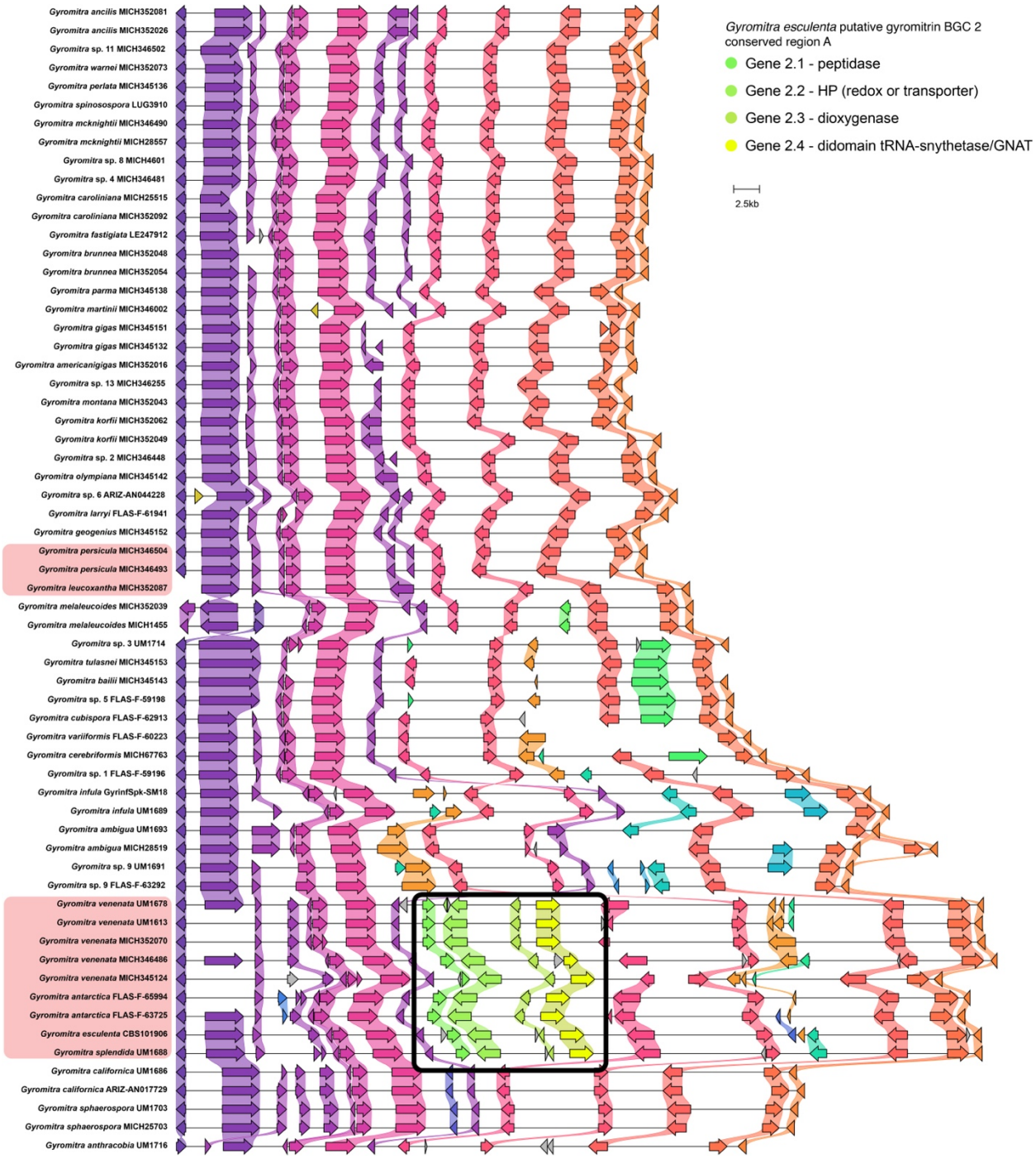
Supplemental Figure 4.2. Phylogenetic trees of the orthogroups corresponding to putative gyromitrin biosynthesis gene cluster 1 inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. Homologs of gene 1.3 were absent from *Gyromitra* subgen. *Leucoxantha* and no phylogeny was generated. See the main text for corresponding protein IDs.



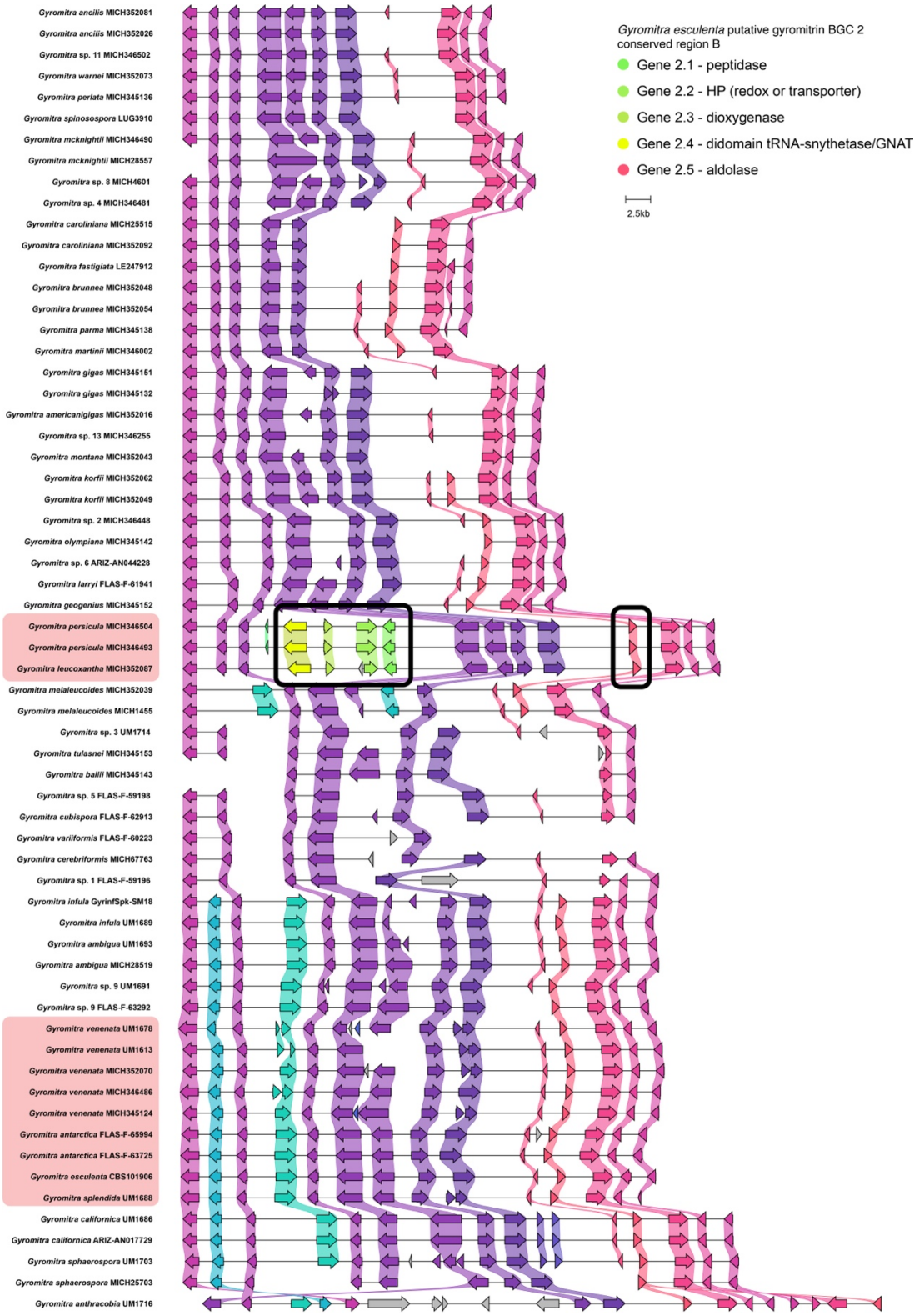
Supplemental Figure 4.3. Phylogenetic trees of orthogroups with all gyromitrin-producing samples present and inferred to form a monophyletic clade with IQ-TREE2 using 1000 ultrafast bootstrap replicates. See Table S1 for corresponding protein IDs.



Supplemental Figure 4.4. Phylogenetic trees of orthogroups corresponding to putative gyromitrin biosynthesis gene cluster 2 and flanking genes inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. The phylogenetic trees are organized in the same order as they appear from left to right for *Gyromitra* subg. *Leucoxantha* (*Gyromitra leucoxantha* and *Gyromitra persicula*) in Figure 6.



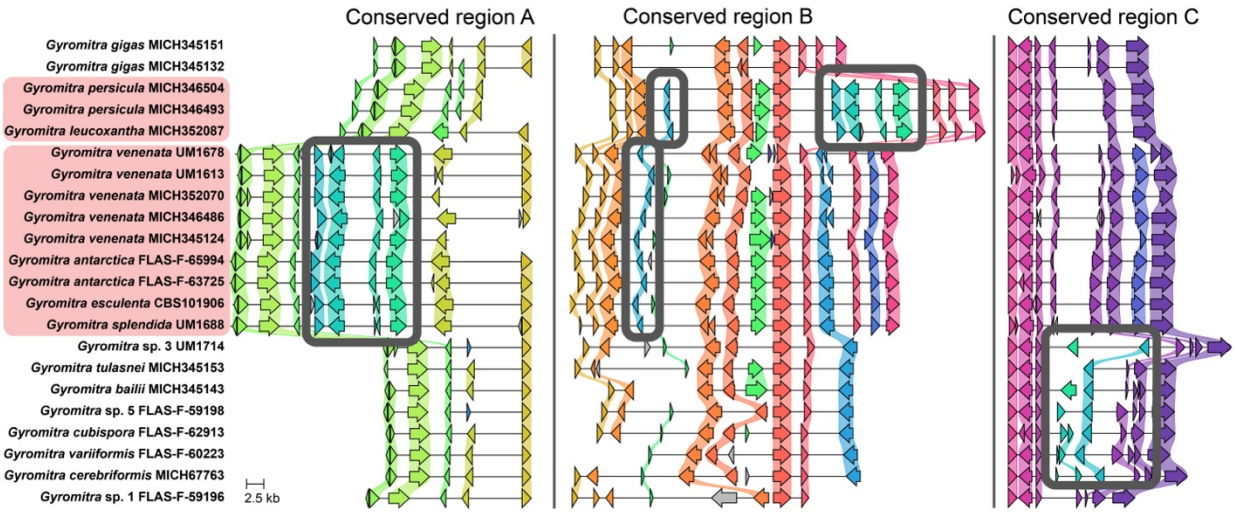
Supplemental Figure 4.5. Gene synteny map of putative gyromitrin biosynthesis gene cluster 2, conserved region A. The putative gyromitrin BGC 2 genes are indicated by a box. Red highlights indicate gyromitrin-producing clades.



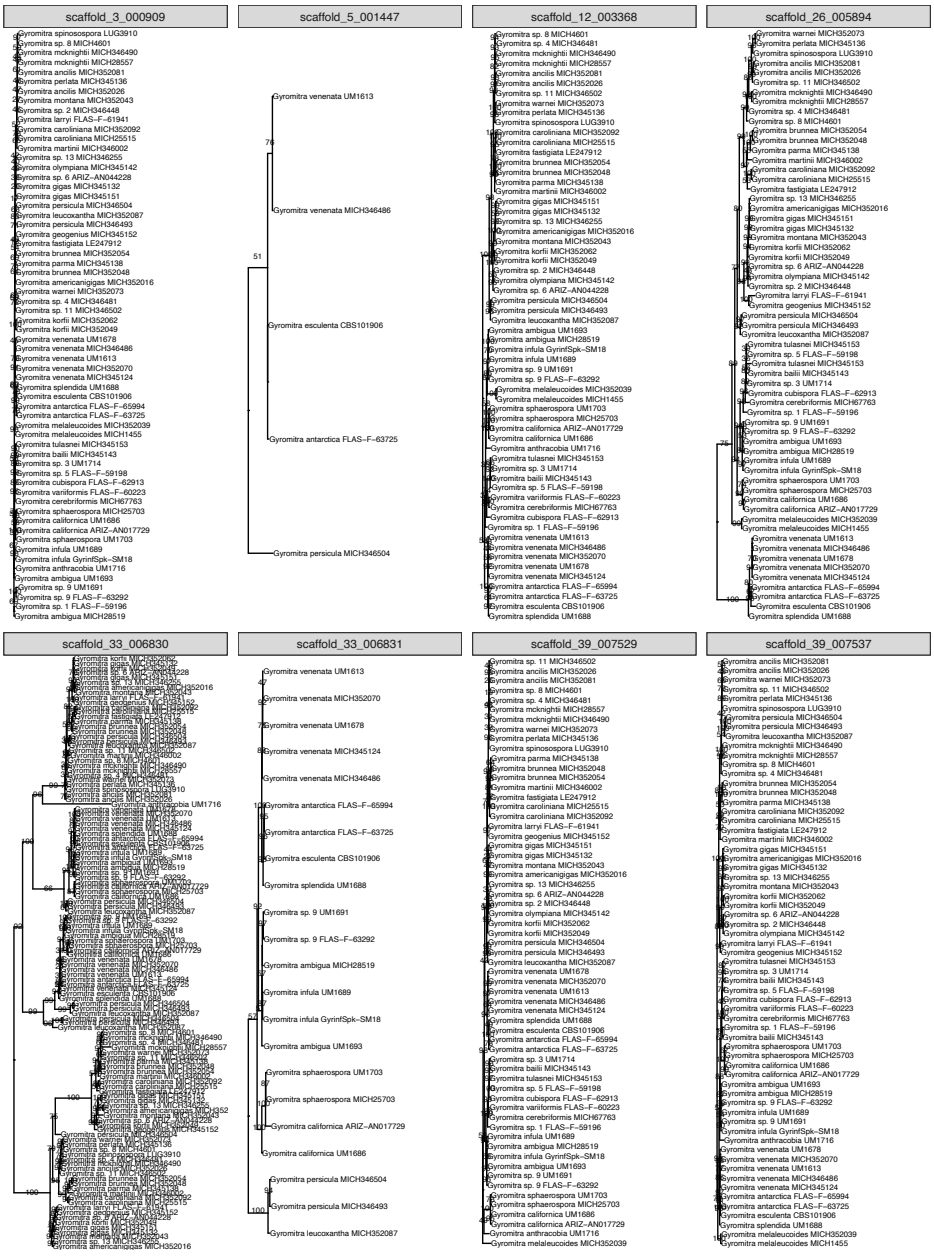
Supplemental Figure 4.6. Gene synteny map of putative gyromitrin biosynthesis gene cluster 2, conserved region B. The putative gyromitrin BGC 2 genes are indicated by a box. Red highlights indicate gyromitrin-producing clades.



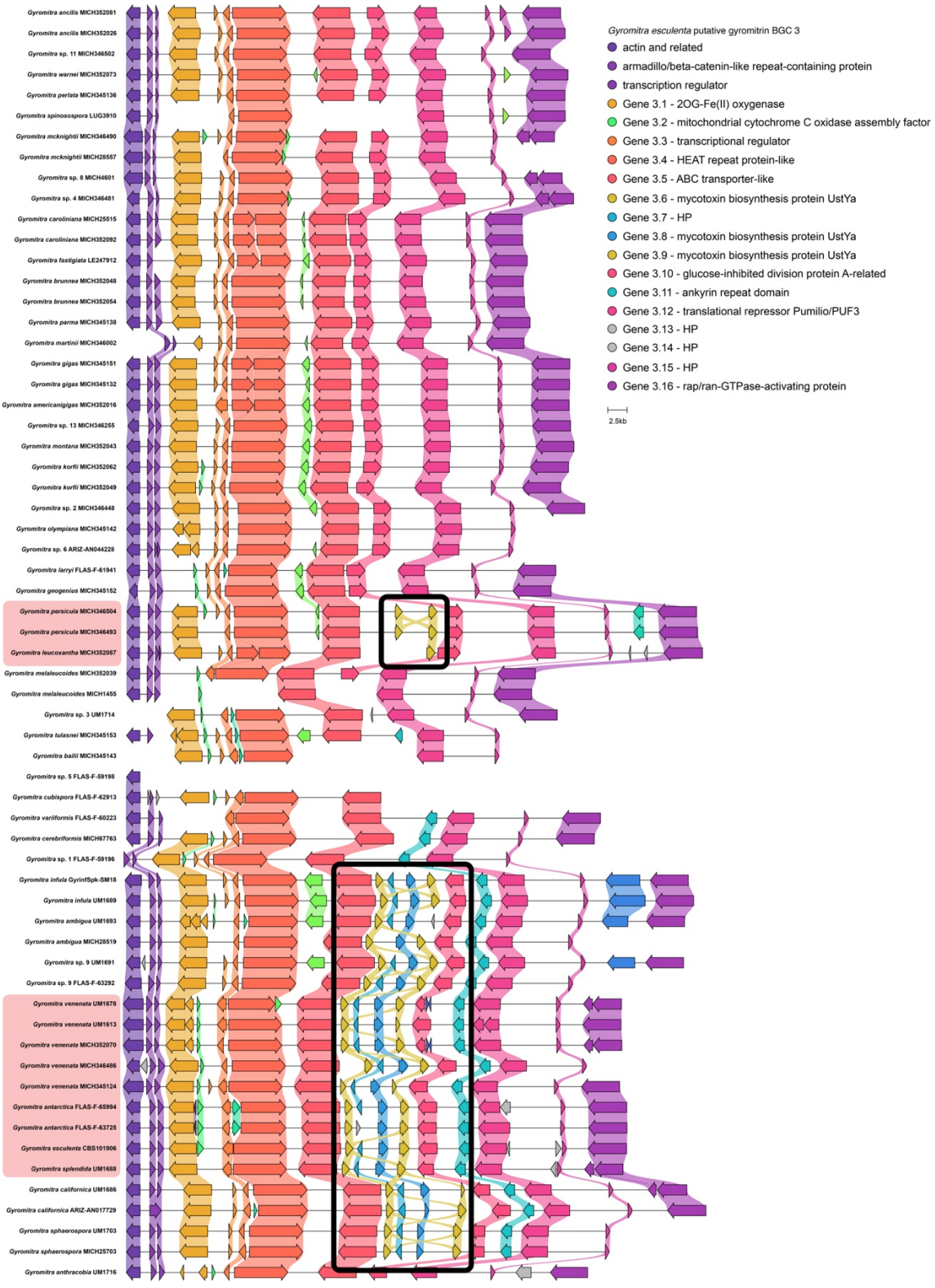
Supplemental Figure 4.7. Gene synteny map of putative gyromitrin biosynthesis gene cluster 2, conserved region C. The putative gyromitrin BGC 2 genes are indicated by a box. Red highlights indicate gyromitrin-producing clades.



Supplemental Figure 4.8. Gene synteny map of conserved regions A, B, and C of putative gyromitrin biosynthesis gene cluster 2. The putative gyromitrin BGC 2 genes are indicated by boxes. Red highlights indicate gyromitrin-producing clades.



Supplemental Figure 4.9. Phylogenetic trees of orthogroups corresponding to *Gyromitra esculenta* CBS101906 biosynthesis gene clusters in which the gyromitrin-producing subgenera were more closely related than expected. The trees were inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support.



Supplemental Figure 4.10. Gene synteny map of putative gyromitrin biosynthesis gene cluster 3. The UstYa-family genes are inside the boxes (gold and blue genes). Red highlights indicate gyromitrin-producing clades.



Supplemental Figure 4.11. Phylogenetic trees of orthogroups corresponding to putative gyromitrin biosynthesis gene cluster 3 (see main text for details) inferred with IQ-TREE2. Branch values show the ultrafast bootstrap support. The phylogenies for the core UstYa-family genes 3.6 and 3.9 are displayed in Figure S9 (scaffold 33 006830) and the main text (Figure 4.6). The remaining missing orthogroup phylogenies were not generated since they were absent from *Gyromitra* subg. *Leucoxantha* samples.

Supplemental Table 4.1. Orthogroups whose phylogenies had gyromitrin-producing taxa forming a monophyletic clade and contained *Gyromitra esculenta* CBS101906. The location refers to the position in the Joint Genome Institute MycoCosm genome browser for *G. esculenta*, and the protein ID is a URL to the MycoCosm page for that protein.

Gene	Number	Orthogroup	Gyromitrin-taxa conserved	Predicted function	Location	Protein ID
1	548	OG0009750	YES	HP	scaffold_2:330662-331510	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=612765
2	1447	OG0012989	NO	polysaccharide deacetylase	scaffold_13:472556-474762	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=578455
3	2006	OG0000516	NO	HP	scaffold_7:12759-13824	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=521052
4	2371	OG0011566	NO	HP	scaffold_8:549882-550812	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=360096
5	2487	OG0000516	NO	HP	scaffold_7:12999-13824	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=510174
6	3272	OG0009706	YES	metal-dependent amidase/aminoacylase/carboxypeptidase	scaffold_12:237664-238997	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=529821
7	3275	OG0010417	YES	oxoglutarate/iron-dependent dioxygenase	scaffold_12:245914-247434	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=44179
8	3276	OG0009479	YES	didomain aminoacyl-tRNA synthetase/GNAT domain	scaffold_12:248482-251259	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=546118
9	3702	OG0011018	NO	GDP-L-fucose synthetase	scaffold_14:242503-243438	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=578595
10	4336	OG0011001	NO	HP similar to nuclear recepter co-repressor 2	scaffold_17:108647-109699	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=549985
11	6117	OG0010298	YES	tetrapyrrole methylase	scaffold_27:516336-517699	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=514041
12	6737	OG0010418	NO	chitin-binding type 1; CAZy CBM18; lectin B	scaffold_32:432805-434008	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=502296
13	7119	OG0013180	NO	amino acid transporter	scaffold_35:450626-452637	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=514975
14	7187	OG0012978	NO	HP	not predicted by JGI	not predicted by JGI
15	7391	OG0008100	YES	deoxyribose-phosphate aldolase	scaffold_38:62961-63758	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=531962

16	7437	OG0011004	NO	transcriptional regulator	scaffold_38:229280-235244	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=525362
17	7588	OG0010599	YES	HP	scaffold_39:366913-367200	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=515409
18	7728	OG0008859	YES	chromate transporter	scaffold_40:424441-426253	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=503637
19	7911	OG0012957	NO	HP	scaffold_42:222353-222757	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=618255
20	8401	OG0011555	NO	5'-3' exoribonuclease 2/3	scaffold_47:55959-59453	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=434718
21	8551	OG0007298	NO	HP	scaffold_49:46377-47954	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=268796
22	8658	OG0010300	NO	HP	scaffold_50:142073-143083	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=281142
23	8866	OG0010790	NO	Dbl homology domain	scaffold_52:250609-257032	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=526429
24	9523	OG0008732	YES	HP	scaffold_60:196808-198424	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=619399
25	10769	OG0009041	NO	myosin-7-like	scaffold_96:81320-84066	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=518544
26	11000	OG0009710	YES	HP	scaffold_119:10172-11920	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=574706
27	11002	OG0009979	YES	HP, HET domian-containing protein	scaffold_119:22677-23759	https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Gyresc1&id=438843

Chapter 5 Conclusion

I set out to expand our understanding of the gyromitrin phenotype in lorchel mushrooms and elucidate the genetics and evolution of the mycotoxin. Along the way, I hoped to augment our general knowledge of lorchels —controversially edible mushroom-forming fungi with significant cultural importance. In Chapter 2, I developed a new analytical technique for the detection of gyromitrin and surveyed a large diversity of species from the lorchel family, *Discinaceae*. I found that gyromitrin has a more limited distribution than expected, being present in subgenera *Gyromitra* and *Leucoxantha* but absent from previously suspected groups. In Chapter 3, I sequenced numerous *Discinaceae* genomes to infer a robust phylogenomic tree for the family. This allowed me to revise the lorchel taxonomy and explore fundamental aspects of their biology, such as their reproduction and ecology. The phylogenomic tree showed that gyromitrin has a discontinuous distribution, being present in two distantly related groups. Finally, in Chapter 4, I took this genomic dataset and conducted bioinformatics analyses to identity putative gyromitrin biosynthesis gene clusters. My top candidate gene cluster showed a prominent pattern of phylogenetic incongruence suggestive of horizontal gene transfer. The rate and mechanisms of horizontal gene transfer in fungi are largely unknown, but the close taxonomic affinity of the lorchels suggests interspecific hybridization could have played a role in a putative transfer event (Gonçalves et al., 2024; He et al., 2020).

Validation of the gyromitrin genes appears at arm's reach because of the work undertaken in this thesis, but further research is required to confirm the hypotheses set out in

Chapter 4. Future studies could employ labelled isotopes, gene knockouts, and heterologous expression to validate the gyromitrin biosynthesis gene cluster (Cano-Prieto et al., 2024). Labelled isotope studies use chemicals with one or more atoms replaced with a heavy isotope, which can be detected with mass spectrometry. By feeding mycelium isotopically labelled amino acids and other building blocks, the substrates involved in gyromitrin biosynthesis can be identified based on their incorporation into the secondary metabolite. This is one method for indirectly testing whether our candidate gene clusters are associated with gyromitrin production. For example, if the tyrosyl-tRNA synthetase is responsible for N–N bond formation, one would expect the isotopic signature of a nitrogen-labelled tyrosine to show up in gyromitrin. Labelled isotope studies could lead to the development of new genetic hypotheses while clarifying the chemicals involved in its biosynthetic pathway.

Gene knockout experiments are a direct means to validate the gyromitrin biosynthesis genes (Keller, 2019). CRISPR-based genetic engineering of filamentous fungi brings this dream closer to reality, but the development of genetic transformation protocols in new organisms is still challenging because of species-specific nuances in cell wall composition and physiology (Schuster and Kahmann, 2019). Another challenge for *Gyromitra* species is that conidia have not been produced in lab conditions to date, and the mycelium is relatively slow growing. This is an obstacle to the procurement of large numbers of transformable units, which are required because transformation success rates are typically low.

Heterologous expression involves the incorporation of genes from a target organism into a donor organism. *Aspergillus* spp. are model molds with well-developed protocols for heterologous expression (Lubertozzi and Keasling, 2009). However, the more unrelated two species are the more difficult heterologous expression can be. Challenges arise if the two

organisms have different codon usage, the biosynthesis pathway incorporates substrates from the primary metabolism pool that are absent in the donor, or if resistance to the effects of the secondary metabolite is not encoded in the expressed genes (Alberti et al., 2017). While these challenges are formidable, they are not insurmountable. Throughout my thesis, I have made significant progress in the protoplasting of *Gyromitra* mycelium, which involves the dissolution of the fungal cell wall to facilitate the insertion of foreign material into the cell. The development of a complete transformation protocol for *Gyromitra* will allow future researchers to validate the gyromitrin genes.

In general, the genetics of mushroom-forming fungi like *Gyromitra* spp. have received limited attention. As seen in this thesis, nonmodel fungi are challenging to study, but their underexplored biology also makes them fascinating research subjects with the potential to reshape biological paradigms regarding secondary metabolism (Fricke et al., 2017; Rokas et al., 2020). Specifically, the number of secondary metabolites produced by noncanonical biosynthesis gene clusters (*i.e.,* gene clusters that lack a stereotypical backbone synthase gene) is unknown. As evidenced by the bioinformatics discovery of the psilocybin biosynthesis gene cluster (Reynolds et al. 2018), and perhaps also the work detailed here for gyromitrin, comparative genomics could be a particularly valuable approach for genotype-phenotype linkage for noncanonical biosynthesis gene clusters, whose contribution to fungal chemodiversity may be underappreciated.

References

Alberti, F., Foster, G.D., Bailey, A.M., 2017. Natural products from filamentous fungi and production by heterologous expression. Appl Microbiol Biotechnol 101, 493–500. https://doi.org/10.1007/s00253-016-8034-2

Cano-Prieto, C., Undabarrena, A., de Carvalho, A.C., Keasling, J.D., Cruz-Morales, P., 2024. Triumphs and challenges of natural product discovery in the postgenomic era. Annu. Rev. Biochem. 93, 20.1-20.35. https://doi.org/10.1146/annurev-biochem-032620-104731

Gonçalves, C., Hittinger, C.T., Rokas, A., 2024. Horizontal gene transfer in fungi and its ecological importance, in: Hsueh, Y.-P., Blackwell, M. (Eds.), Fungal Associations, The Mycota. Springer International Publishing, Cham, pp. 59–81. https://doi.org/10.1007/978-3-031-41648-4_3

He, P., Yu, M., Wang, K., Cai, Y., Li, B., Liu, W., 2020. Interspecific hybridization between cultivated morels *Morchella importuna* and *Morchella sextelata* by PEG-induced double inactivated protoplast fusion. World J Microbiol Biotechnol 36, 58. https://doi.org/10.1007/s11274-020-02835-0

Keller, N.P., 2019. Fungal secondary metabolism: regulation, function and drug discovery. Nat. Rev. Microbiol. 17, 167–180. https://doi.org/10.1038/s41579-018-0121-1

Lubertozzi, D., Keasling, J.D., 2009. Developing *Aspergillus* as a host for heterologous expression. Biotechnology Advances 27, 53–75. https://doi.org/10.1016/j.biotechadv.2008.09.001

Reynolds, H.T., Vijayakumar, V., Gluck-Thaler, E., Korotkin, H.B., Matheny, P.B., Slot, J.C., 2018. Horizontal gene cluster transfer increased hallucinogenic mushroom diversity. Evol Lett 2, 88–101. https://doi.org/10.1002/evl3.42

Schuster, M., Kahmann, R., 2019. CRISPR-Cas9 genome editing approaches in filamentous fungi and oomycetes. Fungal Genetics and Biology 130, 43–53. https://doi.org/10.1016/j.fgb.2019.04.016